Regulation and biological function of the RNA ligase RtcB in *Escherichia coli*

Georgia Mary Hann

Centre for Bacterial Cell Biology

Institute for Cell and Molecular Biosciences

Thesis submitted for the degree Doctor of Philosophy

August 2017

Abstract

All organisms contain RNA ligases for the modification or repair of RNA molecules. The RtcB family have a novel 3′-5′ RNA ligase activity and are conserved in all three domains of life. Metazoan homologues of RtcB have essential roles in tRNA maturation and the unfolded protein response. In *E. coli*, expression of RtcB from the σ^{54}-dependent *rtcBA* operon is regulated by transcriptional activator RtcR. Neither the signals that induce *rtcB* expression nor the natural substrates of RtcB in *E. coli* are known. The objective of this project is to understand the regulation of *rtcB* expression by RtcR and the biological function of RtcB in *E. coli*.

In this study, we confirm the roles of RtcR and σ^{54} in *rtcB* transcription in *E. coli*. Using a series of truncated *rtcBA* promoter fusions with the reporter gene *lacZ*, we have identified an inverted repeat that we hypothesise is the RtcR binding site in the promoter of *rtcB*. Mutation of the inverted repeat supports its involvement in RtcR binding. Using a similar approach, we have identified a possible σ^{70} promoter responsible for transcription of *rtcR* and provide evidence of negative autoregulation of *rtcR*. Moreover, we also demonstrate for the first time the requirement of IHF for RtcR-dependent transcription of *rtcB* and confirm the putative binding site of IHF in the *rtcBA* promoter *in vivo* and *in vitro*.

Using a combination of genetic screening with a chromosomal *rtcBA-lacZ* fusion and quantitative PCR analysis, we have uncovered several conditions where transcription of *rtcB* is activated in *E. coli*. Expression of Rof, an inhibitor of Rho transcription termination, induced *rtcB* transcription in an RtcR-dependent manner, indicating a possible link between RtcB function and disruption of transcription termination. Furthermore, treatment of cells with chloramphenicol and induction of endoribonuclease MazF increased *rtcB* transcription, suggesting that RtcB may respond to the translational status of the cell. We have used these conditions to perform CRAC (crosslinking and analysis of cDNA) of RtcB, in order to identify substrate RNAs that interact with RtcB *in vivo*. Many interactions between RtcB and RNAs were identified including rRNAs, tRNAs, sRNAs and mRNAs and several candidates were further analysed by qPCR.

Taken together, the data suggests that *rtcB* has features typical of a σ^{54}-dependent promoter and that RtcB expression can be activated in a variety of stress conditions, thereby allowing interaction with a range of cellular RNAs in *E. coli*. Nevertheless, the specific signal(s) detected by RtcR and the biological function of RtcB in *E. coli* remain elusive.

Acknowledgements

I would like to thank my supervisor Kenn Gerdes for the opportunity to work on this project and for help and guidance throughout. I also thank Kristoffer Winther for useful scientific discussions and help with experimental procedures. I am grateful to Kathryn Turnbull for many useful discussions, scientific and otherwise. I also acknowledge and thank past and present members of the Gerdes group for their support, technical expertise and suggestions.

I am grateful to members of the Centre for Bacterial Cell Biology, Newcastle University and the Centre for Bacterial Stress Response and Persistence, University of Copenhagen for providing a welcoming and stimulating work environment. I thank Michael Sørensen and his group (University of Copenhagen) for an interesting collaboration.

Finally, I thank Mick Hollier, Susan Hann and David Hann for their constant love, encouragement and moral support.

Contents

List of figures

List of tables

Abbreviations

A	Adenosine
AAA+	ATPase associated with various cellular activities
Amp	Ampicillin
AMP	Adenosine monophosphate
aSD	Anti-Shine-Dalgarno
ATP	Adenosine triphosphate
bEBP	Bacterial enhancer binding protein
bp	Base pairs(s)
C	Cytidine
CARF	CRISPR-associated Rossmann fold
CFU	Colony forming units
Cml	Chloramphenicol
CRAC	Crosslinking and analysis of cDNA
DNA	Deoxyribonucleic acid
ds	Double-stranded
EBP	Enhancer binding protein
ECM	Extracellular matrix
EMSA	Electrophoretic mobility shift assay
ER	Endoplasmic reticulum
G	Guanosine
GMP	Guanosine monophosphate
GTP	Guanosine triphosphate
HTH	Helix-turn-helix
IHF	Integration host factor
IPTG	Isopropyl-β-D-1-thiogalactopyranoside
IR	Inverted repeat
Kan	Kanamycin
Kb	Kilobases
LB	Lysogeny broth
MCS	Multiple cloning site
mRNA	Messenger RNA
NA	Nutrient agar

Nt	Nucleotide
NTP	Nucleoside triphosphate
OD	Optical density
ON	Over night
ORF	Open reading frame
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PNPase	Polynucleotide phosphorase
qPCR	Quantitative PCR
RNA	Ribonucleic acid
RNAi	RNA interference
RNAL	RNA ligase
RNAP	RNA polymerase
rRNA	Ribosomal RNA
RT	Room temperature
RT-PCR	Reverse-transcription PCR
Rtc	RNA terminal cyclase
SD	Shine-Dalgarno
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
sRNA	Small RNA
siRNA	Small interfering RNA
σ	Sigma factor
T	Thymidine
TA	Toxin-antitoxin
tRNA	Transfer RNA
ts	Temperature sensitive
U	Uridine
UAS	Upstream activating sequence
UPR	Unfolded protein response
wHTH	Winged helix-turn-helix
WT	Wild-type
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

Chapter 1 Introduction

1.1 Overview

The *rtcB* gene is highly conserved with homologues in eukaryotes, archaea and bacteria (Englert et al., 2011; Popow et al., 2012). In *Escherichia coli* K-12, *rtcB* is part of the *rtcBA* operon and is expressed from a promoter regulated by the alternative sigma factor σ^{54} (Genschik et al., 1998). Initiation of transcription from a σ^{54}-dependent promoter requires a transcriptional activator, which is itself activated by the detection of metabolic or environmental signals (Bush and Dixon, 2012). Accordingly, transcription from the *rtcBA* promoter requires the product of neighbouring gene *rtcR* (Genschik et al., 1998). RtcR and other transcriptional activators of σ^{54}-dependent transcription are known as bacterial enhancer binding proteins (bEBPs), owing to functional similarity with eukaryotic enhancer binding proteins (EBPs) of RNA polymerase (RNAP) II (Ghosh et al., 2010). bEBPs are a subgroup of the AAA+ (ATPase associated with various cellular activities) family of proteins and initiate transcription from σ^{54}–dependent promoters by binding to upstream activating sequences (UASs) and interacting with the RNAP holoenzyme to facilitate open promoter complex formation (Bush and Dixon, 2012). A signal sensory component of bEBPs restricts this activity to appropriate cellular conditions. Despite extensive knowledge about the process of σ^{54}–dependent transcription in general, little is known regarding the regulation of σ^{54}–dependent *rtcBA* transcription by RtcR.

The protein product of the *rtcB* gene was demonstrated to have activity as an RNA ligase (RNAL) by Tanaka and Shuman (2011). RNALs are ubiquitous enzymes involved in modification or repair of RNA molecules by joining breaks in the backbone of RNA to create phosphodiester bonds (Burroughs and Aravind, 2016; Pascal, 2008; Popow et al., 2012). The RtcB family of RNALs catalyse RNA ligation by a novel mechanism (Chakravarty and Shuman, 2012; Chakravarty et al., 2012; Tanaka et al., 2011a), which, when combined with their high degree of conservation, suggests they have an important role in RNA modification. Indeed, recent studies have identified essential roles for metazoan RtcB in transfer (t)RNA maturation, ligating tRNA halves following endonuclease-mediated intron excision (Popow et al., 2011), and in the splicing of *XBP-1* messenger (m)RNA after intron removal by IRE1 endonuclease, generating an active transcription factor in the unfolded protein response

(UPR) (Jurkin et al., 2014; Kosmaczewski et al., 2014; Lu et al., 2014). However, the extent of its biological function is unclear, particularly in bacteria.

This chapter provides an overview of the current state of knowledge regarding the regulation, mechanism of action and function of RtcB. Section 1.2 gives a general introduction to bacterial transcription and the role of σ factors in promoter specification. In Section 1.3, the characteristics of σ^{54}–dependent transcription and the σ^{54} regulon are discussed, while Section 1.4 describes the genetic organisation and regulatory components of the *rtcBA* operon in *E. coli*. The classification and mechanism of action of RNALs are outlined in Section 1.5, with particular focus on the RtcB family. Section 1.6 concerns the conservation and known physiological functions of RtcB and the function of genes associated with bacterial *rtcB*. Finally, the aims of the study are outlined in Section 1.7.

1.2 Promoter recognition in bacterial transcription

1.2.1 Bacterial transcription and the RNA polymerase holoenzyme

Gene transcription is a fundamental and highly regulated process that defines the ability of cells to grow in and adapt to their surroundings. In all free living organisms, it is executed by the DNA-dependent multi-subunit enzyme RNAP, which shares structural and functional similarity in all domains of life despite differences in complexity (Sekine et al., 2012). There are three main phases of transcription: initiation, which involves promoter recognition and binding, polymerisation of the first few ribonucleotides and promoter clearance by RNAP; elongation, whereby ribonucleotides are sequentially added to the growing RNA transcript according to the sequence of the DNA template; and termination, where RNAP reaches the end of the gene and dissociates from the DNA template, releasing the RNA (Weaver, 2012).

Whereas the majority of eukaryotes have three RNAP enzymes, specialising in transcription of ribosomal (r)RNA, mRNA or tRNA respectively, bacterial organisms utilise a single RNAP to achieve transcription of all classes of RNA (Weaver, 2012). The core bacterial RNAP consists of five subunits, β, β′, two α and ω, which together are competent for RNA synthesis from a DNA template alone. However, in order to mediate initiation of transcription, an additional subunit is required, known as the σ or ‘specificity’ factor, which associates with core RNAP to form the RNAP holoenzyme (Burgess and Travers, 1969; Ghosh et al., 2010).

1.2.2 The σ factor of RNA polymerase holoenzyme

An essential regulatory component of bacterial transcription, the σ factor directs and initiates promoter-specific transcription by associating with core RNAP and recognising conserved motifs upstream of the transcription start site to enable engagement with the promoter. In addition, the σ factor mediates transition from a closed promoter complex with RNAP to an open promoter complex, initiating strand separation of duplex DNA. The roles of σ factors in transcription initiation are reviewed by Feklistov et al. (2014) and Saecker et al. (2011). The σ factor dissociates from core RNAP in a stochastic manner, facilitated by competition for interaction with RNAP from the elongating RNA transcript, usually during the early elongation phase of transcription (Shimamoto et al., 1986; Yang and Lewis, 2010). Following its release from RNAP, it re-joins the cellular pool of σ factors and is free to compete for binding to other core RNAPs (Mooney et al., 2005; Travers and Burgess, 1969). This recycling process is known as the σ cycle, as depicted in Figure 1.

There are usually multiple σ factors within an organism, each recognising distinct promoter sequences to mediate promoter selectivity and differential gene transcription by RNAP. A primary σ factor is responsible for the majority of transcription initiation, including general 'housekeeping' genes necessary for growth, while a variable number of alternative σ factors direct transcription of specific subsets of stress-response genes (Ghosh et al., 2010; Mooney et al., 2005). The differing promoter specificities of σ factors, in combination with the processes regulating their availability in the σ factor pool and thereby driving competition for RNAP binding in the σ cycle, allow bacteria to adjust the transcription profile of the cell in response to changes in environmental conditions. The regulatory mechanisms that govern σ factor interaction with RNAP are reviewed by Osterberg et al. (2011).

There are two classes of σ factor, grouped according to primary sequence similarity, consensus sequence of the DNA elements they recognise and their mode of activation. The majority belong to the σ^{70} class, named for the primary σ factor of *E. coli* and are further sub-divided according to their possession of the four domains conserved in housekeeping σ factors (Wosten, 1998). In contrast to the diversity of the σ^{70} class, σ^{54} is the sole member of the σ^{54} class but is widely dispersed among bacteria, with approximately 60% of bacteria encoding at least one σ^{54} gene (Wigneshweraraj et al., 2008). As discussed in Section 1.3, σ^{54} has a distinct mechanism of transcription initiation to σ factors of the σ^{70} class, recognising promoter sequences that differ in position, sequence and spacing to σ^{70} and require additional regulatory components.

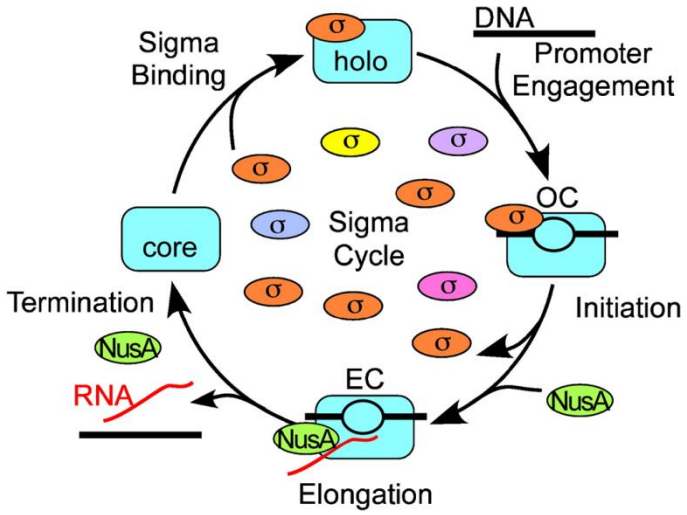


Figure 1 The σ cycle.

Sigma (σ) factors compete for binding to core RNA polymerase. Once the RNAP holoenzyme is formed, the σ factor mediates promoter engagement and formation of an open promoter complex (OC). The σ factor is released stochastically from the elongating complex (EC) and returns to the σ factor pool. RNAP interacts with other regulatory factors during the elongation and termination phases of transcription, such as NusA, which is involved in transcription termination and antitermination. Taken from Mooney et al. (2005).

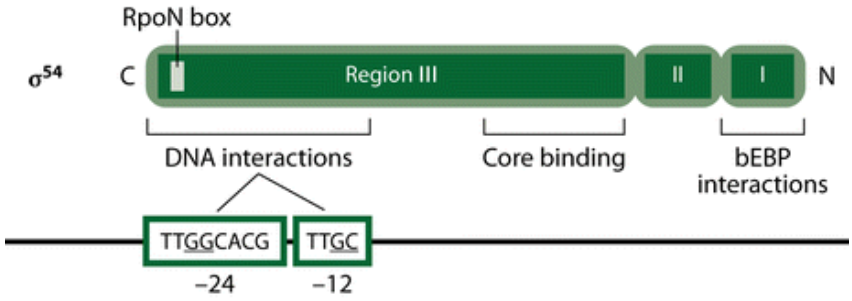


Figure 2 Regions and functions of σ^{54}.

σ^{54} is divided into three regions. Region I, at the N-terminal, mediates interactions with transcriptional activators of σ^{54}-dependent transcription, known as bEBPs. It also contacts the -12 promoter DNA element and has interactions with core RNA polymerase (not shown). Region II links regions I and III. C-terminal region III contains the main core RNA polymerase-binding domains and is also responsible for interaction with promoter elements at -24 and -12, as indicated in the green boxes where the highly conserved GG and GC motifs are underlined. Taken from Osterberg et al. (2011).

E. coli has seven σ factors: primary factor σ^{70} and six alternatives. Five of these, σ^{38}, σ^{32}, σ^{28}, σ^{24} and σ^{19}, belong to the σ^{70} class and transcribe genes involved in the stationary phase of growth, response to heat shock, motility, response to stress on membrane proteins and ferric citrate transport respectively (Ishihama, 2000). Alternative factor σ^{54} transcribes genes in response to nitrogen limitation and other stresses, as discussed in the next section. As the housekeeping σ factor, σ^{70} is the most abundant with the strongest binding affinity for core RNAP, occupying an estimated 78% of holoenzymes during exponential growth, while σ^{54}-associated RNAP constitutes 8% (Ishihama, 2000; Maeda et al., 2000).

1.3 Sigma 54-dependent transcription

1.3.1 Mode of transcription initiation by σ^{54}-associated RNAP

Transcription by the σ^{54}-associated RNAP holoenzyme has been described as a second paradigm of bacterial transcription, differing substantially from that directed by σ^{70}-RNAP. As illustrated in Figure 2, σ^{54} can be divided into three regions, which do not share sequence or structural homology to the σ^{70} class and make distinct contacts with both core RNAP and promoter DNA (Buck and Cannon, 1992; Buck et al., 2000; Wong et al., 1994; Zhang et al., 2016).

The major interactions between σ^{54} and core RNAP are mediated by the core binding domain of region III, which is the most conserved domain of σ^{54} orthologues. Mutations (Tintut and Gralla, 1995) and deletions (Wong et al., 1994) of residues in the core binding domain were found to disrupt the formation of the holoenzyme, while this domain alone was sufficient for high affinity RNAP binding (Gallegos and Buck, 1999). Moreover, the structure of the σ^{54}-RNAP holoenzyme (Figure 3) revealed extensive contacts between the core binding domain and moieties of the α, β and β′ subunits (Yang et al., 2015). Despite the lack of sequence and structural similarity between σ^{54} and σ^{70}, comparison of the holoenzyme structures indicates they contact similar regions of core RNAP, albeit with different functional domains (Yang et al., 2015; Zhang et al., 2016).

σ^{54}-regulated promoters have several features that distinguish them from those regulated by the σ^{70} family and define the process of σ^{54}-dependent transcription. In contrast to σ^{70} recognition sites, which are positioned at -10 and -35 relative to the +1 transcription start site with the exact sequence and spacing varying for each member, σ^{54} recognises -24 and -12 promoter elements (Kustu et al., 1989; Zhao et al., 2010). An extended σ^{54} consensus

sequence was established by calculating the relative nucleotide frequency at each position of 186 σ^{54}–regulated promoters from over 40 species (Barrios et al., 1999). The resulting consensus sequence of mrNrYTGGCACG-N_4-TTGCWNNw (where uppercase and lowercase indicate highly and weakly conserved nucleotides respectively; R represents purines (A/G), Y represents pyrimidines (C/T); W is A/T; M is C/G and N is any nucleotide) spans from -31 to -8, centred around highly conserved GG and GC motifs at -24 and -12 respectively.

Binding of σ-associated RNAP to promoter elements to form a closed promoter complex is the first stage of transcription initiation. As indicated in Figure 2 and Figure 3C, region III of σ^{54} contains the major determinants for promoter DNA binding (Guo and Gralla, 1997; Wong et al., 1994). A cluster of ten residues at the C-terminal end of region III termed the RpoN box, the most conserved feature of σ^{54}, are responsible for binding to the -24 sequence, as demonstrated by mutagenesis (Taylor et al., 1996) and cleavage of RpoN box-adjacent DNA by a proximity-based cleavage reagent conjugated to σ^{54} (Burrows et al., 2003). Structural information revealed the presence of a helix-turn-helix (HTH) motif, a signature of DNA binding proteins, of which residues of the RpoN box form the DNA recognition helix (Doucleff et al., 2005; Doucleff et al., 2007; Yang et al., 2015). In contrast, contact with the -12 element is more complex, involving region I as well as a second HTH of region III (Burrows et al., 2003; Wong et al., 1994).

Unlike σ^{70} family-associated RNAP, which spontaneously isomerise DNA to form an open promoter complex, σ^{54}-RNAP-bound -24/-12 elements remain closed and transcriptionally silent as a result of the inhibitory function of region I of σ^{54}. Recently, Yang et al. (2015) showed that, according to the structure of σ^{54}-RNAP in complex with promoter DNA, region I interacts with the HTH of region III to form a structural unit that is positioned close to the -12 promoter element and impedes DNA entry into the active site of RNAP. This is consistent with biochemical studies that identified residues of regions I and III responsible for -12 binding (Cannon et al., 1994; Chaney and Buck, 1999; Wang and Gralla, 1996) and demonstrated that σ^{54} inhibited association of RNAP with pre-melted DNA but deletion of region I removed this obstacle (Cannon et al., 1999). The interaction of region I and III of σ^{54} with the -12 element has been shown to distort DNA at the base pair immediately downstream, contributing to the formation of a fork junction-like structure called the regulatory centre (Bose et al., 2008; Burrows et al., 2003; Morris et al., 1994; Yang et al., 2015). The regulatory centre is located at the point of nucleation of DNA melting according to the structure of the σ^{70}-RNAP holoenzyme, therefore preventing DNA isomerisation,

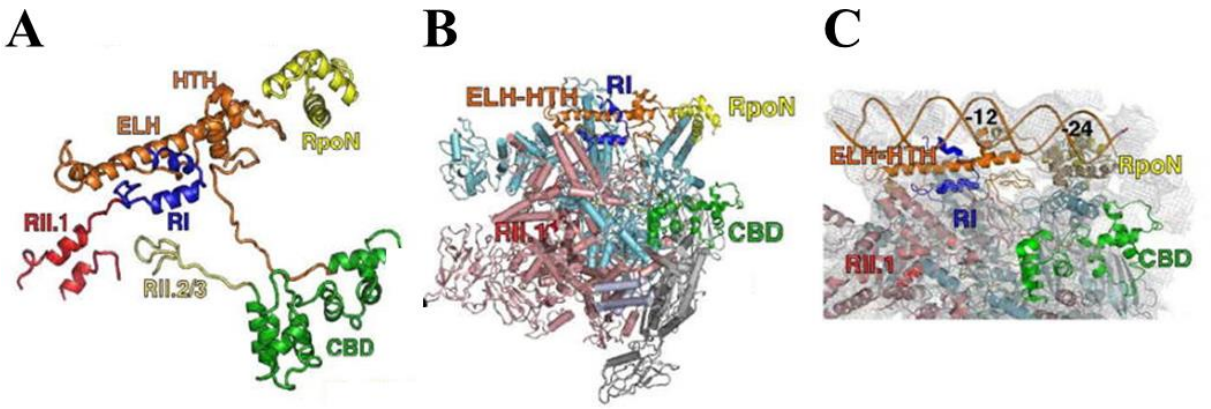


Figure 3 Structures of the σ^{54}.

(A) Structure of σ^{54}. (B) Structure of σ^{54}-associated RNAP holoenzyme. (C) Electron density map and structural model of the σ^{54}- RNAP-promoter DNA complex. In all cases, the structures show region I (RI; purple), region II (divided into RII.1, RII.2 and RII.3; red and yellow) and region III, which is divided into the core binding domain (CBD; green), the extra-long α helix and helix-turn-helix (ELH-HTH; orange), and the RpoN box (RpoN; lime). The subunits of core RNAP are shown as follows: β (blue), β′ (pink), α1 (light grey), α2 (dark grey) and ω (lilac). The -24 and -12 promoter DNA elements are also indicated. Taken from Yang et al. (2015).

formation of the transcription bubble and entry of promoter DNA into the active site of σ^{54}-RNAP (Zhang et al., 2016; Zuo and Steitz, 2015).

In order to overcome this barrier to transcription initiation, σ^{54}-dependent promoters require a transcriptional activator, known as a bEBP, which binds to a UAS and interacts with the σ^{54}-RNAP holoenzyme-promoter DNA complex (Buck et al., 1986; Sasse-Dwight and Gralla, 1988). As the inhibitory determinant of σ^{54}, region I is the main bEBP interaction domain. Structural data suggests region I is positioned favourably for contact with bEBPs and is stabilised upon interaction (Siegel and Wemmer, 2016; Yang et al., 2015). Furthermore, binding to region I has been demonstrated for several members of the bEBP family, including PspF, NifA and NtrC1 (Bordes et al., 2003; Chaney et al., 2001; Siegel and Wemmer, 2016). bEBPs have a highly conserved ATPase domain, which in the adenosine triphosphate (ATP)-bound state, interacts with region I and converts the chemical energy from ATP hydrolysis to mechanical energy to induce conformational change of σ^{54} at the regulatory centre (Bordes et al., 2004; Bose et al., 2008; Burrows et al., 2008; Chaney et al., 2001). This enables promoter melting at the regulatory centre and eliminates the obstruction of promoter DNA from the RNAP catalytic site, resulting in open complex formation.

Figure 4 illustrates the sequence of events during initiation of σ^{54}-dependent transcription.

1.3.2 Bacterial enhancer binding proteins

bEBPs are essential for initiation of transcription from σ^{54}-dependent promoters. The majority of bEBPs have three domains: the regulatory domain at the N-terminal of the protein, the central and highly conserved ATPase domain and the C-terminal DNA binding domain (reviewed by Bush and Dixon (2012) and Shingler (2010)).

1.3.2.1 ATPase domain of bEBPs

The central ATPase domain consists of seven highly conserved motifs, many of which are hallmarks of the AAA+ superfamily of proteins. AAA+ proteins function in a diverse range of cellular processes but all involve the binding and hydrolysis of ATP, converting the energy produced into mechanical force to remodel their substrates (Snider et al., 2008). In the case of bEBPs, the ATPase domain mediates interaction with σ^{54} and drives conformational changes to promote open promoter complex formation by the σ^{54}-RNAP holoenzyme.

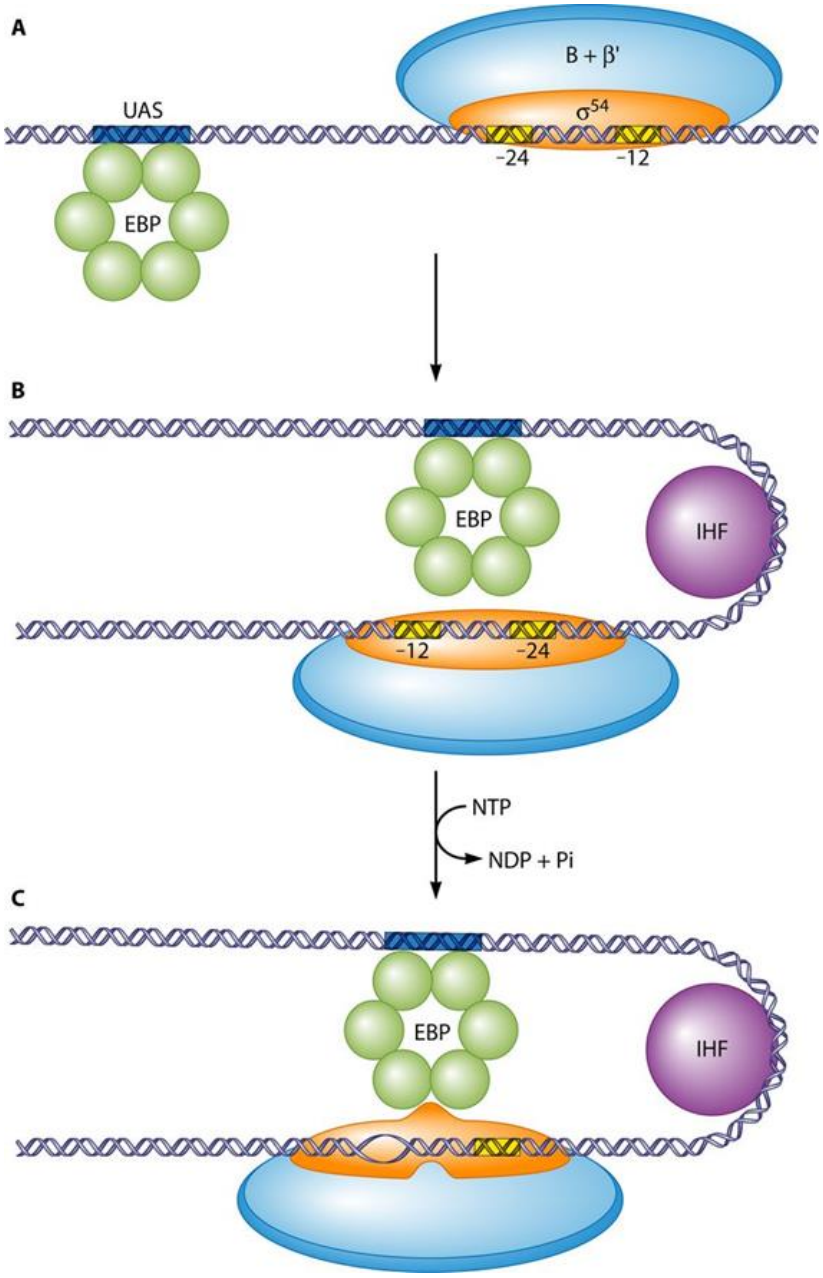


Figure 4 Initiation of transcription at σ^{54}-dependent promoters.

(A) σ^{54}–associated RNAP is directed to promoters with GG and GC motifs at -24 and -12 respectively, relative to the +1 transcription start site. The σ^{54}–RNAP holoenzyme is unable to spontaneously form an open complex with promoter DNA and requires a transcriptional activator (enhancer binding protein; EBP). The activator binds as a dimer to an upstream activating sequence (UAS) and oligomerises (commonly as a hexamer as shown) upon activation by a specific signal. (B) Interaction between the EBP and σ^{54}–RNAP is facilitated by DNA bending, usually mediated by integration host factor (IHF). This loops out the intervening sequence between UASs and -12/-24 elements and mediates correct interfacing between σ^{54}–RNAP and the EBP. (C) The activator hydrolyses ATP and uses the energy to mediate contact with σ^{54}-RNAP at the regulatory centre, removing the repressive interactions between the -12 promoter element and σ^{54}. This enables open promoter complex formation and initiation of transcription. Taken from Bush and Dixon (2012).

Central to this activity are the Walker A and B motifs, which bind and hydrolyse ATP respectively (Snider et al., 2008; Walker et al., 1982). In order to bind ATP, bEBPs oligomerise, predominantly into hexamers, forming a ring shape that creates the ATP binding pocket at the interface between individual proteins (Rappas et al., 2007). Using various ATP and ADP analogues that mimic the ATP-bound 'ground' state, the ATP hydrolysis 'transition' state and the ADP-bound post-hydrolysis state, structural data has revealed that there are substantial rearrangements within the ATPase domain during the ATP hydrolysis cycle, which are coupled to σ^{54} by two surface-exposed loops that are unique to bEBPs (Chen et al., 2007a; Rappas et al., 2007; Sysoeva et al., 2013). In particular, the GAFTGA motif within loop 1, which is the most conserved of the bEBP-unique features, is known to be an essential contact surface with σ^{54} (Dago et al., 2007; Zhang et al., 2009). In the ATP ground state, loops 1 and 2 are extended away from the bEBP oligomer, enabling the GAFTGA motif to contact region I of σ^{54} (Chen et al., 2007a). The transition state stabilises this interaction and region I undergoes a significant conformational rearrangement, increasing its distance from region III and relinquishing the obstruction of promoter DNA from core RNAP (Bose et al., 2008; Chen et al., 2007a; Sharma et al., 2014). In addition, contacts between σ^{54} and the -12 to -8 region of promoter DNA increased during this state, indicating the initial nucleation of DNA melting (Burrows et al., 2008). In the ADP-bound state, the loops flatten and return to their positions on the surface of the bEBP, while the open complex forms, with core RNAP contacting promoter DNA.

Therefore, the central ATPase domain is the essential catalytic component of bEBPs, necessary for σ^{54}–dependent transcription, but relies on the regulatory and DNA binding domains to maintain promoter specificity in appropriate cellular conditions.

1.3.2.2 Regulatory domain of bEBPs

The N-terminal domain is the regulatory component of bEBPs, confining ATPase activity to relevant conditions by responding directly or indirectly to a variety of environmental or metabolic signals. It is the most variable part of the bEBP, with the motifs contained determining the type of signal detected.

The bEBPs characterised to date perceive signals in three ways. Many bEBPs belong to two-component systems and are activated by a sensory histidine kinase that phosphorylates an aspartate residue of the response regulator motif within the regulatory domain (Yamamoto et al., 2005). For example, NtrC is phosphorylated by NtrB in nitrogen limiting conditions

(Austin and Dixon, 1992; Reitzer, 2003), while ZraR is phosphorylated by ZraS, which is thought to sense high concentrations of zinc and lead ions (Leonhartsberger et al., 2001). Others bEBPs directly bind small ligands that are generally substrates for the enzymatic activity of the gene products they regulate. Examples include DmpR and XylR, which are activated by the binding of specific aromatic compounds to vinyl 4 reductase (V4R) motifs in their regulatory domains (O'Neill et al., 1998; Shingler and Moore, 1994) and NorR and NifA, which use GAF motifs to bind nitric oxide and 2-oxoglutarate respectively (D'Autreaux et al., 2005; Little and Dixon, 2003). NifA is also subject to additional regulation by interaction with anti-activator protein NifL, which sequesters NifA in conditions inappropriate for nitrogen fixation. The ATPase activity of PspF, a bEBP that naturally lacks a regulatory domain, is also restricted by binding to an anti-activator, PspA, which dissociates in conditions where membrane integrity is compromised (Elderkin et al., 2005).

Upon deletion of the regulatory domains of several bEBPs, including DmpR, XylR and NorR, the resulting truncated proteins were found to be constitutively active and competent to activate σ^{54}-dependent transcription in the absence of a signal (Fernández et al., 1995; Gardner et al., 2003; O'Neill et al., 1998), as illustrated in Figure 5. In contrast, deletion of the regulatory domain of NtrC produced a constitutively inactive bEBP, unable to initiate transcription (Drummond et al., 1990). Therefore, signal detection by the regulatory domain can positively regulate the ATPase domain and stimulate its activity or, more frequently, relieve repression of the ATPase domain in a mechanism of negative regulation. These regulatory mechanisms can target the activity of the central domain at the level of oligomerisation, ATP hydrolysis and σ^{54} interaction.

1.3.2.3 DNA binding domain of bEBPs

The DNA binding domain provides promoter specificity for the action of the regulatory and ATPase domains of bEBPs. The HTH motif of the domain binds to specific UASs usually between 150 and 200 base pairs (bp) upstream of the transcription start site (Leonhartsberger et al., 2001; Porter et al., 1993; Tucker et al., 2005), tethering the bEBP to the correct promoter and enabling direct interaction with promoter-bound σ^{54} at the regulatory centre.

bEBPs bind in a dimeric conformation to UASs, which take the form of inverted repeats (IRs) (De Carlo et al., 2006; Sallai and Tucker, 2005; Vidangos et al., 2013). The length and spacing of repeats and the number of binding sites varies for each promoter and bEBP. For example, NtrC binds to two UASs upstream of the *glnA* promoter consisting of 5-6 bp half-

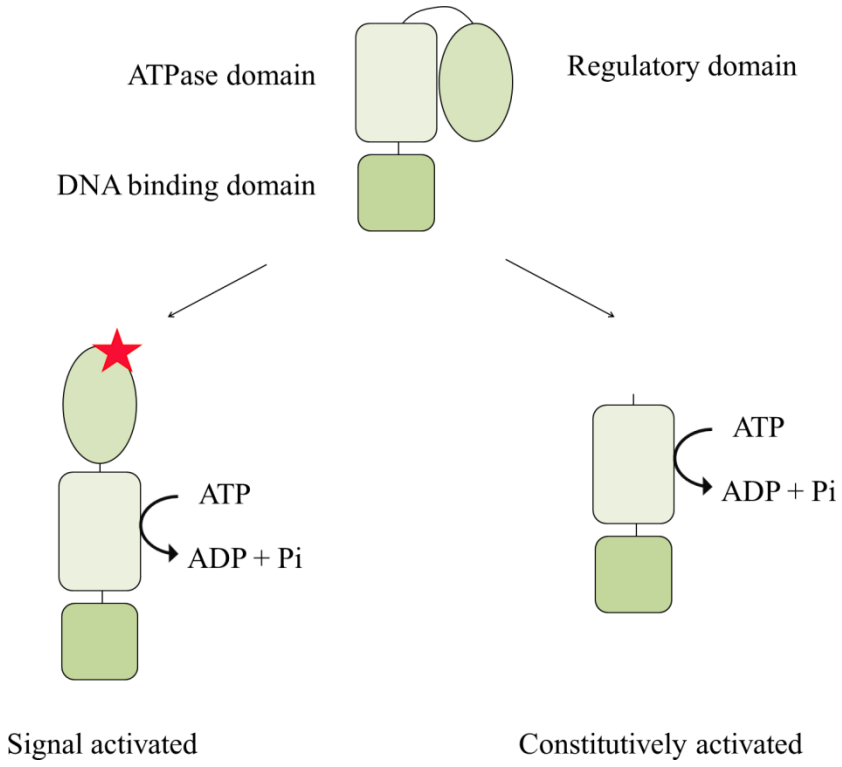


Figure 5 Negative regulation of the ATPase activity of bEBPs.

Schematic illustration of the negative mechanism of control of the central ATPase domain by the N-terminal regulatory domain of bEBPs. The regulatory domain inhibits the activity of the central domain, which is intrinsically competent for ATP binding and hydrolysis. Signal detection (e.g. ligand binding or phosphorylation) by the regulatory domain leads to derepression of the central domain. Deletion of the regulatory domain results in a constitutively activated bEBP.

sites separated by a 7 bp AT-rich region (Porter et al., 1993; Reitzer et al., 1989). In contrast, ZraR binds to a single IR consisting of 16 bp half-sites separated by 14 bp in the *zraP* promoter (Leonhartsberger et al., 2001), while NorR binds to three IRs consisting of 3-5 bp half-sites with a consensus of GT-N_7-AT (Tucker et al., 2005).

Interestingly, NorR must bind all three UASs in the *norVW* promoter in order to form a stable oligomer with the ability to initiate σ^{54}-dependent transcription upon activation by nitric oxide (Bush et al., 2015; Justino et al., 2005). Other bEBPs do not show the same strict DNA-dependency for oligomerisation as NorR, as indicated by the ability of truncated variants lacking the DNA binding domain to activate transcription (Berger et al., 1994; Jovanovic et al., 1999; North and Kustu, 1997). However, concentration of bEBPs at promoter DNA by UAS-binding is commonly found to facilitate and stabilise the process of oligomerisation (De Carlo et al., 2006; Porter et al., 1993).

1.3.3 The role of integration host factor in σ^{54}-dependent transcription

For a productive interaction between the σ^{54}-RNAP holoenzyme bound at the -24/-12 promoter elements and the bEBP tethered to the UAS, topological changes in DNA must occur. In some cases, the flexibility of the intervening sequence enables the interaction, as demonstrated by the *glnA* promoter (Carmona et al., 1997). Those lacking propensity for DNA flexibility rely on interaction with proteins to facilitate DNA looping. Accordingly, binding sites for the global regulator integration host factor (IHF), which alters DNA architecture to enable bending, are frequently found in the region between the UAS and the -24/-12 elements (Arfin et al., 2000; Huo et al., 2009). The position of IHF binding relative to the promoter ensures correct interfacing between σ^{54}-RNAP and bEBP, which is crucial for efficient transcription (Huo et al., 2006).

IHF-mediated topological changes in DNA can have other regulatory effects on σ^{54}-dependent transcription (Shingler 2011). These include preventing non-specific cross-activation by reducing promoter accessibility to other bEBPs (Perez-Martin and De Lorenzo, 1995; Wassem et al., 2000), stimulating formation of an open complex at the DmpR-regulated Po promoter (Sze et al., 2001), mediating cooperative DNA binding with PspF at the *psp* promoter (Jovanovic and Model, 1997) and recruiting σ^{54}-RNAP to the XylR-dependent *Pu* promoter and aiding its interaction with additional stimulatory elements further upstream (Bertoni et al., 1998).

1.3.4 The σ^{54} regulon

A major function of σ^{54}-associated RNAP is to transcribe genes involved in the assimilation of ammonia in order to generate the intracellular nitrogen donors glutamate and glutamine in response to nitrogen limitation (Reitzer, 2003; Reitzer and Schneider, 2001). The bEBP NtrC is the principle regulator during the nitrogen stress response and activates σ^{54}-dependent transcription from over 20 promoters (Brown et al., 2014). Among the proteins expressed are glutamine synthetase, the key enzyme involved in glutamine biosynthesis in low nitrogen conditions, and many transporter proteins involved in uptake of ammonia, glutamine, glutamate and other amino acids, in an effort to scavenge nitrogen-containing compounds for glutamine synthesis.

In addition, the σ^{54} regulon incorporates operons of diverse function including flagellar biosynthesis, maintenance of membrane integrity, toxic metal tolerance and several involved in the utilisation of alternative carbon sources, (Reitzer and Schneider, 2001; Zhao et al., 2010). Therefore, there is no obvious single theme linking the functions of all σ^{54}-regulated genes. According to the Pro54 database, there are 96 annotated σ^{54}-dependent promoters in *E. coli* K-12, (Liang et al., 2017), one of which directs transcription of the *rtcBA* operon.

1.4 Sigma 54-dependent transcription of the *rtcBA* operon in *E. coli*

1.4.1 Features of the *rtcBA* promoter

The *rtcB* gene of *E. coli* K-12 was discovered by Genschik et al. (1998) during an investigation to characterise the product of the downstream *rtcA* gene, so named for its activity as a RNA terminal cyclase, (further discussed in Section 1.6.4.1). Owing to the shared directionality and close proximity of the *rtcB* and *rtcA* genes, with the start codon of *rtcA* immediately following the stop codon of *rtcB*, they were considered to form a hitherto uncharacterised operon.

Examination of the DNA sequence upstream of the *rtcBA* operon led to the identification of a region with high similarity to the σ^{54} consensus sequence, including correctly spaced GG and GC doublets corresponding to the -24 and -12 promoter elements, suggesting that transcription of *rtcBA* was σ^{54}-dependent (Genschik et al., 1998). Supporting this, two putative binding sites for IHF were present further upstream of the -24/-12 elements, one on the other strand to the *rtcB* coding sequence and orientated away from *rtcB* (*ihf1*) and the

other on the same strand and orientated towards *rtcB* (*ihf2*). As discussed in Section 1.3.2, a key requirement for transcription via σ^{54}-associated RNAP is the ATPase activity of a dedicated bEBP to enable open promoter complex formation. Genschik et al. (1998) found that the product of adjacent gene *rtcR* fulfilled this function for *rtcBA* transcription.

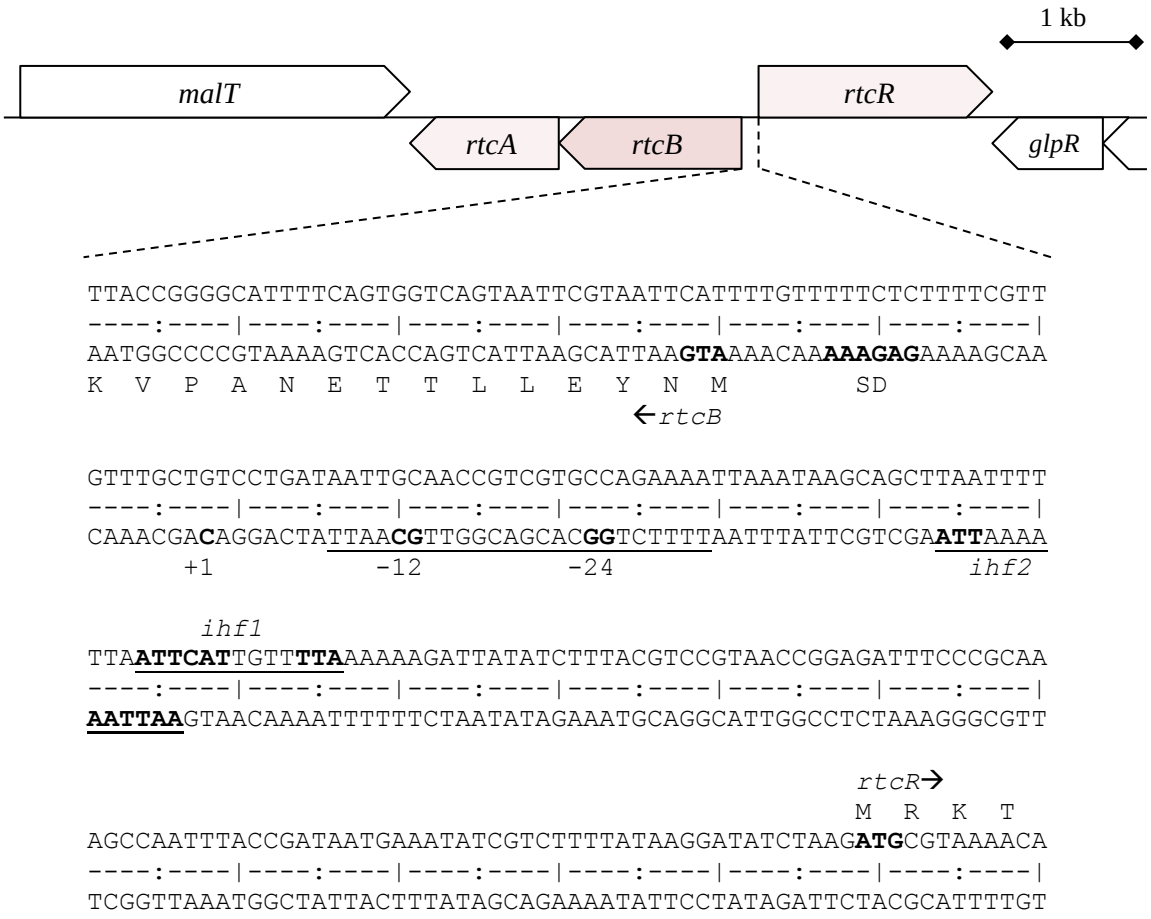
Figure 6 shows the organisation of the *rtcBA-rtcR* locus of *E. coli* and the sequence between *rtcB* and *rtcR*, annotated with the features identified by Genschik et al. (1998). The sequence of the σ^{54} and IHF binding sites are also compared to published consensus sequences.

1.4.2 RtcR, the bEBP of the *rtcBA* operon

The *rtcR* gene was considered as a candidate transcriptional activator of *rtcBA* owing to its resemblance to other bEBPs (Genschik et al., 1998). The central domain of RtcR showed approximately 60% similarity to equivalent domains of other bEBPs, including the seven highly conserved regions necessary for ATPase activity and interaction with σ^{54}. The C-terminus of RtcR was found to contain an HTH motif, consistent with the DNA binding domain of bEBPs for mediating interaction with UAS(s). However, the N-terminal domain of RtcR was found to have no similarity with the regulatory domain of other bEBPs, implying that it may have a novel mechanism of signal detection or recognise a different type of signal.

Recently, the regulatory domain of RtcR was found to contain a divergent CRISPR-associated Rossmann fold (CARF) domain, which are widespread among proteins genetically linked with CRISPR-cas systems but not directly associated with CRISPR-cas function (Makarova et al., 2014). The CARF domain is predicted to bind nucleotides or nucleotide-derived molecules via a Rossman-like fold, which contributes to a well conserved pocket containing a positively charged residue (Lintner et al., 2011; Makarova et al., 2014). The majority of CARF domain-containing proteins also have a winged HTH (wHTH) motif, a HTH variant (Aravind et al., 2005) and an effector domain, most commonly with enzymatic activity against nucleic acids, such as a nuclease (Makarova et al., 2014; Makarova et al., 2013). Together, this has led to the hypothesis that the CARF domain functions to detect ligand binding and transmit the signal to an effector domain, which responds to the stimulus with catalytic activity. This is also consistent with the function of RtcR as a bEBP. As a distantly related member of the CARF domain superfamily, RtcR has a similar domain organisation, with a HTH and an effector ATPase domain. However, RtcR proteins generally lack the highly conserved basic residue of the CARF domain pocket, suggesting they may bind alternative ligands (Makarova et al., 2014).

A



B σ^{54} consensus sequence: mrNrYTGGCACGNNNNTTGCWNNw

σ^{54} binding site, *rtcBA* promoter: TTTTCTGGCACGACGGTTGCAATT

C IHF consensus sequence: WATCAANNNNTTR

ihf1: ATTCATTGTTTTA

ihf2: AATTAAAAAATTA

Figure 6 Organisation and features of the σ^{54}-dependent *rtcBA* operon of *E. coli*.

(A) Genetic organisation of the *rtcBA-rtcR* gene cluster (filled boxes) and surrounding genes (empty boxes), with the sequence of the *rtcR-rtcB* intergenic region. Notable features identified by Genschik et al. (1998) are annotated including: the -24/-12 promoter elements; two putative IHF binding sites (*ihf1* and *ihf2*); the transcription start site of *rtcBA* (+1); putative Shine-Dalgarno (SD) sequence of *rtcB* and start codons of *rtcB* and *rtcR*. (B) Alignment of the σ^{54} consensus sequence (Barrios et al., 1999) with the σ^{54} binding site in the *rtcBA* promoter. (C) Alignment of the IHF consensus sequence (Friedman, 1988; Goodrich et al., 1990) with the proposed binding sites *ihf1* and *ihf2*. For (B) and (C) uppercase letters indicate highly conserved nucleotides; lowercase letters indicate weak conservation; R represents purines (A/G), Y represents pyrimidines (C/T); W is A/T; M is C/G and N is any nucleotide. The highly conserved GG and GC elements at -24 and -12 respectively are bold and nucleotides in the *rtcBA* promoter that match the consensus (including N) are red.

Using reverse-transcription polymerase chain reaction (RT-PCR) to measure transcription from the *rtcBA* promoter, Genschik et al. (1998) demonstrated that *rtcB* and *rtcA* were transcribed in *E. coli* cells expressing a truncated version of RtcR lacking the N-terminal regulatory domain but not full-length RtcR. As truncated RtcR became constitutively active in the absence of its regulatory domain, this suggests that the activity of RtcR is negatively regulated in the absence of its (currently unknown) activating signal, as well as establishing that RtcR acts as a transcriptional activator of *rtcBA*. Consistent with the presence of -24/-12 elements upstream of the *rtcBA* operon, inactivation of the *rpoN* gene that encodes σ^{54} prevented transcription of *rtcBA* during truncated RtcR expression. Transcription of both genes was restored by complementation of the *rpoN*-inactivated strain, demonstrating activation of *rtcBA* transcription by RtcR is σ^{54}-dependent. Further, primer extension analysis using RNA extracted from cells expressing truncated RtcR mapped the transcription start site of *rtcBA* to a C located 12 nucleotides downstream of the -12 element.

1.5 The RNA ligase RtcB

1.5.1 Types and functions of RNA ligases

RNALs are widespread in organisms from all domains of life and catalyse the formation of phosphodiester bonds to join breaks in the backbone structure of RNA during splicing, repair and modification processes. Known RNALs can be categorised into three groups, based on the type of phosphodiester bond created and the origin of the phosphate incorporated into the bond, which is referred to as the junction phosphate (Popow et al., 2012).

1.5.1.1 5′-3′ RNA ligases

The first RNAL activity was discovered in bacteriophages, plants and fungi and involved ATP-dependent ligation of RNA substrates with a 3′-hydroxyl group and a 5′-phosphate group to form a 2′-hydroxyl, 3′,5′-phosphodiester bond (Konarska et al., 1981; Peebles et al., 1979; Silber et al., 1972). This group of enzymes are referred to as 5′-3′ RNALs to indicate that a 5′-phosphate is joined to a 3′-hydroxyl (the phosphate-linked end is named first). ATP is central to ligation, which proceeds through three adenylate transfer steps (Cherepanov and de Vries, 2002; El Omari et al., 2006; Greer et al., 1983b). First, the ligase reacts with ATP to form a ligase-adenosine monophosphate (AMP) intermediate, displacing pyrophosphate. Next, AMP is transferred from the ligase to the 5′-phosphate of RNA, forming an

intermediate with a high energy phosphoanhydride bond, App-RNA. Finally, the 3′-hydroxyl of RNA attacks the double phosphate bond of the App-RNA intermediate, sealing the RNA substrates in a 3′,5′-phosphodiester bond and releasing AMP.

5′-3′ RNALs are frequently associated with domains that catalyse RNA end healing prior to ligation to increase their repertoire to RNAs with 2′,3′-cyclic phosphate and a 5′-hydroxyl terminals, which are commonly occurring products of ribonuclease activity (Burroughs and Aravind, 2016). This involves dephosphorylation of the 3′-end RNA substrate and phosphorylation of the 5′-end RNA substrate to produce suitable terminals for 5′-3′ RNAL activity. Therefore, the junction phosphate of the subsequent 3′,5′-phosphodiester bond originates from the γ-phosphate of NTP used to phosphorylate the 5′-end and not from the substrate-derived phosphate, which is another distinguishing feature of this type of RNAL (Greer et al., 1983b). Different enzymatic activities are responsible for end healing, as demonstrated in the following examples.

Rnl1, a 5′-3′ RNAL identified in bacteriophages and some bacterial and archaeal species, utilises a separate enzyme, Pnpk, to perform end healing prior to RNA ligation (Burroughs and Aravind, 2016; Popow et al., 2012; Zillmann et al., 1991). Pnpk has dual activity as a phosphatase, dephosphorylating the 2′,3′-cyclic phosphate to generate a 2′,3′-diol, and as a kinase, phosphorylating the 5′-hydroxyl group (Das and Shuman, 2013b). The best characterised Rnl1 enzyme is from T4 bacteriophage, which, in combination with T4 Pnpk, has been demonstrated to repair the anticodon loop of host $tRNA^{Lys}$, thereby evading attempts to prohibit phage propagation and allowing continued phage protein synthesis in *E. coli* (Amitsur et al., 1987).

Rnl1-type 5′-3′ RNALs are also widespread in plants and fungi, but unlike the Rnl1 ligases of bacteriophages, bacteria and archaea, all components required to process and ligate RNA terminals are contained within the same enzyme (Burroughs and Aravind, 2016; Popow et al., 2012). Prototypical ligases from *Arabidopsis thaliana* (AtRNL) and *Saccharomyces cerevisiae* (Trl1) have separate domains for cyclic phosphodiesterase, kinase and ligase activity (Remus and Shuman, 2014; Wang and Shuman, 2005). The cyclic phosphodiesterase activity hydrolyses the 2′,3′-cyclic phosphate group, yielding a 2′-phosphate-3′-hydroxyl terminal, while the kinase activity phosphorylates the 5′-hydroxyl group. Therefore, upon ligation the 3′,5′-phosphodiester linkage also bears a 2′-phosphate, which can be hydrolysed to the typical 2′-hydroxyl by the activity of a separate phosphotransferase (Culver et al., 1997; Keppetipola et al., 2007). The function of 5′-3′ RNALs of plants and fungi is well studied,

with essential roles in tRNA splicing, sealing tRNA halves following endoribonuclease-mediated intron removal, and in *XBP-1* mRNA splicing, an essential transcription factor in the UPR (Sidrauski et al., 1996; Yoshihisa, 2014).

Another family of 5′-3′ RNAL, Rnl2, exists in some species as a stand-alone ligase and in others is associated with Pnpk or other accessory phosphatases and kinases (Burroughs and Aravind, 2016). Rnl2 homologues are widespread in protists and viral genomes and to a lesser degree in archaea and bacteria (Popow et al., 2012). The most studied example is the T4 bacteriophage Rnl2, which exhibits a preference for ligating nicks in duplex RNA or RNA-DNA hybrids (Ho et al., 2004; Nandakumar and Shuman, 2004). Little is known about the functions of the Rnl2 family, with the exception of its established role in the mitochondria of kinetoplastids, where it is involved in insertion or deletion of uridine in mRNA, an editing process that creates translatable transcripts (McManus et al., 2001).

5′-3′ RNALs appear to be largely absent from metazoa, and are less prevalent in archaea and bacteria that other types of RNALs (Popow et al., 2012). 5′-3′ RNAL activity has been detected in human HeLa cell extracts but no homology to Rnl1, Rnl2 or any other ligase domain has been identified (Burroughs and Aravind, 2016; Zillmann et al., 1991).

1.5.1.2 2′-5′ RNA ligases

An unusual type of RNAL activity capable of forming a 3′-hydroxyl, 2′,5′-phosphodiester bond between RNA substrates with a 2′,3′-cyclic phosphate at the 3′-terminal and a 5′-hydroxyl at the 5′-terminal was first detected in several bacterial extracts (Greer et al., 1983a). Characterisation of the purified *E. coli* enzyme, named LigT (‘ligase tRNA’), revealed that it had both RNA ligase and nuclease activities, establishing an equilibrium that slightly favoured the nuclease reaction *in vitro* (Arn and Abelson, 1996). In contrast to other RNALs, the 2′-5′ RNAL activity was found to occur independently of nucleoside triphosphate (NTP) and unlike 5′-3′ RNALs, LigT catalysed direct ligation without prior end healing steps, with the junction phosphate donated by the 2′,3′-cyclic phosphate group.

Homologues of LigT are widely found in bacteria and archaea and 2′-5′ RNAL activity has been identified in some eukaryotes, including *Spinacea oleracea* chloroplasts (Arn and Abelson, 1996; Molina-Serrano et al., 2012; Rehse and Tahirov, 2005). However, the function of 2′-5′ RNALs is unknown. Structures of LigT homologues from the archaeon *Pyrococcus horikoshii* and the bacterium *Thermus thermophilus* revealed structural similarity with the cyclic phosphodiesterase domain of plant and fungal 5′-3′ RNALs (Kato et al., 2003;

Rehse and Tahirov, 2005). Accordingly, Remus et al. (2014) determined that LigT from *E. coli* can also hydrolyse the 2′,3′-cyclic phosphate group of an RNA substrate to form a 2′-phosphate-3′-hydroxyl terminal, implicating that LigT has roles in RNA end healing. Although the 2′-5′-phosphodiester bond formed by LigT is considered unusual, examples of its occurrence exist in nature (Trujillo et al., 1987). Therefore, both the cyclic phosphodiesterase and ligase/nuclease functions of LigT may be relevant for its function *in vivo.*

1.5.1.3 3′-5′ RNA ligases

A 3′-5′ RNAL activity, which catalysed ligation of RNA substrates with either 2′,3′ -cyclic phosphate or 3′-phosphate terminals together with 5′-hydroxyl terminals to generate a 2′-hydroxyl, 3′,5′-phosphodiester bond, was initially detected in HeLa cells and archaea (Filipowicz et al., 1983; Kjems and Garrett, 1988). Upon purification of the 3′-5′ RNAL activity, the proteins responsible were found to be homologues of *E. coli* RtcB (Englert et al., 2011; Popow et al., 2011), which was previously annotated as part of an operon with RNA cyclase RtcA, but to which no function had been assigned (Genschik et al., 1998). The mechanism of action of this novel family of RNALs was subsequently investigated using the *E. coli* and archaeal homologues, as discussed in Section 1.5.2. The known functions of RtcB are described in Section 1.6.

1.5.2 Mechanism of action of RtcB

RtcB and its homologues have a distinct mechanism of ligation to the 5′-3′ and 2′-5′ RNALs. Like 2′-5′ RNALs, RtcB was found to recognise RNA with 2′,3′-cyclic phosphate and 5′-hydroxyl ends (Arn and Abelson, 1996; Tanaka and Shuman, 2011). However, it was also able to ligate 3′-phosphate RNA terminals to the 5′-hydroxyl and joined these groups in a different way to 2′-5′ RNALs, forming a phosphodiester bond between the 3′-phosphate and 5′-hydroxyl in a 3′-5′ ligase activity (Filipowicz et al., 1983). In contrast to 5′-3′ RNALs, end healing was unnecessary for ligation and the junction phosphate originated from the RNA substrate rather than the γ-phopsphate of an NTP (Filipowicz and Shatkin, 1983; Zofallova et al., 2000).

Owing to the similarity with 2′-5′ RNALs, RtcB was initially thought to catalyse a single-step nucleophilic attack by the 5′-hydroxyl group on the 2′,3′-cyclic phosphate (Arn and Abelson, 1998). However, this was inconsistent with an early report that inclusion of GTP enhanced

ligation (Kjems and Garrett, 1988). Investigation into the requirements of purified RtcB confirmed that its activity was strictly dependent on the presence of GTP and that RNA ligation was abolished when GTP was omitted from the reaction or exchanged with another NTP (Tanaka et al., 2011a; Tanaka and Shuman, 2011). This indicated that the ligation pathway was more complex, likely involving guanylylation, similar to the nucleotidyltransferase activity involved in 5′-3′ RNA ligation.

Accordingly, RtcB has been demonstrated to react with GTP to form an RtcB-guanosine monophosphate (GMP) intermediate. Following incubation of RtcB with [α-^{32}P]GTP, Tanaka et al. (2011a) observed transferral of the radiolabel to RtcB, which was not the case upon incubation with γ-phosphate-radiolabelled GTP ([γ-^{32}P]GTP), indicating the formation of RtcB-GMP. Crystal structures of an archaeal RtcB homologue from *P. horikoshii* further elucidated the guanylylation pathway of RtcB (Desai et al., 2013; Englert et al., 2012; Okada et al., 2006). The site of guanylylation in *P. horikoshii* was identified as a highly conserved histidine residue within the active site, which formed a covalent bond with the α-phosphate of GMP (Englert et al., 2012). Biochemical analysis of *E. coli* RtcB revealed substitution of the equivalent histidine rendered RtcB inactive (Chakravarty et al., 2012). A highly conserved cysteine residue was also determined to be essential for RtcB activity, owing to its role in the coordination of two Mn^{2+} ions in the RtcB active site (Englert et al., 2012). Comparison of RtcB structures with and without GTP revealed the Mn^{2+} ions were central to the process of guanylylation (Desai et al., 2013). The first Mn^{2+} was associated with RtcB in the absence of GTP and interacted with GTP upon entry to the RtcB active site, while the second Mn^{2+} entered the active site in complex with GTP and orientated it with respect to the histidine residue to promote hydrolysis to GMP and expulsion of pyrophosphate. This is comparable to the nucleotidylation of 5′-3′ RNALs, except that they use Mg^{2+} and ATP as cofactors rather than Mn^{2+} and GTP (Cherepanov and de Vries, 2002; El Omari et al., 2006).

Following guanylylation of RtcB, GMP was transferred from the RtcB-GMP intermediate to the 3′-phosphate terminal of RNA, as demonstrated by incubation of RtcB-GMP with a non-ligatable RNA substrate bearing 5′-phosphate and 3′-^{32}phosphate ends (Chakravarty and Shuman, 2012). This resulted in the accumulation of an RNA intermediate with a 3′-end phosphoanhydride bond (RNA-ppG), distinguishable from the initial substrate by polyacrylamide gel electrophoresis (PAGE). Some residues important for this stage of ligation have also been characterised. For example, mutation of an arginine residue at a predicted 3′-terminal RNA contact site of RtcB (according to structural studies) allowed RtcB

guanylylation but would not permit guanylylation of the 3′-phosphate RNA terminal (Englert et al., 2012).

On the basis of the two guanylate intermediates identified, two mechanisms of RtcB ligation were presented in conflicting reports. Desai and Raines (2012) proposed that RtcB-GMP transfers GMP to 3′-phosphate RNA to form RNA-ppG, which is subject to attack by the adjacent 2′-hydroxyl, leading to cyclisation of the phosphate to generate a 2′,3′-cyclic phosphate suitable for ligation with 5′-hydroxyl RNA. This would imply that guanylylation is dispensible for ligation of RNA substrates with a preformed 2′,3′-cyclic phosphate. An alternative mechanism was presented by Chakravarty and Shuman (2012), whereby RtcB reacts with GTP to form RtcB-GMP and catalyses the hydrolysis of 2′,3′-cyclic phosphate to 3′-phosphate, which is then guanylylated by RtcB-GMP to form RNA-ppG. The phosphoanhydride bond of RNA-ppG is then attacked by the 5′-hydroxyl RNA, forming a 3′,5′-phosphodiester. In this scheme, guanylylation is necessary for ligation of all RtcB substrates.

In support of their model, Chakravarty and Shuman (2012) demonstrated that ligation of RNA with 2′,3′-cyclic phosphate and 5′-hydroxyl ends only occurred at significant levels in the presence of GTP. Furthermore, in the absence of a suitable 5′-hydroxyl substrate to complete ligation, RtcB-GMP has been shown to hydrolyse 2′,3′-cyclic phosphates, yielding 3′-phosphate terminals (Chakravarty and Shuman, 2012; Tanaka et al., 2011a). This cyclic phosphodiesterase activity of RtcB suggests that ligation proceeds via nucleophilic attack on the 3′-phosphate group rather than 2′,3′-cyclic phosphate. The finding that the presence of a 2′-hydroxyl significantly stimulates phosphodiester synthesis between the adjacent 3′-phosphate and an RNA substrate bearing a 5′-hydroxyl also favours hydrolysis of the 2′,3′-cyclic phosphate prior to the final step in ligation (Chakravarty et al., 2012). An arginine residue predicted by structural analysis of RtcB to coordinate interaction between the RNA-ppG and 5′-hydroxyl RNA substrates has also been confirmed to significantly reduce the formation of the phosphodiester bond, but not other ligation steps, when substituted for alanine (Maughan and Shuman, 2016). Therefore, it appears that RtcB plays a direct role in each stage of the ligation pathway: 2′,3′-cyclic phosphodiesterase activity, activation of the 3′-phosphate by guanylylation and sealing of RNA-ppG and 5′-hydroxyl RNA, all of which are dependent on the reaction of RtcB with GTP. Together, the biochemical and structural analyses of RtcB have contributed to the widely accepted mechanism of action proposed by Chakravarty and Shuman (2012), as depicted in Figure 7.

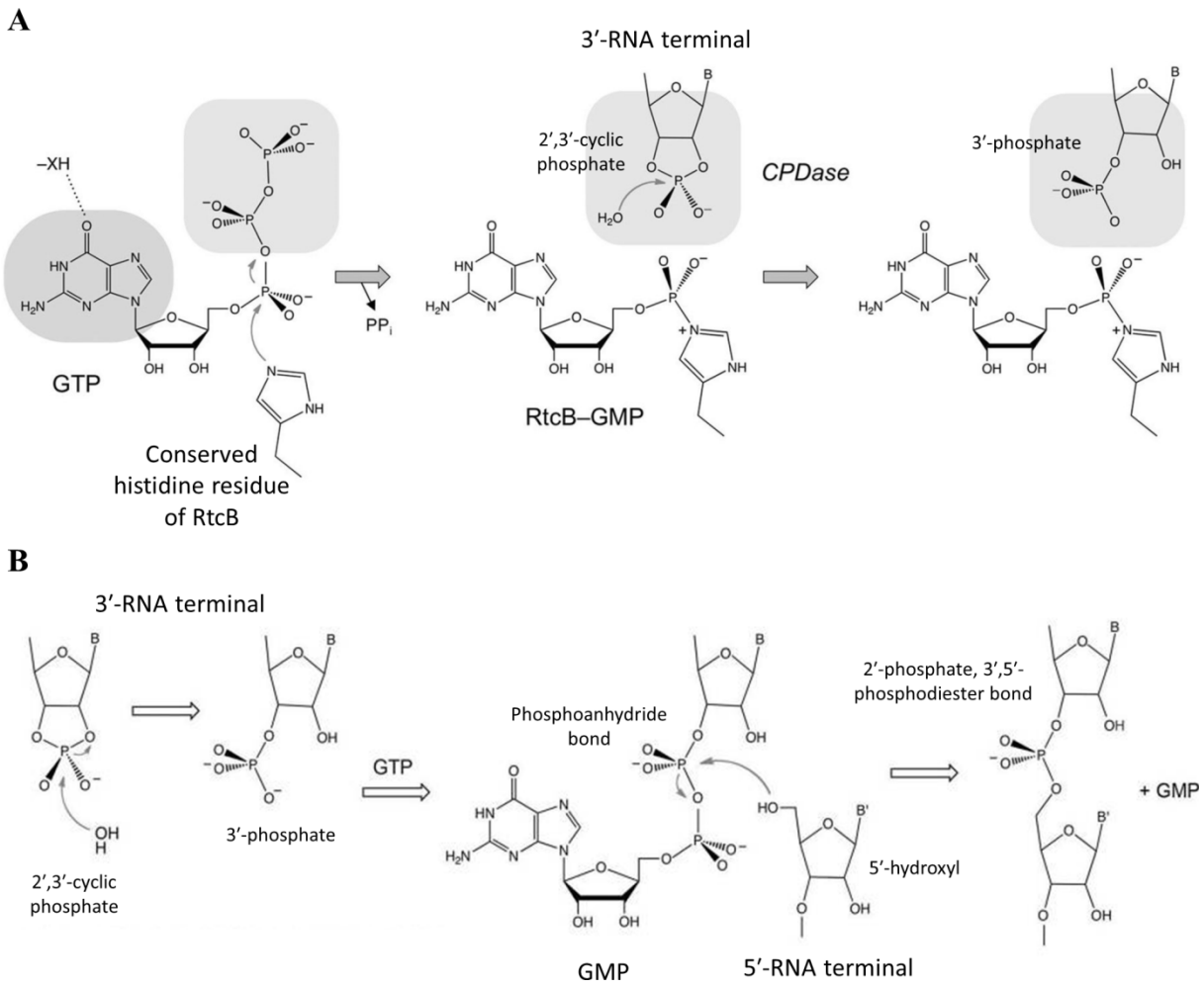


Figure 7 Mechanism of RNA ligation by RtcB.

(A) The cyclic phosphodiesterase activity of RtcB. RtcB is activated by guanylylation of a conserved histidine in the catalytic site, to form an RtcB-GMP intermediate. RtcB-GMP coordinates hydrolysis of RNA substrates with 2′,3′-cyclic phosphate terminals to form open 3′-phosphate ends. (B) The ligase activity of RtcB. Following hydrolysis of the 2′,3′-cyclic phosphate, GMP is transferred from RtcB to the 3′-phosphate RNA terminal, forming an RNA-ppG intermediate. The high energy, activated phosphoanhydride bond of RNA-ppG is attacked by an RNA substrate bearing a 5′-hydroxyl terminal, forming a 2′-hydroxyl, 3′,5′-phosphodiester bond and eliminating GMP. Taken from Chakravarty and Shuman (2012).

1.5.3 Additional activities of RtcB in DNA modification

RtcB has also been implicated in DNA repair or modification. Purified RtcB from *E. coli* was capable of performing intramolecular and intermolecular ligation of a 5′-hydroxyl DNA substrate with a single ribonucleotide at the 3′-end bearing a 3′-phosphate (Das et al., 2013). However, this activity was significantly less efficient than ligation of an equivalent RNA substrate (8% ligation compared to 68% respectively). RtcB also catalysed ligation of an all-DNA broken stem-loop structure, but not non-complementary DNA substrates or nicks in double-stranded DNA, suggesting that DNA ligation by RtcB requires intermolecular complementarity but unpaired 3′-phosphate or 5′-hydroxyl terminals (Das et al., 2013).

Although the final step of ligation between the 3′-DNA-ppG intermediate and 5′-hydroxyl appeared to be slower than ligation of RNA substrates, guanylylation of the 3′-phosphate to generate the DNA-ppG intermediate occurred at a similar rate to RNA-ppG formation and was catalysed on single-stranded, stem-loop and double-stranded DNA substrates (Das et al., 2013). Das et al. (2014) demonstrated that 3′-phosphate guanylylation of a single-stranded DNA substrate rendered it resistant to the activity of *E. coli* exonucleases I and III and the phosphatase domain of T4 Pnpk *in vitro.* Further, formation of DNA-ppG primed synthesis of templated DNA by a variety of DNA polymerases from different families (Chauleau et al., 2015; Das et al., 2014). However, the resulting repaired DNA contains an unconventional pyrophosphate bond, which Chauleau et al. (2015) determined was resistant to the proof-reading activity of DNA polymerases and other repair mechanisms.

Therefore, based on *in vitro* data, 3′-end guanylylation by RtcB could act as a protective cap or an unconventional primer for DNA synthesis, but the cost of incorporating a pyrophosphate into the DNA backbone in the latter scenario is unclear. Although inefficient compared to its RNAL activity, DNA ligation by RtcB suggests that it could also represent a novel family of DNA ligases. However, to date no *in vivo* function of RtcB has been attributed to its activity in DNA modification, either as a ligase or a capping enzyme.

1.6 Function and conservation of RtcB

1.6.1 Members of the RtcB family are present in all domains of life

As described in Section 1.5.1.3, 3′-5′ RNAL activity was first discovered in extracts from human HeLa cell lines (Filipowicz et al., 1983; Filipowicz and Shatkin, 1983) and shortly

thereafter in cell extracts from the archaea *Desulfurococcus mobilis* and *Haloferax volcanii*. (Gomes and Gupta, 1997; Kjems and Garrett, 1988; Zofallova et al., 2000). Purification of this activity from the archaeon *Methanopyrus kandleri* revealed that the protein responsible was a homologue of *E. coli* RtcB (Englert et al., 2011; Genschik et al., 1998). Accordingly, the purified RNAL from HeLa cells, named HSPC117, was also identified as a member of the RtcB family (Popow et al., 2011).

Bioinformatical analysis has revealed that RtcB is prevalent in all domains of life (Englert et al., 2011; Popow et al., 2012). RtcB is present in the majority of eukaryotic organisms, including metazoa and protists, with the notable exception of some lineages of fungi and plants. In addition, almost all known archaea have an RtcB homologue. A member of the RtcB family was identified in 55% of 965 bacteria (Popow et al., 2012), covering a diverse range of species from *E. coli* (Genschik et al., 1998) and *Salmonella enterica* serovar Typhimurium (Chen et al., 2013) to *Mycobacterium tuberculosis* (Selvaraj et al., 2012), *Rhodopirellula baltica* (Studholme and Dixon, 2004) and *Deinococcus deserti* (de Groot et al., 2014). An *rtcB* gene is also encoded by some viruses (Pope et al., 2014; Popow et al., 2012; Senčilo et al., 2013). Phylogenetic analysis of the RtcB family revealed three types that can be grouped at the domain level, with the eukaryotic and archaeal types showing greater similarity than the bacterial (Englert et al., 2011). This suggests that RtcB was present in the last common ancestor, with gene loss in some clades of bacteria and eukaryotes

1.6.2 Function in eukaryotes and archaea

1.6.2.1 tRNA maturation

The presence of introns within tRNA transcripts has been observed in organisms from all three domains of life and must be processed to generate mature and functional tRNA (Chan and Lowe, 2009; Popow et al., 2012). While bacterial tRNA introns are self-splicing, in eukaryotes and archaea tRNA splicing is a two-step enzymatic process, with endonuclease-mediated excision of the intron followed by ligation of the tRNA halves catalysed by an RNAL. The endonuclease reaction is conserved among eukaryotes and archaea, cleaving pre-tRNA at the intron-exon junctions to generate a 2′,3′-cyclic phosphate group at the 3′- end of the 5′-exon and a 5′-hydroxyl group at the 5′-end of the 3′-exon (reviewed in (Yoshihisa, 2014)). As described in Section 1.5.1.1, fungi and plants use 5′-3′ RNALs (exemplified by AtRNL and Trl1) to complete tRNA splicing, first healing the 2′,3′-cyclic phosphate to a 2′-phosphate-3′-hydroxyl group at the 5′-exon and the 5′-hydroxyl to a 5′-phosphate at the 3′-

exon before ligating the two terminals, incorporating γ-phosphate from ATP into the junction in the process (Greer et al., 1983b; Remus and Shuman, 2014). However, as both 3′-5′ and 5′-3′ RNAL activities were detected in HeLa cell extracts and implicated in ligating tRNA halves after intron excision in humans (Filipowicz et al., 1983; Filipowicz and Shatkin, 1983; Zillmann et al., 1991), it was unclear whether both the endonuclease and ligase reactions were conserved in eukaryote tRNA splicing or whether the ligation steps were catalysed by different mechanisms.

This was resolved by Popow et al. (2011), who identified the enzyme responsible for the 3′-5′ RNAL activity as an RtcB homologue known as HSPC117. Characterisation of human RtcB revealed it as the missing component of mammalian tRNA splicing (Popow et al., 2011). Small interfering (si)RNA-mediated depletion of human RtcB significantly reduced processing of pre-tRNA into mature tRNA in HeLa cell extracts, as did mutation of a conserved residue in the predicted catalytic core of RtcB, according to the structure in archaeon *P. horikoshii* (Okada et al., 2006). This was also the case for murine RtcB. Substrate specificity of human RtcB was confirmed by the inability to ligate tRNA exon halves modified by dephosphorylation of the 2′,3′-cyclic phosphate or phosphorylation of the 5′-hydroxyl and by the incorporation of the 2′,3′-cyclic phosphate into the phosphodiester bond, as indicated by Filipowicz and Shatkin (1983). Using inducible pre-tRNA^{Ile} carrying mutations and pre-tRNA^{Phe} from *S. cerevisiae* in order to distinguish from endogenous pre-tRNAs, the role of human RtcB was also investigated *in vivo* (Popow et al., 2011). Depletion of RtcB in HeLa cells resulted in a significant reduction in the formation of mature tRNA^{Ile} and tRNA^{Phe} reporters, confirming the role of RtcB in mammalian tRNA splicing. The use of 5′-3′ RNALs in fungi and plants in tRNA splicing may account for the low conservation of RtcB homologues in these organisms, as reliance on another type for tRNA ligation may have made RtcB dispensible (Englert et al., 2011).

Ligation of endonuclease-cleaved tRNA exons by the 3′-5′ RNAL activity of *H. volcanii* cell extracts (Gomes and Gupta, 1997; Zofallova et al., 2000), suggests that RtcB is also responsible for tRNA ligation in archaea. In support of this, purified RtcB from archaeon *Pyrobaculum aerophilum* was able to ligate pre-tRNA halves cleaved by the archaeal tRNA splicing endonuclease and the authors were able to demonstrate that the 2′,3′-cyclic phosphate from the 5′-exon was the source of the junction phosphate in the resulting phosphodiester bond (Englert et al., 2011). Furthermore, 5′-3′ RNALs are less prevalent in archaea than 3′-5′ RNALs (Popow et al., 2011), favouring a mechanism of tRNA ligation in archaea similar to metazoa rather than fungi and plants.

In mammals, RtcB forms part of a larger protein complex that mediates tRNA ligation (Figure 8). Additional components include the DEAD-box helicase DDX1, CGI-99, FAM98B and ASW. Popow et al. (2011) found that individual silencing of the other proteins in the tRNA ligase complex by RNA interference (RNAi) did not significantly affect tRNA maturation, indicating that RtcB was the only essential component. However, purified human RtcB cannot function indefinitely *in vitro,* which supports a regulatory function for one or more accessory proteins (Popow et al., 2014; Popow et al., 2012). Accordingly, Popow et al. (2014) identified archease as an additional component of the mammalian tRNA ligase complex, which enabled RtcB to consume all available RNA substrates *in vitro* and significantly impaired tRNA maturation upon depletion by siRNA. Archease was determined to stimulate the activity of human RtcB by mediating guanylation to form the active RtcB-GMP intermediate (see Figure 7). This activity was further enhanced by DDX1-mediated ATP hydrolysis, demonstrating a functional role for another member of the tRNA ligase complex. CGI-99 is implicated in mediating association of the components of the tRNA ligase complex as silencing of this protein destabilised the complex (Perez-Gonzalez et al., 2014). The function of FAM98B and ASW in tRNA ligation remain unclear but could include cellular localisation and stabilisation of RtcB or the RNA substrates. Homologues of DDX1, CGI-99, FAM98B and ASW are absent in bacteria and archaea (Popow et al., 2012) but archease is present in all domains of life and has also been found to augment the function of RtcB in archaea (Desai et al., 2014b) and bacteria (Desai et al., 2015), as discussed further in Section 1.6.4.3.

1.6.2.2 The unfolded protein response

The UPR is an intracellular stress response system of eukaryotes that arises from the accumulation of unfolded proteins in the lumen of the endoplasmic reticulum (ER). ER stress is sensed by transmembrane proteins that span the ER membrane, initiating signalling pathways and culminating in significant alteration of gene transcription and translation, the desired outcome of which is to decrease the protein load entering the ER and enhance the protein folding capabilities of the ER. There are three major UPR signalling pathways, which are conserved to varying degrees in eukaryotes (reviewed by Hollien (2013)). One of the three signalling pathways, the IRE1 branch of the UPR, detects the build-up of unfolded proteins via the intra-lumenal domain of transmembrane protein IRE1, triggering its oligomerisation to form an active endonuclease on the cytosolic face of the ER membrane. The IRE1 endonuclease excises the intron from *XBP-1* mRNA, producing spliced *XBP-1*

mRNA for translation. The resulting XBP-1 protein is a transcription factor that upregulates expression of genes to mediate recovery of ER stress, such as protein chaperones (reviewed by Walter and Ron (2011)). The IRE1 pathway is highly conserved in metazoan, fungi and plants, but not protists (Hollien, 2013). In *S. cerevisiae*, the 5′-3′ RNAL Trl1 is required to complete *XBP-1* splicing following IRE1 cleavage (Sidrauski et al., 1996). Following the characterisation of RtcB as the missing RNAL component of the tRNA splicing machinery in humans and other animals, a process catalysed by Trl1 and its homologues in fungi and plants, it was implicated as the ligase responsible for sealing *XBP-1* exons in metazoa. This hypothesis was addressed by three separate approaches.

Lu et al. (2014) conducted a genome-wide RNAi screen in mouse fibroblast cells to search for candidate *XBP-1*-ligating RNALs by utilising a synthetic IRE1 signalling circuit that, upon induction of ER stress and *XBP-1* splicing, initiated apoptosis rather than the UPR. siRNA-mediated depletion of RtcB rescued cells from apoptosis and further investigation revealed that accumulation of XBP-1 protein was significantly reduced, suggesting that RtcB functioned in the pathway at the level of *XBP-1* splicing. A catalytic mutant of RtcB lacking ligase activity was unable to rescue cells from apoptosis, demonstrating that the RNA ligase activity of RtcB was necessary for *XBP-1* splicing. Purified IRE1 and RtcB were sufficient to produce spliced *XBP-1* mRNA *in vitro*, confirming RtcB as the ligase component.

Kosmaczewski et al. (2014) used the nematode *Caenorhabditis elegans* to investigate the function of RtcB in a living metazoan model. Deletion of *rtcB* reduced the life span of nematodes and caused growth defects, which were recovered by expressing intron-less versions of relevant tRNAs, confirming the role of RtcB in tRNA ligation. However, under constant induction of ER stress, nematodes lacking RtcB had drastically reduced life spans irrespective of the presence of intron-less tRNAs, indicating a tRNA maturation-independent function for RtcB. RT-PCR revealed that while the longer, pre-spliced version of *XBP-1* was present in wild-type and *rtcB* null animals, the shorter post-splicing transcript was only detectable in wild-type, demonstrating the role of RtcB in ligation of IRE1-cleaved *XBP-1* mRNA. The phenotypes associated with *rtcB* deletion during the UPR were complemented by expression of wild-type RtcB, but as observed with Lu et al. (2014), a mutated version of RtcB lacking RNA ligase activity was not.

In addition to its stimulatory role in tRNA ligation, Jurkin et al. (2014) found that archease was also an RtcB co-factor for *XBP-1* ligation. While wild-type levels of spliced *XBP-1* mRNA were detectable in HeLa cells upon RNAi depletion of RtcB alone during ER stress,

depletion of both RtcB and archease abolished *XBP-1* splicing. This suggests that low levels of RtcB are sufficient for *XBP-1* exon ligation, with the proviso that archease is also present at wild-type levels. Jurkin et al. (2014) also demonstrated that targeted deletion of *rtcB* in plasma cells in a mouse model, where activation of the UPR is known to be important for production of high levels of antibody, significantly reduced antibody secretion in response to lipopolysaccharide challenge.

Together, Lu et al. (2014), Kosmaczewski et al. (2014) and Jurkin et al. (2014) robustly demonstrated that metazoan RtcB has an essential function in the IRE1 arm of the UPR, catalysing *XBP-1* mRNA ligation (Figure 9). This is of particular interest because of the emerging role of the UPR in diseases, such as cancer and neurodegenerative disorders (Hazari et al., 2016; Xiang et al., 2017), which could make RtcB a relevant therapeutic target. In support of this, an RNAi screen of *C. elegans* expressing α-synuclein as a model for Parkinson's disease identified RtcB as a candidate neuroprotective protein (Hamamichi et al., 2008). In accordance with the accumulation of misfolded proteins such as α-synuclein as a key feature of Parkinson's disease, the regenerative role of RtcB was confirmed to be mediated via the IRE1 arm of the UPR (Ray et al., 2014).

1.6.2.3 Other indications of function

An RNAi screen of *C. elegans* with a mutation in cytoskeletal protein β-spectrin as a model of axonal injury identified RtcB as a candidate regulator of axon regeneration (Nix et al., 2014). Further investigation by Kosmaczewski et al. (2015) revealed that RtcB localised at the tips of neurons damaged by laser axotomy and that deletion of the *rtcB* gene significantly increased axon regeneration in *C. elegans* as little as six hours following neuron damage. The phenotype could be complemented by expression of wild-type RtcB from a neuron-specific promoter but not with a ligase-defective mutant, indicating that the observed effect was caused by the loss of RtcB ligase activity in neurons. Therefore, as an inhibitor of axon regeneration, RtcB could represent a therapeutic target for axonal injury. Interestingly, when the *rtcB* mutant nematodes were provided with pre-spliced tRNA to rescue growth defects associated with impaired tRNA processing, axon regeneration remained high and did not return to the wild-type phenotype. Furthermore, expression of pre-spliced *XBP-1* in neuron had no effect on axon regeneration. This suggests that inhibition of axon regeneration by RtcB is independent of its function in tRNA maturation and the UPR and that RtcB ligase activity has a role in at least one other biological pathway in metazoa.

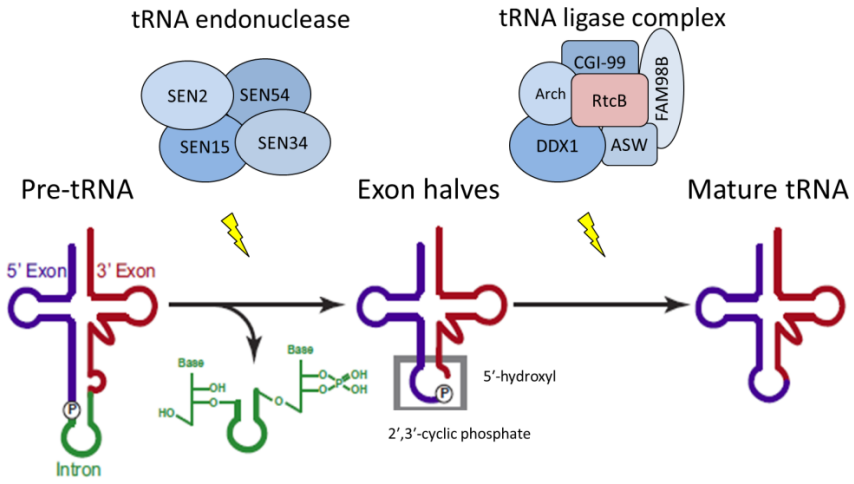


Figure 8 The role of RtcB in metazoan tRNA splicing.

Intron-containing pre-tRNAs undergo splicing to mature. The intron is removed by the tRNA endonuclease (consisting for four splicing endonuclease (SEN) subunits as indicated; speculative organisation) and the 5′- and 3′-exons are sealed by the tRNA ligase complex, which consists of ligase component RtcB and accessory proteins archease (arch), DDX1, CGI-99, FAM98 and ASW (speculative organization). Adapted from Popow et al. (2014).

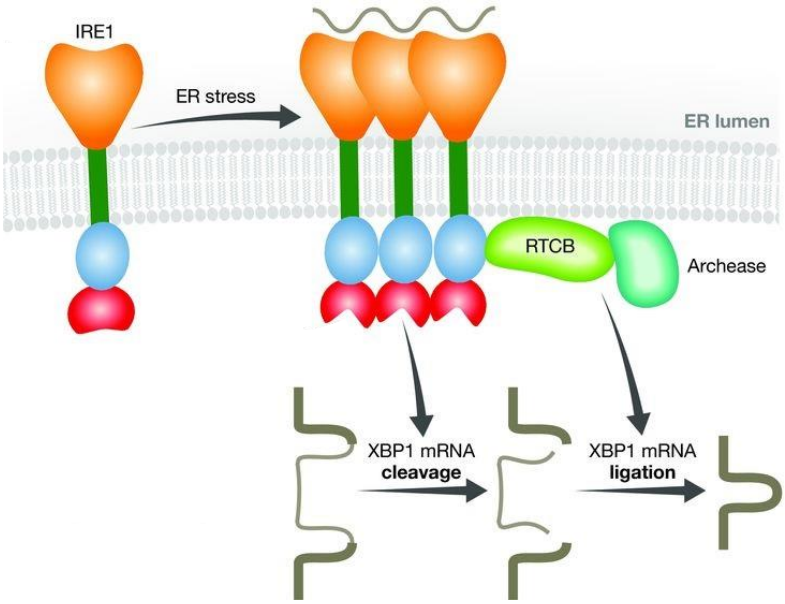


Figure 9 The role of RtcB in the IRE1 pathway of the UPR.

IRE1 senses the accumulation of unfolded proteins in the ER lumen and oligomerises to cleave *XBP-1* mRNA. RtcB and its cofactor archease ligate the exons, producing a spliced *XBP-1* transcript, which encodes a transcription factor that alters the transcription profile of the cell to alleviate ER stress. Taken from Filipowicz (2014).

RtcB was identified as a component of a large multi-protein RNA transporting granule isolated from mouse brain (Kanai et al., 2004). Transport of RNA is thought to be an important mechanism for regulating protein expression in specific cellular compartments, particularly in neurons (Kiebler and Bassell, 2006). Other components of the mammalian tRNA ligase complex, DDX1 and CGI-99, were also associated with the granule. More recently, RtcB, DDX1, FAM98B and CGI-99 were found to be complexed in another granule isolated from both nuclear and cytoplasmic fractions of human embryonic kidney cell line HEK293T (Perez-Gonzalez et al., 2014). RNA was also a significant constituent of the granule and was necessary for its assembly in the nucleus. CGI-99 was found to have roles in stabilisation of the granule. Monitoring of the shuttling activity of the granule revealed nuclear import was dependent on active transcription, while the complex was readily exported to the cytoplasm. The association of RtcB with granules that shuttle RNA around the cell via cytoskeletal interactions (Kanai et al., 2004) may explain the localisation of RtcB at the tips of damaged neurons (Kosmaczewski et al., 2015) and is also consistent with its nuclear function as a tRNA ligase and its cytoplasmic function in the UPR. These findings extend the role of the tRNA ligase complex to cell-wide RNA metabolism, and sequencing of the RNAs associated with these granules may provide further indication as to the function of RtcB in eukaryotes.

RtcB has been implicated in embryo development in mice, as depletion of RtcB was associated with increased foetal death at the post-implantation stage (Wang et al., 2010). The positive correlation between placental expression of RtcB and successful foetal development in the study might be explained by the essential roles of RtcB in tRNA maturation or the UPR. In addition, murein RtcB was found to be a regulator of cell adhesion with the extracellular matrix (ECM) by modulating the interaction of various adhesins (Ding et al., 2010; Hu et al., 2008). RtcB is also involved in cell migration by regulating the expression of various ECM components (Ma et al., 2014). Regulation of cell adhesion and migration by RtcB could also account for its role in foetal development. However, it is currently unclear how RtcB regulates these processes.

Together, this indicates that RtcB and its associated proteins may be involved in RNA ligation activities that affect a range of cellular processes in eukaryotes, many of which have implications in clinical settings.

1.6.3 Function in bacteria

Both biochemical and *in vivo* analyses have revealed a wide range of repair activity for RtcB purified from *E. coli*. *In vitro,* RtcB was able to seal a broken stem-loop RNA molecule cleaved by the site-specific endoribonuclease activity of yeast γ-toxin (Tanaka and Shuman, 2011) and perform intramolecular and intermolecular ligation of a single-stranded RNA substrate to generate a circular RNA product (Tanaka et al., 2011a), with the caveat that the RNAs had either a 2′,3′-cyclic phosphate or a 3′-phosphate group at the 3′-end and a 5′-hydroxyl at the 5′-end (Section 1.5.2). RtcB-mediated capping and ligation of DNA bearing the same terminals has also been demonstrated *in vitro* (Section 1.5.3). Biochemical analysis of three of the six RtcB paralogues present in the bacterium *Myxococcus xanthus* revealed that two resembled the *E. coli* RtcB homologue, with both RNAL and DNA guanylylation activities and weak DNA ligase activity, while the third was only capable of DNA guanylylation (Maughan and Shuman, 2015). This raises the possibility that some RtcB proteins they may have acquired functional specialisation for one or more of the range of activities identified *in vitro.*

The first *in vivo* indication of bacterial RtcB function was the observation that recombinant RtcB from *E. coli* complemented the normally lethal deletion of *S. cerevisiae* RNAL Trl1 (Tanaka et al., 2011b). This suggested it was capable of performing all essential RNA ligations in this organism, despite the difference in mechanism of the two ligases. Specifically, RtcB sealed tRNA halves following intron excision and cleaved *XBP-1* mRNA during the UPR, demonstrating RtcB-mediated ligation of both tRNA and mRNA substrates. Moreover, RtcB could also protect cells against the activity of fungal ribotoxin PaT, which, like γ-toxin, targets the anticodon loop of a specific tRNA (Tanaka et al., 2011b; Tanaka and Shuman, 2011).

Despite demonstrating the same functional capabilities as its eukaryotic homologues, RtcB must fulfil a different function in bacteria because bacterial introns are ribozymes capable of self-splicing (Hausner et al., 2014; Toro et al., 2007). Further, no introns are present in tRNA in *E. coli* (Chan and Lowe, 2009). Therefore, RtcB in *E. coli* and other bacteria is postulated to function as a repair response to stress-mediated RNA damage (Tanaka and Shuman, 2011). The regulation of *rtcBA* transcription in *E. coli* by alternative σ factor σ^{54} supports this notion, suggesting that RtcR detects damaged RNA or other stress signals resulting from RNA damage via its CARF-like regulatory domain, and initiates a response by activating *rtcBA* transcription.

Consistent with the hypothesis that RtcB catalyses the repair of stress-induced damage of cellular RNAs, a role for RtcB was recently identified in the response to RNA cleavage by the endoribonuclease MazF in *E. coli* (Temmel et al., 2016). MazF is the toxin component of the *mazEF* toxin-antitoxin (TA) system and is neutralised by interaction with its cognate antitoxin MazE (Pedersen et al., 2002). Degradation of labile MazE in stressful conditions, such as amino acid starvation and antibiotic treatment, permits the toxic effect of stable MazF (Christensen et al., 2003; Hazan et al., 2004; Sat et al., 2001). MazF is a ribosome-independent endoribonuclease, cleaving single-stranded RNA at ACA sites, resulting in inhibition of translation (Zhang et al., 2003; Zhang et al., 2005). However, Vesper et al. (2011) discovered that selective translation of a subset of mRNAs was still permitted during MazF induction. The translated mRNAs were characterised by the presence of ACA sites in the vicinity of the start codon, which, upon MazF cleavage, generated leaderless mRNAs lacking the Shine-Dalgarno (SD) sequence. Specialised 'stress' ribosomes lacking the anti-Shine-Dalgarno (aSD) sequence usually required to interact with the SD sequence of incoming mRNA and mediate translation initiation, were found to be responsible for translating the leaderless mRNAs. Moreover, the stress ribosomes were themselves created as a result of the endoribonuclease activity of MazF, which cleaved an ACA site at the 3′-terminal of 16S rRNA, removing the final 43 nucleotides including the aSD sequence. Thus, the authors considered the global changes in translation elicited by MazF to be a novel post-transcriptional stress response mechanism. Further investigation to elucidate the MazF regulon by polysome profiling followed by RNA deep sequencing revealed that *rtcB* mRNA was one of the selectively translated transcripts during MazF expression (Sauert et al., 2016). The generation of leaderless *rtcB* mRNA via MazF-cleavage at an ACA five nucleotides upstream of the start codon was verified by primer extension analysis.

There was no apparent functional clustering of the products of the 330 leaderless mRNAs identified in the MazF regulon to suggest an overall function for the proposed stress response mechanism (Sauert et al., 2016). However, Temmel et al. (2016) were able to demonstrate a role for RtcB. As the 43 nucleotide fragment of 16S rRNA cleaved by MazF remained stable and associated with the 30S ribosome subunit during MazF induction, this suggested that the 16S rRNA may be repaired and as a member of the MazF regulon, RtcB was an attractive candidate ligase. As such, RtcB was found to mediate the recovery of cells from MazF induction. Northern blot analysis revealed that the cleaved 43 nucleotide fragment was detectable for a longer period in an *rtcB* deletion strain compared to wild-type *E. coli* during recovery, while levels of intact 16S rRNA were lower. Furthermore, a reporter system in

which 16S rRNA was engineered to contain an SD sequence in place of the aSD to enable specific translation of modified *gfp* preceded by the aSD sequence, demonstrated *in vivo* that RtcB was responsible for ligating 16S rRNA fragments after MazF expression and that the resulting ribosomes were functional for translation of canonical mRNAs.

Other functions of RtcB in bacteria have been proposed in recent months by Engl et al. (2016). Surprisingly, they found that deletion of the *rtcBA* operon in *E. coli* produced a phenotype in standard laboratory conditions, contrary to previously published data (Englert et al., 2011; Genschik et al., 1998) Cells lacking *rtcB* had a reduced growth rate compared to wild-type cells and accordingly, RNA deep sequencing demonstrated changes in the transcriptome of the two strains. Over 570 genes were differentially expressed by over four-fold, of which the biggest functional grouping was genes associated with translation, which led the authors to propose a link between RtcB and the translation apparatus. Further investigation revealed changes in the ribosome profile of the *rtcB* deletion strain, indicative of increased rRNA degradation, which was complemented by expression of RtcB from a plasmid. RNA sequencing data in combination with qPCR showed that 16S rRNA was more unstable in the absence of RtcB but no difference in 23S or 5S rRNA was observed. This led to the suggestion that RtcB functions in ribosome homeostasis in *E. coli* (Engl et al., 2016). Although Temmel et al. (2016) also defined a role for RtcB in repair of 16S rRNA, their findings do not corroborate those of Engl et al. (2016), as they did not observe any difference in 16S rRNA stability upon *rtcB* deletion in the absence of MazF induction.

In addition, the RNA sequencing data showed that genes involved in chemotaxis and motility were upregulated in the *rtcB* deletion (Engl et al., 2016). Motility assays confirmed that the spread of cells increased in the absence of RtcB, while complementation reversed this effect, suggesting that the activity of RtcB negatively regulates cell motility under standard growth conditions. Although the mechanism behind this phenotype is unclear, this is in line with the finding that human RtcB is involved in mediating cell-cell interactions and migration (Hu et al., 2008; Ma et al., 2014).

1.6.4 Genes commonly associated with *rtcB* in bacteria

1.6.4.1 The RNA cyclase RtcA

In *E. coli*, *rtcB* is localised in an operon with *rtcA* that is transcriptionally regulated by the σ^{54}-RNAP holoenzyme in an RtcR-dependent manner (Genschik et al., 1998). RtcA belongs

to a family of 3′-terminal phosphate cyclases, which are also highly conserved with homologues in the eukaryal, archaeal and bacterial domains (Genschik et al., 1997). Bioinformatic analysis identified 32 bacteria that share the genetic arrangement of an *rtcBA* operon with a neighbouring *rtcR* gene, which in the majority of cases appeared in the same orientation as in *E. coli* (Das and Shuman, 2013a). While this is a relatively small proportion of the bacteria known to have an RtcB homologue, suggesting that *rtcB* and *rtcA* are not generally genetically linked, the *rtcR-rtcBA* gene cluster is conserved across four different phyla, namely Proteobacteria, Acidobacteria, Chlamydiae and Planctomycetes. The presence of an adjacent *rtcR* gene implies that, as demonstrated in *E. coli*, the *rtcBA* operon is transcriptionally regulated by RtcR from a σ^{54}-dependent promoter in these bacteria.

The activity of RtcA was first characterised in HeLa cell extracts and found to catalyse ATP-dependent conversion of a 3′-terminal phosphate to a 2′,3′-cyclic phosphates at the 3′-terminal of RNA (Filipowicz et al., 1983; Filipowicz et al., 1985). This activity was found to be conserved in human and *E. coli* RtcA and was specific to RNA and not DNA (Genschik et al., 1997; Genschik et al., 1998). The nucleotidyltransferase steps of the RtcA cyclisation pathway are analogous to the mechanism of action of RtcB. First, RtcA is adenylylated on a conserved histidine residue by ATP, coordinated by a divalent metal cofactor (Mn^{2+} or Mg^{2+}), forming an RtcA-AMP intermediate (Chakravarty et al., 2011; Tanaka et al., 2010). Next, the RtcA-AMP is hydrolysed, transferring AMP to the 3′-phosphate group of substrate RNA to generate an RNA-ppA intermediate (Desai et al., 2014a; Reinberg et al., 1985). Finally, nucleophilic attack of the activated phosphoanhydride bond of the intermediate by the terminal 2′-hydroxyl group of the RNA produces a 2′,3′-cyclic phosphodiester at the RNA 3′-terminal, releasing AMP (Filipowicz et al., 1985).

In addition to its 3′-terminal phosphate cyclase activity, the repertoire of RtcA has been expanded. Chakravarty and Shuman (2011) established that RtcA has activity as a 5′-adenylyltransferase, transferring of AMP from RtcA to a 5′-phosphate group of RNA or DNA in a manner similar to 5′-3′ RNALs and DNA ligases (Cherepanov and de Vries, 2002; El Omari et al., 2006). Das and Shuman (2013a) demonstrated that RtcA also had 2′-terminal phosphate cyclase activity, albeit at a five-fold slower rate than the 3′-phosphate cyclase activity.

The genetic linkage of *rtcB* and *rtcA* led to initial speculation that they are also mechanistically linked, with *E. coli* RtcA postulated to provide or maintain 2′,3′-cyclic phosphate terminals for RtcB-mediated RNA repair (Englert et al., 2011; Filipowicz et al.,

1983). However, Tanaka and Shuman (2011) found that inclusion of RtcA did not augment ligation of broken synthetic tRNA stem loops by RtcB *in vitro* and further elucidation of the substrate specificity of RtcB showed that it can use 3′-phosphate RNA terminals for ligation directly, precluding the need for RtcA-mediated 3′-phosphate cyclisation (Chakravarty et al., 2012). Although the 3′-terminal phosphate cyclase activity of RtcA would appear to be redundant in the case of RtcB, its ability to also catalyse the conversion of a 2′-phosphate to a 2′,3′-cyclic phosphate provides an RtcB-ligatable substrate and represents a plausible scenario for activity of RtcA and RtcB in the same pathway.

Although three *in vitro* activities have been reported for RtcA, the function *in vivo* is unclear. Engl et al. (2016) observed the same increased motility phenotype as described for an *rtcB* deletion strain with an *rtcA* deletion, which could be complemented by RtcA expression from a plasmid. However, no evidence was presented to suggest that RtcA was also implicated in 16S rRNA maintenance and repair (Engl et al., 2016; Temmel et al., 2016). Thus it is unclear whether RtcB and RtcA are involved in the same reaction pathways or act independently *in vivo*. Co-transcription of *rtcBA* from the σ^{54}-dependent promoter indicates that RtcB and RtcA at least function in the same cellular conditions, if not the same reactions.

1.6.4.2 Ro, an RNA scavenger

In other bacterial species, the *rtcBA* operon often includes extra genes. The most common addition is *ro*, which was found in the *rtcBA* locus in ten bacteria across three phyla, out of the 32 with *rtcBA* and an adjacent *rtcR* (Das and Shuman, 2013a). As observed in *E. coli*, transcription of the *ro-rtcBA* operon of *S.* Typhimurium was activated by expression of constitutively active RtcR lacking the N-terminal regulatory domain (Chen et al., 2013), confirming this as a conserved feature of the *rtcBA* operon.

Ro is a ring-shaped protein that associates with two to four small non-coding RNAs known as Y RNAs to form a functional ribonucleoprotein (Wolin et al., 2013). It is conserved in most metazoa and found in approximately 5% of bacterial species but is absent in *E. coli* (Sim and Wolin, 2011). The Ro ribonucleoprotein forms a complex with exoribonuclease polynucleotide phosphorylase (PNPase), which increases the RNA degradation activity of PNPase for structured RNAs, but not single-stranded RNAs (Chen et al., 2013). Ro ribonucleoprotein interacts with other exoribonucleases involved in RNA metabolism during environmental stress (Chen et al., 2007b). As Ro binds misfolded RNAs in addition to Y RNAs (Sim and Wolin, 2011), this suggests that it scavenges cellular RNAs and directs their

degradation, altering the specificity and activity of various RNA degrading proteins via Y RNA-mediated tethering.

1.6.4.3 The RtcB cofactor archease

The *archease* gene is found in all domains of life and has a strikingly conserved genomic context, commonly positioned adjacent to genes encoding proteins that function in RNA or DNA processing, such as *rtcB* (Canaves, 2004). Although not widely distributed in bacteria (it is absent from *E. coli*), *archease* is often found in an operon with *rtcB* in bacteria and also in archaea (Desai et al., 2014b).

As discussed, archease is a confirmed cofactor of RtcB in animals and is necessary for its full activity in tRNA and *XPB-1* ligation by mediating guanylylation of RtcB (Jurkin et al., 2014; Popow et al., 2014). In contrast, archease from archaeon *P. horikoshii* accelerated both guanylylation of the 3′-end phosphate of the 5′-RNA substrate and attack of the guanylylated 3′-end by the 5′-hydroxyl group of the 3′-substrate (Desai et al., 2014b). Further, *P. horikoshii* archease reduced the specificity of RtcB for GTP, enabling the use of other NTPs efficiently. Archease is also a cofactor of bacterial RtcB: the RNA ligase activity of RtcB from bacterium *T. thermophilus* was significantly enhanced in the presence of endogenous archease and archease from *P. horikoshii* (Desai et al., 2015). However, the addition of archease had no effect on the activity of RtcB from *Thermobifida fusca*, a bacterial species that does not encode *archease* (Desai et al., 2015). Therefore, RtcB homologues can be catagorised as archease-dependent or –independent and it appears that coevolution of RtcB and archease is necessary to enable a functional interaction. Archease has also been characterised as an accessory protein of 5-methylcytosine tRNA methytransferase, increasing its specificity for methylation of a specific cytosine in archaeon *Pyrococcus abyssi* (Auxilien et al., 2007). Therefore, archease may function as a general modulator of proteins involved in RNA metabolism.

As demonstrated with archease, the genomic context of RtcB supports its proposed function in RNA repair in bacteria, where it is localised and often co-transcribed with other genes encoding enzymes involved in RNA metabolism, as exemplified here by RtcA, Ro and archease.

1.7 Aims of this work

As discussed in this chapter, the *rtcB* gene was first annotated by Genschik et al. (1998) as part of the σ^{54}–dependent *rtcBA* operon of *E. coli*. Since then, the activity of the highly conserved RtcB family as 3′-5′ RNA ligases has been characterised *in vitro* (Chakravarty and Shuman, 2012; Chakravarty et al., 2012; Tanaka et al., 2011a; Tanaka and Shuman, 2011). Furthermore, functional roles for RtcB in tRNA maturation and the UPR in metazoa (Jurkin et al., 2014; Kosmaczewski et al., 2014; Lu et al., 2014; Popow et al., 2011) and in ribosome homeostasis and repair of MazF-induced stress ribosomes in *E. coli* (Engl et al., 2016; Temmel et al., 2016) have been reported. However, at the outset of this project there was no published information regarding the role of RtcB in bacteria, and to date the extent of its function remains unclear. Further, the regulation of σ^{54}–dependent *rtcBA* transcription by RtcR has yet to be fully characterised.

Therefore, the objective of this project was to understand the regulation and biological function of RtcB in *E. coli* K-12. Specifically, we aimed to characterise the roles of RtcR and IHF as regulators of the *rtcBA* operon, to identify conditions where transcription of *rtcBA* is activated by RtcR *in vivo* and to identify the natural RNA substrates of RtcB.

Chapter 2 Materials and methods

2.1 Bacterial strains, media and growth conditions

E. coli strains used in this study (Table 1) were cultured at 37°C and 160 rpm in LB media (1% tryptone, 1% NaCl, 0.5% yeast extract, pH 7.5). Nutrient agar (NA; 0.5% peptone, 0.2% yeast extract, 0.1% Lab-Lemco powder, 1.5% agar; 0.5% NaCl) was used for bacterial growth on solid media at 37°C for 14 to 18 hours, unless otherwise stated.

Growth media was supplemented as required with 25 µg/ml kanamycin (kan), 30 µg/ml or 100 µg/ml ampicillin (amp; for low- and high-copy-number plasmids respectively) and 50 µg/ml chloramphenicol (cml). Growth media was supplemented with 20 µg/ml L-glutamine as indicated to allow good growth in the absence of σ^{54}. Transcription from LacI-regulated promoters was induced by the addition of 0.1 to 1 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) as indicated. Transcription from AraC-regulated promoters was induced by 0.2% arabinose. 40 µg/ml 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) was used to detect β-galactosidase activity on solid medium.

2.2 Standard DNA procedures

Molecular cloning to construct plasmids used in this study (Table 2) was conducted using standard methods (Sambrook et al., 1989). Q5 High-Fidelity DNA polymerase was used to amplify DNA fragments for cloning by PCR according to the manufacturer's instructions (NEB). PCR amplifications for colony PCR and arbitrary PCR were performed using GoTaq DNA polymerase (Promega) or DreamTaq DNA polymerase (Thermo Fisher Scientific) and in accordance with the instructions of the respective manufacturer's. Oligonucleotides used are listed in Table 3. DNA fragments were resolved by gel electrophoresis on a 1% TAE (40 mM Tris-base, 20 mM acetic acid, 1 mM EDTA) agarose gel containing 0.5 µg/ml Ethidium bromide (Sigma-Aldrich) using 10X Orange G (25% ficoll, 0.4% Orange G dye) loading dye and a 1 kb ladder (GeneRuler, Thermo Fisher Scientific) at 125 V for 30 minutes. DNA was purified from gels and enzymatic mixtures using the QIAquick Gel Extraction or QIAquick PCR Purification kits (both Qiagen) according to the provided instructions.

PCR products and plasmids were digested with the relevant Fast-Digest restriction endonucleases from Thermo Fisher Scientific and ligated using the compatible T4 DNA

ligase, according to the manufacturer's recommendations. Digested plasmids were treated with Fast Alkaline Phosphatase (Thermo Fisher Scientific) before ligation to minimise self-ligation. Plasmid DNA was purified from bacterial culture using the QIAprep Spin Miniprep kit (Qiagen) or Untergasser's rapid boiling plasmid extraction protocol (Untergasser, 2006). With the exception of the low-copy temperature sensitive (ts) R1 plasmids (Table 2), which were diluted from overnight (ON) culture into fresh LB medium grown exponentially then cultured for 2-4 hours at 42°C to induce uncontrolled plasmid replication, all plasmids were purified from ON culture. All cloned inserts were sequenced using DNA Sequencing & Services, Dundee, Scotland or Eurofins Genomics, Ebersberg, Germany.

Table 1 Strains used in this study

Strain	Genotype	Keio collection strain (if applicable)	Source or reference
MG1655	*E. coli* K-12 MG1655 F^- λ^- *ilvG*$^-$ *rfb-50 rph-1*		Laboratory collection
TB28	MG1655Δ*lacIZYA*		Laboratory collection
C41	*E. coli* BL21 B F^- *ompT gal dcm lon hsdS*$_B$(*r*$_B$-*m*$_B$-) λ(DE3 [*lacI lacUV5-T7p07 ind1 sam7 nin5*]) [*malB*$^+$]$_{K-12}$(λ^S)		(Miroux and Walker, 1996)
DH5α λpir	*E. coli* K-12 DH5α pir$^+$ F^-*endA1 glnV44 thi-1 recA1 gyrA96 relA1 deoR nupG purB20* φ80d*lacZ*ΔM15 Δ(lacZYA-*argF*)U169 hsdR17(*r*$_K$-*m*$_K^+$) λ^-		Laboratory collection
SC28	*E. coli* K-12 MC1000Δ*mazEF*		(Pedersen et al., 2002)
GH01	MG1655Δ*rpoN*		This study
GH02	TB28Δ*rpoN*		This study
GH03	TB28*rtcBA-lacZ*		This study
GH04	TB28*rtcBA-lacZ*Δ*ihfA::kan*	JW1702	This study
GH05	TB28*rtcBA-lacZ*Δ*ihfB::kan*	JW0895	This study
GH06	MG1655Δ*rtcR*		This study
GH07	TB28Δ*rtcR*		This study
GH08	MG1655Δ*rtcB*	JW3384	This study
GH09	MG1655Δ*rtcA*	JW5688	This study
GH10	MG1655Δ*rtcBA*		This study
GH11	MG1655Δ*rtcBAR*		This study
GH12	MG1655*rtcR*^{M1X}		This study
GH13	TB28*rtcR*^{M1X}		This study
GH14	SC28Δ*rtcB*	JW3384	This study

Table 2 Plasmids used in this study

Plasmid	Description	Source or reference
pOU254	mini-R1 *ts*, *lacZYA* transcriptional fusion, Amp^R	Laboratory collection
pOU254K	mini-R1 *ts*, *lacZYA* transcriptional fusion, Kan^R	Laboratory collection
pGH254	pOU254 ter_{rrnB}	This study
pGH254K	pOU254K ter_{rrnB}	This study
pGH254BA0	pGH254 P_{rtcBA}^{BA0}*::lacZYA*	This study
pGH254BA1	pGH254 P_{rtcBA}^{BA1}*::lacZYA*	This study
pGH254BA2	pGH254 P_{rtcBA}^{BA2}*::lacZYA*	This study
pGH254BA3	pGH254 P_{rtcBA}^{BA3}*::lacZYA*	This study
pGH254BA3-IR_{M1}	pGH254BA3 IR_{M1} mutation	This study
pGH254BA3-IR_{M2}	pGH254BA3 IR_{M2} mutation	This study
pGH254BA3-IR_{M3}	pGH254BA3 IR_{M3} mutation	This study
pGH254BA3-IR_{M4}	pGH254BA3 IR_{M4} mutation	This study
pGH254BA3-IR_{M5}	pGH254BA3 IR_{M5} mutation	This study
pGH254BA3-IR_{M6}	pGH254BA3 IR_{M6} mutation	This study
pGH254BA3-IR_{M7}	pGH254BA3 IR_{M7} mutation	This study
pGH254BA3-IR_{M8}	pGH254BA3 IR_{M8} mutation	This study
pGH254BA3-IR_{M9}	pGH254BA3 IR_{M9} mutation	This study
pGH254BA3-IR_{M10}	pGH254BA3 IR_{M10} mutation	This study
pGH254BA3-IHF_{M1}	pGH254BA3 IHF_{M1} mutation	This study
pGH254BA3-IHF_{M2}	pGH254BA3 IHF_{M2} mutation	This study
pGH254BA3-IHF_{M3}	pGH254BA3 IHF_{M3} mutation	This study
pGH254BA3-IHF_{M4}	pGH254BA3 IHF_{M4} mutation	This study
pGH254BA3-IHF_{M5}	pGH254BA3 IHF_{M5} mutation	This study
pGH254BA3-IHF_{M6}	pGH254BA3 IHF_{M6} mutation	This study
pGH254BA3-IHF_{M7}	pGH254BA3 IHF_{M7} mutation	This study
pGH254R0	pGH254 P_{rtcR}^{R0}*::lacZYA*	This study
pGH254R1	pGH254 P_{rtcR}^{R1}*::lacZYA*	This study
pGH254R2	pGH254 P_{rtcR}^{R2}*::lacZYA*	This study
pGH254R3	pGH254 P_{rtcR}^{R3}*::lacZYA*	This study
pGH254R3$_{mut10}$	pGH254R3 -10 mutation	This study
pGH254KR3	pGH254K P_{rtcR}^{R3}*::lacZYA*	This study
pGH254KBA2	pGH254K P_{rtcBA}^{BA2}*::lacZYA*	This study
pBR322	pMB1, Amp^R Tet^R	(Bolivar et al., 1977)
pGH10	pBR322::*yaeP-rof-tilS*	This study
pGH12	pBR322::*yaeP-rof-tilS-yaeR*	This study
pGH13	pBR322::*yaeP-rof-tilS*	This study
pBAD322K	pBR322, *araC* P_{BAD}, Kan^R	(Cronan, 2006)
pGH322R	pBAD322K P_{BAD}::SD_{opt}::*rtcR*	This study
pGH322RΔN	pBAD322K P_{BAD}::SD_{opt}::*rtcRΔN*	This study
pBAD322C	pBR322, *araC* P_{BAD}, Cml^R	(Cronan, 2006)
pGH322CR	pBAD322C P_{BAD}::SD_{opt}::*rtcR*	This study
pGH322CRΔN	pBAD322C P_{BAD}::SD_{opt}::*rtcRΔN*	This study
pMG25	pUC, $lacI^q$ $P_{A1/O4/O3}$, Amp^R	Laboratory collection
pGH25rpoN	pMG25 $P_{A1/O4/O3}$::SD_{opt}::*rpoN*	This study
pGH25rtcBA	pMG25 $P_{A1/O4/O3}$::SD_{opt}::*rtcBA*	This study
pGH25H_6RC	pMG25 $P_{A1/O4/O3}$::SD_{opt}::H_6-$rtcR_{C\text{-}term}$	This study

Plasmid	Description	Source or reference
pGH251	pMG25 *rtcR-rtcBA* intergenic sequence	This study
pGH252	pMG25 *rtcR-rtcBA* intergenic sequence, - *rtcO*	This study
pBAD33	p15, *araC* P_{BAD}, Cml^R	(Guzman et al., 1995)
pSC3326	pBAD33 P_{BAD}:: *mazF*	(Pedersen et al., 2002)
pKW82	pBAD33 P_{BAD}::*vapC*	(Winther and Gerdes, 2009)
pKP3035	pBAD33 P_{BAD}::*relE*	(Pedersen et al., 2002)
pKP3065	pBAD33 P_{BAD}::*yafQ*	Laboratory collection
pMCD8	pBAD33 P_{BAD}::*parE*	Laboratory collection
pGH33H_6R	pBAD33 P_{BAD}::SD_{opt}::H_6-*rtcR*	This study
pGH33H_6RΔN	pBAD33 P_{BAD}::SD_{opt}::H_6-*rtcRΔN*	This study
pGH33H_6IHF	pBAD33 P_{BAD}::SD_{opt}::H_6-*ihfA*::SD_{opt}::*ihfB*	This study
pGH33rof	pBAD33 P_{BAD}::SD_{opt}::*rof*	This study
pGH33rof_{ext}	pBAD33 P_{BAD}::*rof* (extended)	This study
pGH33yaeP	pBAD33 P_{BAD}::SD_{opt}::*yaeP*	This study
pGH33yaeP-rof	pBAD33 P_{BAD}::*yaeP-rof*	This study
pGH33tilS	pBAD33 P_{BAD}::SD_{opt}::*tilS*	This study
pNDM220	mini-R1 *ts*, $lacI^q$ $P_{A1/O4/O3,}$ Amp^R	Laboratory collection
pSC228	pNDM220 $P_{A1/O4/O3}$::*mazE*	(Pedersen et al., 2002)
pKW220	mini-R1 *ts*, $lacI^q$ $P_{O5/O4}$, N-terminal FLAG-TEV-H_6 tag, Amp^R	Laboratory collection
pKW220rtcB	pKW220, $P_{O5/O4}$::FTH-*rtcB*	Kristoffer Winther
pNKBOR	R6K, mini-Tn10::*kan*, Kan^R	(Rossignol et al., 2001)
pMiniHimar	R6K, mini-mariner::*kan*, Kan^R	(Rubin et al., 1999)
pMG25::*frt::kan::frt*	pMG25, *frt::kan::frt*	Laboratory collection
pGH25lacZ	pMG25::*lacZ::frt::kan::frt*	This study
pKD3	R6K, Cml^R, Amp^R	(Datsenko and Wanner, 2000)
pKD46	pSC101 *ts*, P_{araB}::*γ-β-exo* (λ red), Amp^R	(Datsenko and Wanner, 2000)
pWRG100	pKD3, I-SceI recognition site, Cml^R, Amp^R	(Blank et al., 2011)
pWRG99	pKD46, P_{tetA}::*I-SceI*, Amp^R	(Blank et al., 2011)
pCP20	Rep *ts*, λ P_R::*flp*, λ cI857, Amp^R Cml^R	(Cherepanov and Wackernagel, 1995)

Table 3 Oligonucleotides used in this study

Oligonucleotide name	Sequence (5′ → 3′)
CLONING	
GH-ter-mfeI-cw	CCCCCCAATTGGGCATCAAATAAAACGAAAGGC
GH-ter-ecoRI-ccw	CCCCCGAATTCAGTTTGTAGAAACGCAAAAAGG
GH-ter-eagI-cw	CCCCCCGGCCGCTTTTGACGTCTTTGGCATCAAATAAAACGAAAGGCTC
RTCR1-optSD-f	CCCCCGGTACCGAATTCAAAATAAGGAGGAAAAAAAAATGCGTAAAACAGTGGCTTTTGG
RTCR2	CCCCCGAATTCAAGCTTCAGGTGCTGACGGACCACC
RTCR1-optSD-trunc-f	CCCCCGGTACCGAATTCAAAATAAGGAGGAAAAAAAAATGACCATGATTACCCCGCTTGATTTTCTTAAGTCCGG
GH-optSD-rpoN	CCCCCGGATCCGGTACCAAAATAAGGAGGAAAAAAAAATGAAGCAAGGTTTGCAAC
GH-rpoN-ccw	CCCCCGAATTCAAGCTTTCAAACGAGTTGTTTACGCTGGTTTG
rtcB-NcoI-BamHI-f	CCCCCCCATGGGGATCCAAAATAAGGAGGAAAAAAAAATGAATTACGAATTACTG
rtcA-SalI-XhoI-rv	CCCCCGTCGACCTCGAGTCATTCAATGCTC
RTCB1	CCCCCGAATTCAGACGATATTTCATTATCGG
RTCB2	CCCCCGGATCCATTTTCAGTGGTCAGTAATTCG
RTCB4	CCCCCGAATTCTTCATTATCGGTAAATTGG
RTCR6	CCCCCGAATTCGCATAAACAGAGTGTCGG
RTCR11	CCCCCGAATTCGCATCTTAGATATCCTTATAAAAG
RTCBmut9-f	GTTTTACGCATCTTAGGCGCTCTTATAAAAGACGAT
RTCBmut9-rv	ATCGTCTTTTATAAGAGCGCCTAAGATGCGTAAAAC
RTCBmut10-f	GTTTTACGCATCTTAGGTGTTCTTATAAAAGACGAT
RTCBmut10-rv	ATCGTCTTTTATAAGAACACCTAAGATGCGTAAAAC
RTCBmut11-f	GTTTTACGCATCTTAGTTTTGCTTATAAAAGACGAT
RTCBmut11-rv	ATCGTCTTTTATAAGCAAAACTAAGATGCGTAAAAC
RTCBmut3-f	ATCCTTATAAAAGACAGCGCTTCATTATCGGTAAA
RTCBmut3-rv	TTTACCGATAATGAAGCGCTGTCTTTTATAAGGAT
RTCBmut4-f	ATCCTTATAAAAGACAACACTTCATTATCGGTAAA
RTCBmut4-rv	TTTACCGATAATGAAGTGTTGTCTTTTATAAGGAT
RTCBmut5-f	ATCCTTATAAAAGACCAAAATTCATTATCGGTAAA
RTCBmut5-rv	TTTACCGATAATGAATTTTGGTCTTTTATAAGGAT
RTCBmut6-f	TTACGCATCTTAGGTGTTCTTATAAAAGACAACACTTCATTATCGGTA
RTCBmut6-rv	TACCGATAATGAAGTGTTGTCTTTTATAAGAACACCTAAGATGCGTAA
RTCBmut7-f	TTACGCATCTTAGGTGTTCTTATAAAAGACTAGAGTTCATTATCGGTA
RTCBmut7-rv	TACCGATAATGAACTCTAGTCTTTTATAAGAACACCTAAGATGCGTAA
RTCBmut8-f	TTACGCATCTTAGGTAGCCTTATAAAAGACGGTAGTTCATTATCGGTA
RTCBmut8-rv	TACCGATAATGAACTACCGTCTTTTATAAGGCTACCTAAGATGCGTAA
RTCBmut2-f	CTTAGATGGCCTTATAAAAG
RTCBmut2-rv	ACCGATAATGAAATCCCGTC
RTCBmutIHF1-f	AGATATAATCTTTTTTGAAACAATGAATTAAA
RTCBmutIHF1-rv	TTTAATTCATTGTTTCAAAAAAGATTATATCT
RTCBmutIHF2-f	CAATGAATTAAAAAATCAAGCTGCTTATTTAA
RTCBmutIHF2-rv	TTAAATAAGCAGCTTGATTTTTTAATTCATTG
RTCBmutIHF7-f	CTTTTTTAAAACAATGAATTGCAAAACCCAGCTGCTTATTTAAT

Oligonucleotide name	Sequence (5′ → 3′)
RTCBmutIHF7-rv	ATTAAATAAGCAGCTGGGTTTTGCAATTCATTGTTTTAAAAAAG
RTCBmutIHF3-f	CTTTTTTAAAACAATGAACTAAAAAATTAAGCTG
RTCBmutIHF3-rv	CAGCTTAATTTTTTAGTTCATTGTTTTAAAAAAG
RTCBmutIHF4-f	AGATATAATCTTTTTTGAAACAATGAATTAAAAAATCAAGCTGCTTATTTAAT
RTCBmutIHF4-rv	ATTAAATAAGCAGCTTGATTTTTTAATTCATTGTTTCAAAAAAGATTATATCT
RTCBmutIHF5-f	AGATATAATCTTTTTTGAAACAATGAACTAAAAAATCAAGCTGCTTATTTAAT
RTCBmutIHF5-rv	ATTAAATAAGCAGCTTGATTTTTTAGTTCATTGTTTCAAAAAAGATTATATCT
RTCBmutIHF8-f	AGATATAATCTTTTTTGGGAACAGCAAGTTGCAAAACCCAGCTGCTTATTTAAT
RTCBmutIHF8-rv	ATTAAATAAGCAGCTGGGTTTTGCAACTTGCTGTTCCCAAAAAGATTATATCT
GH-optSD-his6-rtcR-cw	CCCCGGTACCGAATTCAAAATAAGGAGGAAAAAAAAATGCATCACCATCACCATCACCGTAAAACAGTGGCTTTTGG
GH-optSD-his6-rtcR-trunc-cw	CCCCCGGTACCGAATTCAAAATAAGGAGGAAAAAAAAATGCATCACCATCACCATCACCTTGATTTTCTTAAGTCCGG
GH-optSD-his6-rtcR-Cterm-cw	CCCCCGGTACCGAATTCAAAATAAGGAGGAAAAAAAAATGCATCACCATCACCATCACTTTGCCACTAGCGGACGC
GH-optSD-his6-ihfA-cw	CCCCCGGTACCGAATTCAAAATAAGGAGGAAAAAAAAATGCATCACCATCACCATCACGCGCTTACAAAAGCTGA
GH-ihfA2-ccw	CCCCCTCTAGAGGATCCTTACTCGTCTTTGGGCGAAG
GH-optSD-ihfB2-cw	CCCCCTCTAGAGGATCCAAAATAAGGAGGAAAAAAAAATGACCAAGTCAGAATTGA
GH-ihfB2-ccw	CCCCCCTGCAGGTCGACTTAACCGTAAATATTGGCGCG
pRTCR12	CCCCCGAATTCGAAATATCGTCTTTTATAAG
pRTCR3	CCCCCGGATCCGACAAAGCCAAAAGCCACTG
pRTCR7	CCCCCGAATTCAAGCCAATTTACCGATAATG
pRTCR6	CCCCCGAATTCGATTATATCTTTACGTCCG
pRTCR1	CCCCCGAATTCTGTTTGCTGTCCTGATAATTG
pRTCRmut1-f	CAAAGCCAATTTACCGGGAAGGAAATATCGTCTTTT
pRTCRmut1-rv	AAAAGACGATATTTCCTTCCCGGTAAATTGGCTTTG
GH-optSD-rof-cw	CCCCCGGATCCGGTACCAAAATAAGGAGGAAAAAAAAATGAATGATACGTATCAACC
GH-rof-ccw	CCCCCGAATTCAAGCTTTCAGGATTCGCTTACCACC
GH-rof-cw	CCCCCGGATCCGGTACCCATACGCTGCGGCAAAC
GH-rof-rep15-ccw	CCCCCGAATTCAAGCTTGGCGAAACGCTGATTG
GH-optSD-yaeP-cw	CCCCCGGATCCGGTACCAAAATAAGGAGGAAAAAAAAGTGGAAAAATATTGTGAG
GH-yaeP-ccw	CCCCCGGATCCGGTACCAGTGCTGACAGGAGGGCA
GH-yaeP-cw	CCCCCGAATTCAAGCTTTCATTCATTGACATAATCGC
GH-optSD-tilS-cw-	CCCCCGGATCCGGTACCAAAATAAGGAGGAAAAAAAAATGACACTCACGCTCAATAG
GH-tilS-ccw	CCCCCGAATTCAAGCTTCCTGCCAGCATCATTTAAT
RECOMBINATION	
GH-pWRG100-rpoN-f	AAGTTTGCGACGTTTTAGCAGGAGAGTACGATTCTGAACCGCCTTACGCCCCGCCCTGC

Oligonucleotide name	Sequence (5′ → 3′)
GH-pWRG100-rpoN-rv	GTAATGTTGAGCTGCATAGTGTCTTCCTTATCGGTTGGGCTAGACTATATTACCCTGTT
GH-rpoN-scarless-1	GAAGTTTGCGACGTTTTAGCAGGAGAGTACGATTCTGAACCCCAACCGATAAGGAAGACACTATGCAGCTCAACATTACC
GH-rpoN-scarless-2	GGTAATGTTGAGCTGCATAGTGTCTTCCTTATCGGTTGGGGTTCAGAATCGTACTCTCCTGCTAAAACGTCGCAAACTTC
GH-lacZ-SD4-ecoRI-cw	CCCCCGAATTCAAGGAGGATAAATGACCATGATTACGGATTC
GH-lacZ-bamHI-ccw	CCCCCGGATCCTTATTTTTGACACCAGACC
GH-rtcBA-SD4lacZ-cw	TTAGTTTGATAGAAACAGATGGCGTAACGCGGGTGAGCATTGAATGATGCAAGGAGGATAAATGACCATG
GH-rtcBA-lacZ-ccw	ACGGCGTGTAAGTTTAGCCGGATAACGCGCCAGATCCGGCTTACATCTCTCCTCCTTAGTTCCTATTCC
GH-pWRG100-rtcR-f	ACCGATAATGAAATATCGTCTTTTATAAGGATATCTAAGCGCCTTACGCCCCGCCCTGC
GH-pWRG100-rtcR2-rv	CTTTGCCTGGCGGCAGATAGCGATAACGTGTTCCAGTTGCTAGACTATATTACCCTGTT
GH-rtcR2-scarless-1	AAGCCAATTTACCGATAATGAAATATCGTCTTTTATAAGGATATCTAAGCAACTGGAACACGTTATCGCTATCTGCCGCCAGGCAAAGTCGCTTTCCGC
GH-rtcR2-scarless-2	GCGGAAAGCGACTTTGCCTGGCGGCAGATAGCGATAACGTGTTCCAGTTGCTTAGATATCCTTATAAAAGACGATATTTCATTATCGGTAAATTGGCTT
GH-pKD3-rtcBA-f	ATTAAAAAATTAAGCTGCTTATTTAATTTTCTGGCACGACGTGTAGGCTGGAGCTGCTTC
GH-pKD3-rtcBA-rv	TTAGCCGGATAACGCGCCAGATCCGGCTTACATCTCTGCACATATGAATATCCTCCTTAG
GH-pKD3-rtcBAR-f	AAGCTGGCGTCCGGCTGCGGAAAGCGACTTTGCCTGGCGGGTGTAGGCTGGAGCTGCTTC
GH-pWRG100-rtcRmut-rv	CAGTACGGTACCGACAAAGCCAAAAGCCACTGTTTTACGCTAGACTATATTACCCTGTT
GH-rtcRmut-1	AAGCCAATTTACCGATAATGAAATATCGTCTTTTATAAGGATATCTAAGTAGCGTAAAACAGTGGCTTTTGGCTTTGTCGGTACCGTACTGGATTATGC
GH-rtcRmut-2	GCATAATCCAGTACGGTACCGACAAAGCCAAAAGCCACTGTTTTACGCTACTTAGATATCCTTATAAAAGACGATATTTCATTATCGGTAAATTGGCTT
COLONY PCR AND SEQUENCING	
GH-rtcA-orf-cw	GAGGTGGTTGCGGCACAG
GH-rtcA-flank-rv	ACAACACGCCCAGCAATTGC
pGH254amp-cw	TAGGGGTTCCGCGCACATTTCCC
pGH254-ccw	GCGATTAAGTTGGGTAAC
pGH254kan-cw	TTGAATATGGCTCATAACAC
pBAD-fw	CGCAACTCTCTACTGTTTCTC
pBAD-rv	CCGCTTCTGCGTTCTG
pMG25-f	GAGTGTTGACTTGTGAGCGG
pMG25-rv	GGAGTTCTGAGGTCATTACTGG
GH-ihfA-cw	GATTCCAGGCATCATTGAGG
GH-ihfA-ccw	GCAGAGCGGCCTTTTTAG
GH-ihfB-cw	ATTTGCCTTTAAGGAACCGG
GH-ihfB-ccw	GTGCTTTTCTCTCGTTCAAG

Oligonucleotide name	Sequence (5′ → 3′)
GH-rpoN-flank-cw	CACGCCTACAGAAATCTTAC
GH-rpoN-flank-ccw	GAACAACATAGACCTGGTTGATTC
RTCR1	CCCCCGGATCCGGTACCTACCGATAATGAAATATCGTC
delta-rtcB-f	ACGGTTGCAATTATCAGG
delta-rtcB-rv	CGCCATCCAGCGCAATC
Seq-rtcB-f2	GTAATGAGCCGAACTAAAG
GH-rtcR-flank-end-f	GCATGGTAGATGCCGTCTAC
pBR322-cw	CACTATCGACTACGC
pBR322-ccw	CGATGCGTCCGGCGT
pEG25-f	GCGTATCACGAGGCCCTTTC
ARBITRARY PCR	
OR7	GTGACACAGGAACACTTAACGGCTGACATGG
OR8	GTCGACGTCATATGGATCCG
OR9	CATTAAACGCGTATTCAGGCTGACCCTGCG
OR10	CCGGATTACAGGGATCGATC
ARB1A	CCACGCGTCGACTAGTACNNNNNNNNNNNGATAT
ARB2A	CCACGCGTCGACTAGTAC
EMSA	
RTCR9-f	CAGTACGGTACCGACAAAG
RTCB1-f	AGACGATATTTCATTATCGG
RTCB3-rv	GTTTGCTGTCCTGATAATTG
hicA-f2	CTGGAATCTTCCTTCCTGAT
hicA-rv	CTCGCTTTGTTTCACATCGCC
CRAC	
miRCat-33 linker	rApp-TGGAATTCTCGGGTGCCAAGG-ddC
L5Aa linker	InvddT-ACACrGrArCrGrCrUrCrUrUrCrCrGrArUrCrUrNrNrNrUrArArGrC
L5Ab linker	InvddT-ACACrGrArCrGrCrUrCrUrUrCrCrGrArUrCrUrNrNrNrArUrUrArGrC
L5Ad linker	InvddT-ACACrGrArCrGrCrUrCrUrUrCrCrGrArUrCrUrNrNrNrCrGrCrUrUrArGrC
L5Bb linker	InvddT-ACACrGrArCrGrCrUrCrUrUrCrCrGrArUrCrUrNrNrNrGrUrGrArGrC
miRCat RT	CCTTGGCACCCGAGAATT
PE_miRCat	CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGGCCTTGGCACCCGAGAATTCC
P5	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT

(A denotes adenosine; T denotes thymidine; C denotes cytidine; G denotes guanosine; U denotes uridine; N denotes random nucleotides; r denotes a ribonucleotide rather than a deoxyribonucleotide; App denotes an activated A; ddC denotes a di-deoxycytidine; InvddT denotes an inverted di-deoxythymidine)

2.3 Strain construction

GH01 (MG1655Δ*rpoN*), GH02 (TB28Δ*rpoN*). A scarless *rpoN* deletion was made in the chromosome of *E. coli* strains MG1655 and TB28 using the phage λ red recombination and mutagenesis method outlined in Blank et al. (2011). First, primers GH-pWRG100-rpoN-f and GH-pWRG100-rpoN-rv were used with plasmid pWRG100 as a template to generate a PCR

product with 40 bp homology extensions flanking the chloramphenicol resistance gene and the I-SceI restriction endonuclease recognition site. 200 ng of PCR product were electroporated (Section 2.6) into the relevant strain of *E. coli* and recombined into the chromosome by inducing expression of the phage λ red recombinase genes from pWRG99 with 0.2% arabinose. Successful recombinants were selected by growing on NA supplemented with 50 µg/ml chloramphenicol. In the second step, the chloramphenicol resistance gene and the I-SceI recognition site were replaced by double-stranded (ds)DNA fragments, which were generated by annealing equal amounts (500 nM) of oligonucleotides GH-rpoN-scarless-1 and GH-rpoN-scarless-2 in the presence of 150 mM NaCl at 95°C for 15 minutes, cooling to RT then desalting the resultant dsDNA on Illustra MicroSpin G-25 columns (GE Healthcare). The resulting dsDNA fragment consists of 40 bp of homologous sequence immediately upstream and 40 bp of homologous sequence immediately downstream of *rpoN* and, as with the first step of recombination, was introduced into the chromosome by expressing phage λ red recombinase genes from pWRG99. Successful recombinants with a large colony phenotype indicative of resistance to I-SceI were selected on NA supplemented with 100 µg/ml ampicillin (to maintain pWRG99) and 0.5 µg/ml anhydrotetracycline (to induce expression of I-SceI from pWRG99). Correct insertion of the chloramphenicol resistance marker and I-SceI site and subsequent generation of a scarless and markerless *rpoN* deletion was verified by colony PCR and sequencing using primers GH-rpoN-flank-cw and GH-rpoN-flank-ccw.

GH03 (TB28*rtcBA-lacZ*). The *lacZ* gene was recombined into the *E. coli* TB28 chromosome by phage λ red recombination as described by Datsenko and Wanner (2000). Primers GH-lacZ-SD4-ecoRI-cw and GH-lacZ-bamHI-ccw were used to amplify the *lacZ* gene with Shine-Dalgarno sequence AAGGAGGATAA (SD_4) from genomic MG1655 template DNA. The PCR product was digested with EcoRI and BamHI and ligated into pMG25::*frt::kan::frt*, a pMG25 variant carrying a kanamycin resistance gene flanked by *frt* sites, generating pGH25lacZ. Primers GH-rtcBA-SD4lacZ-cw and GH-rtcBA-lacZ-ccw were used with pGH25lacZ as template DNA to amplify the SD_4-*lacZ::frt::kan::frt* cassette flanked by 40 bp homology extensions for the end of the *rtcA* gene. 100 ng of PCR product was electroporated (Section 2.6) into *E. coli* strain TB28 and recombined into the chromosome by inducing expression of the phage λ red recombinase genes from pKD46 with 0.2% arabinose. Successful recombination events were screened for by growing selectively on NA with 25 µg/ml kanamycin. Subsequently, the antibiotic resistance marker was removed by *frt* site recombination using pCP20 (Cherepanov and Wackernagel, 1995). Correct insertion of *lacZ*

into the end of the *rtcBA* operon and subsequent removal of the antibiotic resistance gene was verified by colony PCR and sequencing using primers GH-rtcA-orf-cw and GH-rtcA-flank-rv.

GH04 (TB28*rtcBA-lacZ*Δ*ihfA::kan*). The *ihfA* gene was deleted from the chromosome of TB28*rtcBA-lacZ* (recipient strain) by P1 transduction, using the relevant strain of the Keio collection (Baba et al., 2006) as donor. Donor strain JW1702 was diluted from an ON culture into fresh LB medium containing 25 µg/ml kanamycin, 10 mM $MgSO_4$ and 10 mM $CaCl_2$ and grown exponentially at 37°C with aeration. The culture was incubated for 1 hour at 37°C upon addition of 100µl of stock P1 phage then transferred to 37°C with aeration until cell lysis occurred. The lysate was incubated for 10 minutes at room temperature (RT) with 0.1 volume of chloroform then centrifuged for 5 minutes at 5000 rpm and the supernatants (P1 lysates) collected. A single colony of recipient strain TB28*rtcBA-lacZ* was grown at 37°C with aeration to stationary phase. 25 mM $CaCl_2$ was added to the culture and incubated at 37°C with aeration for 20 minutes. 100 µl donor P1 lysate was mixed with 1 ml recipient culture and incubated at 37°C for 10 minutes, after which 500 µl LB supplemented with 10 mM sodium citrate was added. Following 60 minutes incubation at 37°C, cells were harvested, plated on NA containing 25 µg/ml kanamycin and 10 mM sodium citrate and incubated overnight at 37°C. The resulting colonies carry the kanamycin resistance gene in place of the *ihfA* gene. Insertion of the kanamycin resistance gene at the correct position was verified by colony PCR using primers GH-ihfA-cw and GH-ihfA-ccw.

GH05 (TB28*rtcBA-lacZ*Δ*ihfB::kan*). The *ihfB* gene was deleted from the chromosome of recipient strain TB28*rtcBA-lacZ* using donor strain JW0895 from the Keio collection by P1 phage transduction, as described for GH04. Insertion of the kanamycin resistance gene at the *ihfB* locus was verified by colony PCR using primers GH-ihfB-cw and GH-ihfB-ccw.

GH06 (MG1655Δ*rtcR*), GH07 (TB28Δ*rtcR*). A scarless *rtcR* deletion was made in MG1655 and TB28 using the phage λ red recombination and mutagenesis method outlined in Blank et al. (2011). The details of the two-step recombination are the same as described for GH01 and GH02. Primers GH-pWRG100-rtcR-f and GH-pWRG100-rtcR2-rv were used with template pWRG100 to generate a PCR product for step 1. For step 2, dsDNA was generated using primers GH-rtcR2-scarless-1 and GH-rtcR2-scarless-2, containing 40 bp of homologous sequence immediately upstream and 40 bp of homologous sequence at the 3′-end of *rtcR* (leaving the transcription terminator which overlaps the 3′-end of *rtcR* intact). Correct insertion of the chloramphenicol resistance marker and I-SceI site (step 1) and subsequent

generation of a scarless and markerless deletion (step 2) was verified by colony PCR and sequencing using primers RTCR1 and RTCR2.

GH08 (MG1655Δ*rtcB*). The *rtcB* gene was deleted from the chromosome of recipient strain MG1655 using donor strain JW3384 from the Keio collection by P1 phage transduction, as described for GH04. The antibiotic resistance marker replacing *rtcB* in the resulting colonies was removed by transformation of pCP20 and culturing according to the protocol described in Cherepanov and Wackernagel (1995). Insertion of the kanamycin resistance gene at the *rtcB* locus and its subsequent removal was verified by colony PCR using primers delta-rtcB-f and delta-rtcB-rv.

GH09 (MG1655Δ*rtcA*). The *rtcA* gene was deleted from the chromosome of recipient strain MG1655 using donor strain JW5688 from the Keio collection by P1 phage transduction, as described for GH04. The antibiotic resistance marker replacing *rtcA* in the resulting colonies was removed by transformation of pCP20 (Cherepanov and Wackernagel, 1995). Insertion of the kanamycin resistance gene at the *rtcA* locus and its subsequent removal was verified by colony PCR using primers Seq-rtcB-f2 and GH-rtcA-flank-rv.

GH10 (MG1655Δ*rtcBA*). The *rtcBA* operon was deleted from MG1655 by phage λ red recombination as described by Datsenko and Wanner (2000). Primers GH-pKD3-rtcBA-f and GH-pKD3-rtcBA-rv were used with plasmid pKD3 as a template to generate a PCR product with 40 bp homology extensions flanking a chloramphenicol resistance gene. 100 ng of PCR product was electroporated (Section 2.6) into the relevant strain of *E. coli* and recombined into the chromosome by inducing expression of the phage λ red recombinase genes from pKD46 with 0.2% arabinose. Successful recombination events were screened for by growing selectively on NA containing 50 µg/ml chloramphenicol. Subsequently, the antibiotic resistance marker was removed using pCP20 (Cherepanov and Wackernagel, 1995). Correct insertion and subsequent removal of the antibiotic resistance gene was verified by colony PCR and sequencing using primers delta-rtcB-f and GH-rtcA-flank-rv.

GH11 (MG1655Δ*rtcBAR*). The *rtcBAR* locus was deleted from MG1655 by phage λ red recombination as described for GH10, using primers GH-pKD3-rtcBAR-f and GH-pKD3-rtcBA-rv with template pKD3 to generate the PCR product. Correct insertion and subsequent removal of the antibiotic resistance gene was verified by colony PCR and sequencing using primers GH-RtcR-flank-end-f and GH-rtcA-flank-rv.

GH12 (MG1655*rtcR*^{M1X}), GH13 (TB28*rtcR*^{M1X}). Site-directed mutations of the *rtcR* start codon were made in the chromosome of MG1655 and TB28 using the phage λ red recombination and mutagenesis method outlined in Blank et al. (2011). The details of the two-step recombination are the same as described for GH01 and GH02. Primers GH-pWRG100-rtcR-f and GH-pWRG100-rtcRmut-rv were used with template pWRG100 to generate a PCR product for step 1. For step 2, primers GH-rtcRmut-1 and GH-rtcRmut-2 were used to generate dsDNA containing 40 bp of homologous sequence immediately upstream of *rtcR*, followed by the nucleotides to replace the *rtcR* start codon and 40 bp of homologous sequence immediately downstream of the *rtcR* start codon, producing an in-frame and scarless substitution of the *rtcR* start codon to a stop codon (ATG→TAG; M1X). Correct insertion of the chloramphenicol resistance marker and I-SceI site (step 1) and subsequent generation of a scarless and markerless deletion (step 2) was verified by colony PCR and sequencing using primers RTCR1 and RTCR2.

GH14 (SC28Δ*rtcB*). The *rtcB* gene was deleted from the chromosome of recipient strain SC28 using donor strain JW3384 from the Keio collection by P1 phage transduction, as described for GH04. The antibiotic resistance marker replacing *rtcB* in the resulting colonies was removed by transformation of pCP20 (Cherepanov and Wackernagel, 1995). Insertion of the kanamycin resistance gene at the *rtcB* locus and its subsequent removal was verified by colony PCR using primers delta-rtcB-f and delta-rtcB-rv.

2.4 Plasmid construction

2.4.1 Transcriptional fusions

All transcriptional fusion plasmids were verified by colony PCR and sequencing using primers pGH254amp-cw and pGH254-ccw for ampicillin-resistant derivatives of pOU254 and primers pGH254kan-cw and pGH254-ccw for kanamycin-resistant derivatives of pOU254K.

pGH254. Transcription terminators of the *rrnB* operon (ter$_{rrnB}$) were amplified from genomic *E. coli* K-12 MG1655 template DNA using primers GH-ter-mfeI-cw and GH-ter-ecoRI-ccw. The PCR product was digested with MfeI and EcoRI and ligated into the EcoRI-digested *lacZYA* transcriptional fusion vector pOU254. The resulting ampicillin resistant plasmid has two terminators upstream of the multiple cloning site (MCS).

pGH254K. Transcription terminators of the *rrnB* operon (ter_{rrnB}) were amplified using primers GH-ter-eagI-cw and GH-ter-ecoRI-ccw. The PCR product and *lacZYA* transcriptional fusion vector pOU254K was digested with EagI and EcoRI and ligated. The resulting kanamycin resistant plasmid has two terminators upstream of the MCS.

pGH254BA0. A fragment starting at -133 bp relative to the *rtcBA* transcription start site (+1) and containing the first 30 bp of *rtcB* was amplified with primers RTCB4 and RTCB2 using genomic *E. coli* MG1655 template DNA. The PCR product and pGH254 were digested with EcoRI and BamHI and ligated. The resulting plasmid carries a transcriptional *rtcBA-lacZYA* fusion known as BA0.

pGH254BA1. Primers RTCB1 and RTCB2 were used to amplify a fragment starting at -142 bp relative to the *rtcBA* transcription start site (+1) and containing the first 30 bp of *rtcB* from genomic *E. coli* MG1655 template DNA. The PCR product and pGH254 were digested with EcoRI and BamHI and ligated. The resulting plasmid carries a transcriptional *rtcBA-lacZYA* fusion known as BA1.

pGH254BA2. A fragment starting at -164 bp relative to the *rtcBA* transcription start site (+1) and containing the first 30 bp of *rtcB* was amplified with primers RTCR11 and RTCB2 using genomic *E. coli* MG1655 template DNA. The PCR product and pGH254 were digested with EcoRI and BamHI and ligated. The resulting plasmid carries a transcriptional *rtcBA-lacZYA* fusion known as BA2.

pGH254KBA2. As described for pGH254BA2, except the PCR product was ligated into pGH254K.

pGH254BA3. Primers RTCR6 and RTCB2 were used to amplify a fragment starting at -262 bp relative to the *rtcBA* transcription start site (+1) and containing the first 30 bp of *rtcB* from genomic *E. coli* MG1655 template DNA. The PCR product and pGH254 were digested with EcoRI and BamHI and ligated. The resulting plasmid carries a transcriptional *rtcBA-lacZYA* fusion known as BA3.

pGH254BA3-IR_{M1}. To introduce mutations into the putative RtcR binding site *rtcO*, primers RTCBmut9-rv and RTCBmut9-f carrying mutations in the inverted repeat (IR) were combined with primers RTCR6 and RTCB2 respectively, using pGH254BA3 template DNA. 0.2 pmol of each of the PCR fragments, which are homologous at the terminals carrying mutations, were combined as templates for a second PCR reaction with primers RTCR6 and RTCB2. The PCR product and pGH254 were digested with EcoRI and BamHI and ligated.

The resulting plasmid carries a transcriptional *rtcBA-lacZYA* fusion known as BA3-IR_{M1}, which has ATATC to GCGCT substitutions in the first half of the *rtcO*.

pGH254BA3-IR_{M2}. As described for pGH254BA3-$IR_{M1,}$ except that primers RTCBmut10-rv and RTCBmut10-f carrying mutations in *rtcO* were combined with primers RTCR6 and RTCB2 respectively in the first round of PCR. The resulting plasmid carries a transcriptional *rtcBA-lacZYA* fusion known as BA3-IR_{M2}, which has substitutions at the 1st (A to G), 3rd (A to G) and 5th (C to T) positions in the first half of the IR.

pGH254BA3-IR_{M3}. As described for pGH254BA3-$IR_{M1,}$ except that primers RTCBmut11-rv and RTCBmut11-f carrying mutations in *rtcO* were combined with primers RTCR6 and RTCB2 respectively in the first round of PCR. The resulting plasmid carries a transcriptional *rtcBA-lacZYA* fusion known as BA3-IR_{M3}, which has substitutions at the 1st (A to T), 3rd (A to T) and 5th (C to G) positions in the first half of the IR.

pGH254BA3-IR_{M4}. As described for pGH254BA3-$IR_{M1,}$ except that primers RTCBmut3-rv and RTCBmut3-f carrying mutations in *rtcO* were combined with primers RTCR6 and RTCB2 respectively in the first round of PCR. The resulting plasmid carries a transcriptional *rtcBA-lacZYA* fusion known as BA3-IR_{M4}, which has GATAT to AGCGC substitutions in the second half of the IR.

pGH254BA3-IR_{M5}. As described for pGH254BA3-$IR_{M1,}$ except that primers RTCBmut4-rv and RTCBmut4-f carrying mutations in *rtcO* were combined with primers RTCR6 and RTCB2 respectively in the first round of PCR. The resulting plasmid carries a transcriptional *rtcBA-lacZYA* fusion known as BA3-IR_{M5}, which has substitutions at the 1st (G to A), 3rd (T to C) and 5th (T to C) positions in the second half of the IR.

pGH254BA3-IR_{M6}. As described for pGH254BA3-$IR_{M1,}$ except that primers RTCBmut5-rv and RTCBmut5-f carrying mutations in *rtcO* were combined with primers RTCR6 and RTCB2 respectively in the first round of PCR. The resulting plasmid carries a transcriptional *rtcBA-lacZYA* fusion known as BA3-IR_{M6}, which has substitutions at the 1st (G to C), 3rd (T to A) and 5th (T to A) positions in the second half of the IR.

pGH254BA3-IR_{M7}. As described for pGH254BA3-$IR_{M1,}$ except that primers RTCBmut6-rv and RTCBmut6-f carrying mutations in *rtcO* were combined with primers RTCR6 and RTCB2 respectively in the first round of PCR. The resulting plasmid carries a transcriptional *rtcBA-lacZYA* fusion known as BA3-IR_{M7}, which has substitutions at the 1st (A to G), 3rd (A

to G) and 5th (C to T) positions in the first half and at the 1st (G to A), 3rd (T to C) and 5th (T to C) positions in the second half of the IR.

pGH254BA3-IR_{M8}. As described for pGH254BA3-$IR_{M1,}$ except that primers RTCBmut7-rv and RTCBmut7-f carrying mutations in *rtcO* were combined with primers RTCR6 and RTCB2 respectively in the first round of PCR. The resulting plasmid carries a transcriptional *rtcBA-lacZYA* fusion known as BA3-IR_{M8}, which has substitutions at the 1st (A to G), 3rd (A to G) and 5th (C to T) positions in the first half and at the 1st (G to T), 3rd (T to G) and 5th (T to G) positions in the second half of the IR.

pGH254BA3-IR_{M9}. As described for pGH254BA3-$IR_{M1,}$ except that primers RTCBmut8-rv and RTCBmut8-f carrying mutations in *rtcO* were combined with primers RTCR6 and RTCB2 respectively in the first round of PCR. The resulting plasmid carries a transcriptional *rtcBA-lacZYA* fusion known as BA3-IR_{M9}, which has substitutions at the 1st (A to G) and 4th (T to G) positions in the first half and at the 2nd (A to G) and 5th (T to G) positions in the second half of the IR.

pGH254BA3-IR_{M10}. As described for pGH254BA3-$IR_{M1,}$ except that primers RTCBmut2-rv and RTCBmut2-f carrying mutations in *rtcO* were combined with primers RTCR6 and RTCB2 respectively in the first round of PCR. The resulting plasmid carries a transcriptional *rtcBA-lacZYA* fusion known as BA3-IR_{M10}, which has substitutions at the 3rd (A to G) and 4th (T to G) positions in the first half and at the 2nd (A to G) and 3rd (T to G) positions in the second half of the IR.

pGH254BA3-IHF_{M1}. To introduce mutations into the putative IHF binding sites, primers RTCBmutIHF1-rv and RTCBmutIHF1-f carrying a mutation in IHF site 1 (*ihf1*) were combined with primers RTCR6 and RTCB2 respectively, using pGH254BA3 template DNA. 0.2 pmol of each of the PCR fragments, which are homologous at the terminals carrying the mutation, were combined as templates for a second PCR reaction with primers RTCR6 and RTCB2. The PCR product and pGH254 were digested with EcoRI and BamHI and ligated. The resulting plasmid carries a transcriptional *rtcBA-lacZYA* fusion known as BA3-IHF_{M1}, which has an A to G substitution in *ihf1*.

pGH254BA3-IHF_{M2}. As described for pGH254BA3-$IHF_{M1,}$ except that primers RTCBmutIHF2-rv and RTCBmutIHF2-f carrying mutations in IHF site 2 (*ihf2*) were combined with primers RTCR6 and RTCB2 respectively in the first round of PCR. The

resulting plasmid carries a transcriptional *rtcBA-lacZYA* fusion known as BA3-IHF$_{M2}$, which has an T to C substitution in *ihf2*.

pGH254BA3-IHF$_{M3}$. As described for pGH254BA3-IHF$_{M1,}$ except that primers RTCBmutIHF7-rv and RTCBmutIHF7-f carrying mutations in *ihf2* were combined with primers RTCR6 and RTCB2 respectively in the first round of PCR. The resulting plasmid carries a transcriptional *rtcBA-lacZYA* fusion known as BA3-IHF$_{M3}$, which has AA----TTA to GC----CCC substitutions in *ihf2*.

pGH254BA3-IHF$_{M4}$. As described for pGH254BA3-IHF$_{M1,}$ except that primers RTCBmutIHF3-rv and RTCBmutIHF3-f carrying mutations in the IHF binding sites were combined with primers RTCR6 and RTCB2 respectively in the first round of PCR. The resulting plasmid carries a transcriptional *rtcBA-lacZYA* fusion known as BA3-IHF$_{M4}$, which has a T to C substitution in the overlapping region of *ihf1* and *ihf2*.

pGH254BA3-IHF$_{M5}$. As described for pGH254BA3-IHF$_{M1,}$ except that primers RTCBmutIHF4-rv and RTCBmutIHF4-f carrying mutations in the IHF binding sites were combined with primers RTCR6 and RTCB2 respectively in the first round of PCR. The resulting plasmid carries a transcriptional *rtcBA-lacZYA* fusion known as BA3-IHF$_{M5}$, which has an A to G substitution in *ihf1* and a T to C substitution in *ihf2*.

pGH254BA3-IHF$_{M6}$. As described for pGH254BA3-IHF$_{M1,}$ except that primers RTCBmutIHF5-rv and RTCBmutIHF5-f carrying mutations in the IHF binding sites were combined with primers RTCR6 and RTCB2 respectively in the first round of PCR. The resulting plasmid carries a transcriptional *rtcBA-lacZYA* fusion known as BA3-IHF$_{M6}$, which has an A to G substitution in *ihf1*, a T to C substitution in *ihf2* and a T to C substitution in the overlapping region of *ihf1* and *ihf2*.

pGH254BA3-IHF$_{M7}$. As described for pGH254BA3-IHF$_{M1,}$ except that primers RTCBmutIHF8-rv and RTCBmutIHF8-f carrying mutations in the IHF binding sites were combined with primers RTCR6 and RTCB2 respectively in the first round of PCR. The resulting plasmid carries a transcriptional *rtcBA-lacZYA* fusion known as BA3-IHF$_{M7}$, which has TAA----ATG--T-AA----TTA to GGC----GCA--C-GC----CCC substitutions in *ihf1* and *ihf2*.

pGH254R0. A fragment beginning 30 bp upstream of the start codon of *rtcR* and containing the first 30 bp of *rtcR* was amplified with primers pRTCR12 and pRTCR3 using genomic *E. coli* MG1655 template DNA. The PCR product and pGH254 were digested with EcoRI and

BamHI and ligated. The resulting plasmid carries a transcriptional *rtcR-lacZYA* fusion known as R0.

pGH254R1. A fragment starting 49 bp upstream of *rtcR* and containing the first 30 bp of *rtcR* was amplified with primers pRTCR7 and pRTCR3 using MG1655 template DNA. The PCR product and pGH254 were digested with EcoRI and BamHI and ligated. The resulting plasmid carries a transcriptional *rtcR-lacZYA* fusion known as R1.

pGH254R2. A fragment starting 87 bp upstream of *rtcR* and containing the first 30 bp of *rtcR* was amplified with primers pRTCR6 and pRTCR3 using MG1655 template DNA. The PCR product and pGH254 were digested with EcoRI and BamHI and ligated. The resulting plasmid carries a transcriptional *rtcR-lacZYA* fusion known as R2.

pGH254R3. A fragment starting 169 bp upstream of *rtcR* and containing the first 30 bp of *rtcR* was amplified with primers pRTCR1 and pRTCR3 using MG1655 template DNA. The PCR product and pGH254 were digested with EcoRI and BamHI and ligated. The resulting plasmid carries a transcriptional *rtcR-lacZYA* fusion known as R3.

pGH254KR3. As described for pGH254R3, except the PCR product was ligated into pGH254K.

pGH254R3$_{mut10}$. To introduce mutations into the hypothetical -10 site of *rtcR*, primers pRTCRmut1-rv and pRTCRmut1-f were combined with primers pRTCR1 and pRTCR3 respectively, using pGH254R3 template DNA. 0.2 pmol of each of the PCR fragments, which are homologous at the terminals carrying the mutation, were combined as templates for a second PCR reaction with primers pRTCR1 and pRTCR3. The PCR product and pGH254 were digested with EcoRI and BamHI and ligated. The resulting plasmid carries a transcriptional *rtcR-lacZYA* fusion known as R3$_{mut10}$, which has GATAAT to GGGAAG substitutions in the hypothetical -10 of *rtcR*.

2.4.2 Overexpression of *rtc* genes

All genes were expressed from an optimal Shine-Dalgarno sequence (UAAGGAGG) with a spacing of 8 nucleotides as described by Ringquist et al. (1992), which was encoded in the forward primer, unless otherwise stated.

pGH332R. The complete open reading frame (ORF) of *rtcR* was PCR amplified using genomic *E. coli* MG1655 template DNA with primers RTCR1-optSD-f and RTCR2. The

PCR product was digested with EcoRI and HindIII and ligated into pBAD322K, producing a plasmid expressing wild-type (WT) RtcR with an optimised Shine-Dalgarno (SD_{opt}) sequence from the AraC-regulated P_{BAD} promoter. Correct clones were verified by colony PCR and sequencing using pBAD-fw and pBAD-rv.

pGH332CR. As described for pGH332R, except the PCR product was ligated into pBAD322C.

pGH332RΔN. Oligonucleotides RTCR1-optSD-trunc-f and RTCR2 were used to amplify a truncated form of *rtcR* (*rtcRΔN*), which lacks the first 534 nucleotides of *rtcR* (Genschik et al., 1998), from MG1655 template DNA. The PCR product was digested with EcoRI and HindIII and ligated into pBAD322K. The resulting plasmid expresses an N-terminally truncated RtcR with an SD_{opt} sequence from AraC-regulated P_{BAD} promoter. Correct clones were verified by colony PCR and sequencing using pBAD-fw and pBAD-rv.

pGH332CRΔN. As described for pGH332R, except the PCR product was ligated into pBAD322C.

pGH25rtcBA. Oligonucleotides rtcB-NcoI-BamHI-f and rtcA-SalI-XhoI-rv were used to amplify *rtcBA* from MG1655 template DNA. The PCR product and vector pMG25 were digested with BamHI and SalI and ligated. The resulting plasmid expresses RtcB and RtcA from the LacI-regulated $P_{A1/O4/O3}$ promoter. Correct clones were verified by colony PCR and sequencing using pMG25-f and pMG25-rv.

2.4.3 Expression of tagged *rtc* genes

All genes were expressed from an optimal Shine-Dalgarno sequence (UAAGGAGG) with an optimal spacing of 8 nucleotides as described by Ringquist et al. (1992), which was encoded in the forward primer, unless otherwise stated.

pGH33H_6R. Primers GH-optSD-his6-rtcR-cw and RTCR2 were used to amplify *rtcR* from genomic *E. coli* MG1655 template DNA. The PCR product and vector pBAD33 were digested with KpnI and HindIII and ligated. The resulting plasmid expresses His_6-tagged RtcR with an SD_{opt} sequence from the AraC-regulated P_{BAD} promoter. Correct clones were verified by colony PCR and sequencing using pBAD-fw and pBAD-rv.

pGH33H_6RΔN. Primers GH-optSD-his6-rtcR-trunc-cw and RTCR2 were used to amplify *rtcRΔN* from MG1655 template DNA. The PCR product and vector pBAD33 were digested

with KpnI and HindIII and ligated. The resulting plasmid expresses His_6-tagged RtcRΔN with an SD_{opt} sequence from the AraC-regulated P_{BAD} promoter. Correct clones were verified by colony PCR and sequencing using pBAD-fw and pBAD-rv.

pGH25H_6RC. Primers GH-optSD-his6-rtcR-Cterm-cw and RTCR2 were used to amplify truncated *rtcR* lacking the first 1272 bp (*rtcR*$_{C\text{-}term}$) from MG1655 template DNA. The PCR product and vector pMG25 were digested with EcoRI and HindIII and ligated. The resulting plasmid expresses His_6-tagged $RtcR_{C\text{-}term}$ with an SD_{opt} sequence from the LacI-regulated $P_{A1/O4/O3}$ promoter. Correct clones were verified by colony PCR and sequencing using pMG25-f and pMG25-rv.

2.4.4 Others

All genes were expressed from an optimal Shine-Dalgarno sequence (UAAGGAGG) with an optimal spacing of 8 nucleotides as described by Ringquist et al. (1992), which was encoded in the forward primer unless otherwise stated.

pGH25rpoN. Primers GH-optSD-rpoN-cw and GH-rpoN-ccw were used to ampify the *rpoN* gene from genomic *E. coli* MG1655 template DNA. The PCR product and vector pMG25 were digested with BamHI and HindIII and ligated. The resulting plasmid expresses σ^{54} with an SD_{opt} sequence from the LacI-regulated $P_{A1/O4/O3}$ promoter. Correct clones were verified by colony PCR and sequencing using pMG25-f and pMG25-rv.

pGH33H_6IHF. Oligonucleotides GH-optSD-his6-ihfA-cw and GH-ihfA2-ccw were used to amplify the *ihfA* gene from MG1655 template DNA. The PCR product and vector pBAD33 were digested with KpnI and XbaI and ligated. Next, primers GH-optSD-ihfB2-cw and GH-ihfB2-ccw were used to amplify *ihfB* from MG1655 template DNA. The PCR product was digested with XbaI and PstI and ligated into pBAD33::P_{BAD}::SD_{opt}::H_6-*ihfA* digested with the same restriction enzymes. The resulting plasmid expresses His_6-tagged IHFα with an SD_{opt} sequence and IHFβ with an SD_{opt} sequence from the AraC-regulated P_{BAD} promoter. Correct clones were verified by colony PCR and sequencing using pBAD-fw and pBAD-rv.

pGH251. Primers pRTCR1 and RTCR11 were used to amplify the *rtcBA-rtcR* intergenic sequence from MG1655 template DNA. The PCR product and vector pMG25 were digested with BamHI and EcoRI and ligated. Correct clones were verified by colony PCR and sequencing using pMG25-f and pMG25-rv.

pGH252. Primers pRTCR1 and RTCB4 were used to amplify the *rtcBA-rtcR* intergenic sequence excluding *rtcO* from MG1655 template DNA. The PCR product and vector pMG25 were digested with BamHI and EcoRI and ligated. Correct clones were verified by colony PCR and sequencing using pMG25-f and pMG25-rv.

pGH33rof. Oligonucleotides GH-optSD-rof-cw and GH-rof-ccw were used to amplify the *rof* gene from MG1655 template DNA. The PCR product and vector pBAD33 were digested with KpnI and HindIII and ligated. The resulting plasmid expresses Rof with an SD_{opt} sequence from the AraC-regulated P_{BAD} promoter. Correct clones were verified by colony PCR and sequencing using pBAD-fw and pBAD-rv.

pGH33rof$_{ext}$. Primers GH-rof-cw and GH-rof-rep15-ccw were used to amplify the *rof* gene and the intergenic region between *rof* and *tilS* from MG1655 template DNA. The PCR product and vector pBAD33 were digested with KpnI and HindIII and ligated. The plasmid produced expresses Rof with its native SD sequence from the AraC-regulated P_{BAD} promoter. Correct clones were verified by colony PCR and sequencing using pBAD-fw and pBAD-rv.

pGH33yaeP. Primers GH-optSD-yaeP-cw and GH-yaeP-ccw were used to amplify the *yaeP* gene from MG1655 template DNA. The PCR product and vector pBAD33 were digested with KpnI and HindIII and ligated. The resulting plasmid expresses YaeP with an SD_{opt} sequence from the AraC-regulated P_{BAD} promoter. Correct clones were verified by colony PCR and sequencing using pBAD-fw and pBAD-rv.

pGH33yaeP-rof. Primers GH-yaeP-cw and GH-rof-ccw were used to amplify the *yaeP* and *rof* genes from MG1655 template DNA. The PCR product and vector pBAD33 were digested with KpnI and HindIII and ligated. The resulting plasmid expresses YaeP and Rof with their native SD sequences from the AraC-regulated P_{BAD} promoter. Correct clones were verified by colony PCR and sequencing using pBAD-fw and pBAD-rv.

pGH33tilS. Oligonucleotides GH-optSD-tilS-cw and GH-tilS-ccw were used to amplify the *tilS* gene from MG1655 template DNA. The PCR product and vector pBAD33 were digested with KpnI and HindIII and ligated. The resulting plasmid expresses TilS with an SD_{opt} sequence from the AraC-regulated P_{BAD} promoter. Correct clones were verified by colony PCR and sequencing using pBAD-fw and pBAD-rv.

2.5 Preparation and transformation of TSB chemically competent cells

The protocol was adapted from (Chung and Miller, 1988). An ON culture inoculated from a single colony of an appropriate bacterial strain was diluted into fresh LB medium and grown at 37°C with aeration until OD_{600} 0.3-0.4. Cells were harvested by centrifugation for 5 minutes at 5000 rpm and resuspended in 0.2 volume cold sterile TSB (10% polyethylene glycol 3350, 5% dimethylsulphoxide, 20 mM $MgCl_2$ solution, diluted in LB broth). 20-200 ng plasmid DNA was mixed with a 200 µl aliquot of TSB chemically-competent cells and incubated on ice for 30 minutes. Cells were heat-shocked at 42°C for 45 seconds and recovered by addition of 800 µl LB media, followed by phenotypic expression at 37°C for 60 minutes. Cells were centrifuged and pellets resuspended in approximately 150 µl of supernatant, plated on NA containing appropriate selection and incubated overnight at 37°C**.**

2.6 Preparation and electroporation of competent cells

An ON bacterial culture was diluted into fresh media and grown at 37°C with aeration until reaching OD_{600} 0.4-0.5. Cells were rapidly chilled in ice-water slurry for 5-10 minutes then harvested at 5000 rpm and 4°C for 5 minutes. The pellet was washed three times with the original culture volume of sterile ice-cold H_2O. The pellet was finally resuspended in a small volume of ice-cold H_2O (1 ml per 50 ml original culture) and separated into 50 µl aliquots. 10-200 ng DNA was mixed with an aliquot of electrocompetent cells on ice. The mixture was added to a pre-cooled 0.2 cm Gene Pulser cuvette (BioRad), pulsed for a few seconds (2.5 kV; 200 ohms; 25 µF) and recovered by addition of 800 µl LB media, followed by phenotypic expression at 37°C with aeration for 60 minutes. Cells were harvested and the pellet was resuspended in approximately 150 µl of supernatant, plated on NA containing appropriate selection and incubated overnight at 37°C.

2.7 Genomic DNA purification

100 ml of ON culture of the relevant strain was centrifuged at 5000 rpm for 20 minutes. The cell pellet was washed once with 25 ml 0.9% NaCl then resuspended in 20 ml lysis buffer (50 mM Tris-HCl (pH 8), 20 mM EDTA (pH 8), 400 µg/ml lysozyme, 10 µg/ml RNase A) and incubated at 37°C for 10 minutes. Following addition of 1ml of 10% SDS, the lysate was incubated at 37°C for a further 20 minutes. An equal volume of phenol (pH 8) was added, vortexed thoroughly and centrifuged at 5000 rpm for 10 minutes. The aqueous phase of the

biphasic mixture was transferred to a fresh vessel and phenol extraction repeated twice, followed by two chloroform extractions. DNA was precipitated upon addition of 0.6 volume of 2-propanol with gentle mixing then 20 minutes incubation at RT. DNA was harvested by centrifugation at 5000 rpm for 20 minutes at 4°C. The DNA pellet was washed twice with 5 ml ice-cold 70% ethanol and once with 5 ml ice-cold 96% ethanol after which the pellet was air dried and dissolved in an appropriate volume of distilled (d)H_2O.

Rapid, small-scale chromosomal DNA extraction was performed according to the method developed by Cheng and Jiang (2006), following the recommendations for Gram-negative bacteria. The Bacterial Genomic DNA Purification kit (EdgeBio) was also routinely used for chromosomal DNA purification, following the instructions provided by the manufacturer.

2.8 Genomic DNA library construction

Chromosomal DNA from *E. coli* K-12 TB28 was partially digested with Sau3AI (Thermo Fisher Scientific) using a protocol optimised from Sambrook et al. (1989). First, a scaled-up digest mix of 40 µg of DNA, 20 µl of Sau3AI digest buffer and dH_2O to a total volume of 200 µl was made. 1 µl of 0.25 U/µl stock Sau3AI was added to a 40 µl aliquot of the digest mix and two-fold serial dilutions were made using 20 µl aliquots of the digest mix. The final 20 µl aliquot was used as a control lacking Sau3AI. The digests were incubated at 37°C for 50 minutes then Sau3AI was inactivated according to the manufacturer's instructions. Partially digested DNA was resolved on a 1% agarose gel stained with ethidium bromide and DNA between 2 and 6 Kb was purified as described in Section 2.2. The 2-6 Kb fragments were ligated at a 4:3 insert to vector ratio into 100 ng BamH1-digested and Alkaline Phosphatase-treated pBR322 and incubated ON at 4°C. T4 DNA ligase was inactivated according to the manufacturer's instructions (Thermo Fisher Scientific). 1 µl of the resulting pBR322 library was used per 200 µl transformation reaction (see Section 2.5). Equivalent amounts of pBR322 ligated alone were also transformed in parallel, to gauge to digestion and ligation efficiency.

2.9 Mapping transposon insertion sites

Arbitrary PCR was used to map mini-Tn10 transposon insertion sites. The method was optimised from (Das et al., 2005). In the first round of PCRs, transposon-specific primers (Table 3; OR7 and 9) were paired in separate reactions with an arbitrary primer (ARB1A), and a small amount of colony providing template DNA. ARB1A binds promiscuously and

any annealing in the region of the transposon insertion produced an amplicon with the transposon-specific primers, which was purified. In the second round, nested primers for the transposon-specific primers (OR8 and OR10 respectively) and a primer for the non-arbitrary flanking sequence at the 5′-end of ARB1A (ARB2A) were paired to specifically amplify PCR product from the first round. The resulting PCR products were purified and sequenced.

Another method used for mapping transposition sites exploited the self-cloning properties of the transposon (Rossignol et al., 2001). Briefly, genomic DNA was extracted from each strain (Cheng and Jiang, 2006), digested with a restriction enzyme unable to cut within the transposon (MfeI and EcoRI were used) and ligated, producing a collection of circular genomic DNA from each strain. This was transformed into *E. coli* DH5α λpir and plated selectively with kanamycin. Resulting colonies carried circularised DNA containing the transposon, which were stably propagated owing to the presence of a conditional origin of replication, R6K, in the transposon and R6K π replicase in DH5α λpir (Rossignol et al., 2001). Subsequently, plasmids were purified and sequenced with transposon-specific primers OR8 and OR10.

2.10 Beta-galactosidase assay

ON cultures inoculated from single bacterial colonies were diluted to OD_{600} 0.05 in fresh LB media supplemented as appropriate with antibiotics and cultured at 37°C with aeration to OD_{600} 0.5. The cultures were again diluted to OD_{600} 0.05 in supplemented LB medium and cultured at 37°C with aeration. 0.5 ml samples of culture were generally taken between OD_{600} 0.1 and 1 as stated and diluted as necessary. Where appropriate, protein expression from various plasmids was induced by addition of 0.2% arabinose or 0.1-1 mM IPTG as indicated. β-galactosidase activity of the *lacZ* fusions in the samples taken were assayed as described by Miller (1972) using the chloroform and SDS protocol variant. Experiments were carried out in triplicate with two technical repeats per sample.

2.11 Total RNA purification

Total RNA purification was performed using the RNeasy Mini kit (Qiagen). The manufacturer's recommendations were adapted to optimise RNA extraction from bacteria. 5 ml bacterial culture was harvested and lysed by addition of 100 µl TE buffer (10 mM Tris-HCl (pH 8), 1 mM EDTA (pH 8) containing 1 mg/ml lysozyme. Following 3 minutes incubation at RT, 350 µl buffer RLT (Qiagen) supplemented with 1% β-mercaptoethanol was

added to the sample, vortexed for 1 minute and centrifuged at 13,000 rpm for 2 minutes. The supernatant was transferred to a fresh tube and 250 µl of ice-cold 100% ethanol was added and thoroughly homogenised by vortexing. The mixture was added to an RNeasy Spin column, centrifuged for 30 seconds and washed twice with 500 µl buffer RW1 (Qiagen). RNA was treated on the column with RNase-free DNase I (Qiagen) according to the manufacturer's instructions for 60 minutes at RT. The column was washed once with 350 µl buffer RW1 and twice with 500 µl buffer RPE and RNA was eluted in 50 µl RNase-free H_2O. The integrity of purified RNA was verified on a 1% agarose gel stained with ethidium bromide and quantified using a NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific).

2.12 Quantitative PCR

Following total RNA extraction from cell samples (Section 2.11), 1 µg of RNA was reverse transcribed to cDNA using the High Capacity cDNA Reverse Transcription kit (Applied Biosystems) according the provided instructions. The resulting cDNA was diluted 100-fold in sterile dH_2O, of which 6 µl was added to 10 µl 2X GoTaq qPCR Master Mix (Qiagen) and 2 µl each of forward and reverse primers from a 10 µM stock. The oligonucleotides used for quantitative (q)PCR are listed in Table 1Table 4. cDNA was amplified using either a Rotor-Gene Q Real-Time PCR Cycler (Qiagen) or a C1000 Thermal Cycler (BioRad) in 40 cycles of 95°C for 10 seconds and 60°C for 30 seconds, preceded by 3 minutes at 95°C. Quantification was performed by the associated software and the data analysed using the Pfaffl method of relative quantification (Pfaffl, 2001). The reference genes used for normalisation were *tatA*, *rpoB*, *tus* and *gyrA*. Primer efficiencies required for accurate relative quantification were calculated by amplification of 10-fold serial dilutions of cDNA and were between 91 and 104%.

Table 4 Oligonucleotides used in this study for qPCR

Oligonucleotide name	Sequence (5′ → 3′)	Target gene
qRtcB2-f	TGGCCGTTGTAAACGTGATA	*rtcB*
qRtcB2-rv	CAGCGTTCCCAGGTGTTTAT	
qRtcA-f	GACCAACTGGTGCTACCGAT	*rtcA*
qRtcA-rv	GCGTTACGCCATCTGTTTCT	
qRpoN-f	GCAACTCAGGCTTAGCCAAC	*rpoN*
qRpoN-rv	TCCAGCGTTTCACTGTCTTG	
qTatA-f	TGAGCGATGATGAACCAAAG	*tatA*
qTatA-rv	TCGCGTCTTCTGTTTTAGCC	

Oligonucleotide name	Sequence (5′ → 3′)	Target gene
qRpoB-f	TCCGTATTCCCGATTCAGAG	*rpoB*
qRpoB-rv	TCACCAGACGCAGTTTAACG	
qTus-f	TACCTTTCGCCAGATGGAAC	*tus*
qTus-rv	GCGATTTTCGGACTGTTGTT	
qGyrA-f	TGATGGAAGTGATCCGTGAA	*gyrA*
qGyrA-rv	AACGTAGCCCTGGTGAGAGA	
qRplL-f	GGCGCTAACAAAGTTGCTGT	*rplL*
qRplL-rv	CTTTGCTCACGCCTTCTTTC	
qRplD-f	CCGAACTGGTACGTCAGGAT	*rplD*
qRplD-rv	ATGCCAGCATCTCCTCAACT	
qRplM-f	GGCAACAAGCGTACTGACAA	*rplM*
qRplM-rv	CGTTACCCGCGTAAACTTTC	
qRpmA-f	GGTAGCATCATCGTTCGTCA	*rpmA*
qRpmA-rv	GGTTTTTCGGGCCTTTAACT	
qRplU-f	GAAGTGCTGATGATCGCAAA	*rplU*
qRplU-rv	CAGTGAACCACTGACGATGG	
qRpsB-f	AAACCGTTCGTCAGTCCATC	*rpsB*
qRpsB-rv	ACCGCCCATGTCTTTGATAC	
qRpsJ-f	TCCGTACTCACTTGCGTCTG	*rpsJ*
qRpsJ-rv	GGCTGATCTGCACGTCTACA	
qGcvB-f	GAATGCGTGTTCTGGTGAACT	*gcvB*
qGcvB-rv	GCAATTAGGCGGTGCTACAT	
qGlmY-f	GACTTATGTCAGCCCCTTCG	*glmY*
qGlmY-rv	CCCGGCTTTGTTATGGAATA	
qIlvL-f	ATGACAGCCCTTCTACGAG	*ilvL*
qIlvL-rv	AAGCCTTTCCTCGTCCAAGT	
qRraA-f	GGTTGCTGTACGATCTGCTC	*rraA*
qRraA-rv	TGCCGATATCCAACTCTTCC	

2.13 Protein purification of His_6-tagged RtcR and IHF

E. coli strain C41 carrying either pGH33H_6R, pGH33H_6RΔN, pGH25H_6RC or pGH33H_6IHF, which express N-terminally His_6-tagged RtcR, RtcRΔN, $RtcR_{C\text{-}term}$ and IHF respectively (see Sections 2.4.3 and 2.4.4), were diluted from ON cultures into 1 l fresh LB media supplemented with 100 µg/ml ampicillin or 50 µg/ml chloramphenicol as required and incubated at 37°C with aeration. At OD_{600} 0.3, protein expression was induced by addition of 1 mM IPTG or 0.2% arabinose as appropriate and cultures were incubated for a further 1-4 hours as indicated. Cells were harvested at 6000 rpm and 4°C for 20 minutes and resuspended in 40 ml ice-cold lysis buffer (50 mM Tris-HCl (pH 7.5), 300 mM NaCl, 10 mM imidazole, 5 mM β-mercaptoethanol, "complete" EDTA-free protease inhibitor (Roche)). Cells were lysed by sonication (Digital Sonifier, Branson) for 2 seconds followed by 2 seconds rest at 60 % amplification for 2.5 minutes. Sonication was repeated following 2.5

minutes on ice. Cell lysates were centrifuged at 16,000 rpm and 4°C for 20 minutes and the supernatants were incubated with washed 0.75 ml Ni-NTA agarose beads (Qiagen) for 1 hour at 4°C with gentle aeration. The mixture was then applied to a Poly-Pre Chromatography column (BioRad) which was subsequently washed several times with ice-cold wash buffer (50 mM Tris-HCl (pH 7.5), 300 mM NaCl, 35 mM imidazole, 1 mM β-mercaptoethanol). Protein was eluted from the Ni-NTA agarose by step-wise addition of 5 ml ice-cold elution buffer (50 mM Tris-HCl (pH 7.5), 300 mM NaCl, 1 mM β-mercaptoethanol) containing increasing concentrations of imidazole (62.5-250 mM). Small samples from each elution, along with samples from the culture, lysate and washing steps, were mixed with 4X NuPAGE LDS sample buffer (Thermo Fisher Scientific) according to the manufacturer's instructions and resolved on a NuPAGE Novex 4-12% Bis-Tris protein gel (Invitrogen) by SDS-PAGE, using Precision Plus Protein Dual-Xtra Prestained Protein Standards (BioRad) as a marker of protein size. Gels were stained using Instant Blue (Expedeon) stain for 1 hour and destained using dH_2O. Elution fractions containing protein of the correct size and purity were pooled and concentrated if necessary using Centrifugal Filter Units (Amicon) according to the manufacturer's recommendations. Proteins were then dialysed ON at 4°C in Slide-A-Lyzer Dialysis Cassettes (Thermo Fisher Scientific) with 5 l dialysis buffer 1 (50 mM Tris-HCl (pH 7.5), 300 mM NaCl, 2 mM β-mercaptoethanol). Finally, dialysis cassettes were transferred into 1 l dialysis buffer 2 (50 mM Tris-HCl (pH 7.5), 300 mM NaCl, 5 mM dithiothreitol (DTT), 50% glycerol) for 6 hours at 4°C and the dialysed proteins stored at -20°C. The concentration of dialysed proteins was assessed using a Bradford assay (Bradford, 1976) with Protein Assay Dye Reagent Concentrate (BioRad).

2.14 Electrophoretic mobility shift assay

DNA fragments were prepared by PCR using *E. coli* MG1655 template DNA with oligonucleotides RTCR9-f and RTCB3-rv for electrophoretic mobility shift assays (EMSAs) with H_6-RtcR (DNA_{RtcR}) and with oligonucleotides RTCB1-f and RTCB3-rv for EMSAs with H_6-IHF (DNA_{IHF}) . Mutations in the predicted IHF binding sites were introduced by PCR with primers RTCBmutIHF5-f and RTCBmutIHF5-rv ($DNA_{IHF\text{-}M5}$) and RTCBmutIHF8-f and RTCBmutIHF8-rv ($DNA_{IHF\text{-}M8}$). Control DNA ($DNA_{control}$) was amplified using primers hicA-f2 and hicA-rv (kindly provided by Kathryn Turnbull). 2-5 pmol DNA were then labelled with [γ-^{32}P]ATP using T4 polynucleotide kinase (PNK; Thermo Fisher Scientific) according to the manufacturer's protocol and purified on Illustra MicroSpin G-25 columns (GE Healthcare).

3 nM labelled DNA was mixed with purified protein or an equivalent volume of dialysis buffer 2 as a control and binding buffer (20 mM Tris-HCl (pH 8), 100 mM KCl, 2 mM MgCl, 1mM DTT, 50 µg/ml BSA, 100 µg/ml sheared salmon sperm DNA (Ambion)) to a total volume of 10 µl. The binding reaction was then incubated at 37°C for 20 minutes, mixed with 3X loading buffer (0.01% bromophenol blue, 0.01% xylene cyanol, 50% glycerol) and 5 µl loaded onto a running native acrylamide gel, which was electrophoresed at 200 V for 90 minutes. Gels consisted of either 6% or 8% acrylamide from a 40% Acrylamide/Bis-Acrylamide 19:1 stock (Sigma), 374 mM Tris-HCl (pH 8) 1X Tris-Glycine (Thermo Fisher Scientific), 0.08% ammonium persulphate and 0.08% tetramethylethylenediamine and the running buffer was 1X Tris-Glycine. Following electrophoresis, the gel was fixed (40% methanol, 20% acetic acid), dried and incubated at RT in a phosphor cassette for imaging.

2.15 CRAC analysis

Crosslinking and analysis of cDNA (CRAC) of RtcB was performed by comparison of four samples: *E. coli* MG1655Δ*rtcB* carrying pKW220rtcB; *E. coli* MG1655Δ*rtcB* carrying pKW220rtcB and treated with chloramphenicol; *E. coli* MG1655Δ*rtcB* carrying pKW220rtcB and pSC3326 (pBAD33::*mazF*) and *E. coli* MG1655Δ*rtcB* carrying empty pKW220 as a control. See Table 2 for details of the plasmids used. ON cultures from single bacterial colonies were diluted to OD_{600} 0.005 in 1 l of LB supplemented with 30 µ/ml ampicillin and, for the sample carrying pSC3326, 50 µg/ml chloramphenicol and cultured at 37°C with aeration. At OD_{600} 0.25, 100 µM IPTG was added to the cultures to induce expression of N-terminally FLAG-TEV-His_6 (FTH)-tagged RtcB from pKW220rtcB. After 10 minutes incubation at 37°C with aeration, the relevant cultures were treated with 50 µg/ml chloramphenicol or 0.2% arabinose for induction MazF expression from pSC3326. Following 30 minutes at 37°C with aeration, the cultures were irradiated with 1800 mJ of UV-C for 100 seconds (Van Remmen UV Techniek), harvested at 5000 rpm and 4°C for 15 minutes and washed in ice-cold phosphate-buffered saline (PBS). Pellets were snap frozen in liquid nitrogen and stored at -80°C.

The protocol used to purify FTH-RtcB, ligate linkers to the crosslinked RNA and generate libraries from cDNA is described in detail by Tree et al. (2014) in the supplementary information, including buffer recipes and timescales. Briefly, the cell pellets collected were lysed and FTH-RtcB was harvested from the lysates using M2 anti-FLAG resin (Sigma). The FLAG tag was removed by Halo-TEV protease cleavage (Promega) and RNA crosslinked to

H_6-RtcB was trimmed with RNaceIT Ribonuclease Cocktail (Agilent). H_6-RtcB-RNA complexes were immobilised on Ni-NTA resin (Qiagen) and alkaline phosphatase (Promega) treated to dephosphorylate the 5′ RNA ends prior to linker ligation. The 3′ miRCat-33 RNA linker (Integrated DNA Technologies) was ligated to 3′-ends of crosslinked RNA using T4 RNA Ligase I (New England Biolabs), then the 5′-ends were phosphorylated with [γ-^{32}P]ATP using T4 PNK (Thermo Fisher Scientific) and subsequently ligated to the 5′ L5 RNA linker. Four L5 linkers were used, each with a unique barcode at the 3′-end and assigned to one of the four samples tested (L5Aa, L5Ab, L5Ad and L5Bb). H_6-RtcB-RNA complexes were eluted from Ni-NTA agarose, resolved by SDS-PAGE (Section 2.13) and transferred to a Hybond C-Extra nitrocellulose membrane (Amersham) using a Trans-Blot Cell wet transfer system (BioRad). The membrane was imaged by autoradiography using a Carestream Biomax MS film (Sigma Aldrich.) and Glogos II Autorad Chemi-luminescent Markers (for alignment; Agilent). Smears of H_6-RtcB-RNA were cut out of the membrane from just above the molecular weight of RtcB alone, by aligning the membrane and autoradiograph with the Glogos markers. H_6-RtcB was digested with Proteinase K and RNA was extracted by phenol:chloroform:isoamylalcohol and chloroform and precipitated with ice-cold ethanol. Purified RNA was reverse transcribed (SuperScript III, Thermo Fisher Scientific) with primer miRCat RT and cDNA was PCR amplified using Takara LA Taq polymerase (Clontech) with primers P5 and PE_miRCat. PCR products were separated on a 2% MetaPhor agarose gel (Lonza) dissolved and ran in 1X TBE buffer (5X stock: 54 g Tris base, 27.5 g boric acid, 20 ml 0.5 M EDTA (pH 8) in 1 l) and stained with SYBR Safe (Thermo Fisher Scientific). The smear of amplicons between 150-300 bp, according to the GeneRuler 50 bp DNA ladder (Thermo Fisher Scientific) and above primer dimers also detectable in the control sample were extracted using the MinElute Gel Extraction kit (Qiagen), generating CRAC cDNA libraries. Primers used for 5′ and 3′ linkers, reverse transcription and PCR amplification are listed in Table 3.

The quality of the cDNA libraries were assessed by cloning with the TOPO TA Cloning kit (Invitrogen) and sequencing five clones per sample. The libraries for each sample were pooled and sent for high throughput sequencing on the illumina MiSeq platform at GATC Biotech. The sequencing output was analysed by Dr Kristoffer Winther (Gerdes Laboratory, University of Copenhagen) using the pyCRAC software (Webb et al., 2014).

2.16 Western blot analysis

A Western blot of the culture and lysate samples collected for CRAC analysis of RtcB (Section 2.15) was conducted to detect FTH-RtcB. SDS-PAGE of the samples was performed as described in Section 2.14 then proteins were transferred onto a PVDF membrane (GE Healthcare) using the semi-dry transfer method. The membrane was first soaked in 100% methanol then in semi-dry transfer buffer (25 mM Tris, 192 mM glycine, 0.1% SDS, 20% methanol) for 5 mins and the gel was equilibrated in transfer buffer for 5 minutes. The membrane and gel were then assembled on the TE70X semi-dry transfer unit (Hoefer) in the correct orientation, sandwiched between four sheets of transfer buffer-soaked blotting paper. Protein transfer was carried out for 1 hour at 115 mA. The membrane was blocked at RT for 1 hour in PBS-T (PBS with 0.1% Tween-20) containing 5% milk powder and subsequently incubated at RT for 1 hour with mouse Penta-His monoclonal antibodies (Qiagen) diluted 1:2000 in PBS-T containing 5% milk powder. Next, the membrane was incubated at RT for 1 hour with horseradish peroxidase-conjugated rabbit anti-mouse antibodies (Sigma Aldrich) diluted 1:20,000 in PBS-T containing 5% milk powder, then washed three times for 5 minutes with PBS-T. Finally, the blot was detected by assaying horseradish peroxidase activity using the Pierce ECL Western Blotting Substrate (Thermo Fisher Scientific; see manufacturer's instructions) followed by chemiluminescent imaging.

Chapter 3 Regulation of the *rtcBA* operon

3.1 Introduction

The *rtcBA* operon of *E. coli* was previously demonstrated to be regulated by σ^{54}-associated RNAP (Genschik et al., 1998). In the same publication, *rtcR*, the gene adjacent to the *rtcBA* operon, was found to encode a protein able to activate *rtcBA* transcription, consistent with the requirement of σ^{54}-regulated promoters for a bEBP. Although the general features of the *rtcBA* promoter are consistent with that of σ^{54}-dependent promoters, *rtcBA* regulation has not been characterised in detail.

bEBPs generally initiate σ^{54}-dependent transcription from a binding site 150 to 200 bp upstream of the transcription start site, known as a UAS. DNA binding is mediated by the C-terminal DNA binding domain of bEBPs, which binds as a dimer to an IR of variable length and spacing (Leonhartsberger et al., 2001; Porter et al., 1993; Tucker et al., 2005). The UAS of RtcR in the *rtcBA* promoter is presently unknown.

Furthermore, although sequences with high similarity to the consensus sequence for IHF binding were previously identified in the *rtcBA* promoter (Genschik et al., 1998), a role for IHF in *rtcBA* transcription has not been described. IHF is well known to mediate transcription from σ^{54}-dependent promoters by changing the topology of DNA between the UAS and the -12/-24 σ^{54} binding elements to favour a more stable interaction between bEBPs and σ^{54} (Huo et al., 2006).

Finally, there is no published information concerning the regulation of *rtcR* itself. The *rtcR* promoter is unmapped and the mode of transcription of *rtcR* is unknown. As a transcriptional activator of *rtcBA*, any regulatory mechanism affecting the levels of RtcR would also have consequences for *rtcBA* expression.

In this chapter, we address unanswered questions concerning the regulation of the *rtcBA* operon, including the location of the RtcR binding site in the *rtcBA* promoter, the role of IHF in transcription of *rtcBA* and the regulation of *rtcR* transcription.

3.2 Results

3.2.1 The *rtcBA* operon behaves as a typical σ^{54}-dependent promoter

3.2.1.1 Confirmation of RtcR-dependency of *rtcBA* transcription

To confirm transcriptional activation of the *rtcBA* operon by RtcR, a similar method to that used by Genschik et al. (1998) was adopted, in which transcription of *rtcBA* from the *E. coli* chromosome was compared during expression of wild-type RtcR and a truncated form of RtcR lacking the N-terminal regulatory domain (RtcRΔN). As described in Chapter 1, the ATPase activity of the central domain of bEBPs is generally negatively regulated by interaction with the regulatory domain in the absence of an activating signal (Fernández et al., 1995). Signal detection by the regulatory domain relieves repression of the central domain, which is then able to activate transcription. Therefore, RtcRΔN should be constitutively active due to the unrestricted activity of the central domain in the absence of the regulatory domain.

E. coli MG1655 carrying pGH322R (RtcR), pGH322RΔN (RtcRΔN) and pBAD322K (empty plasmid control) were cultured exponentially then overexpression of RtcR and RtcRΔN was induced for 30 minutes. Samples were collected and analysed by qPCR using *rtcB* and *rtcA*-specific primers to assess the relative levels of *rtcBA* mRNA in the strains. Figure 10A shows that levels of both *rtcB* and *rtcA* RNA are significantly increased in cells expressing RtcRΔN but not full-length RtcR, relative to the plasmid control, confirming the role of RtcR as a transcriptional activator of the *rtcBA* promoter.

3.2.1.2 Confirmation of σ^{54}-dependency of *rtcBA* transcription

Levels of *rtcBA* mRNA in a strain lacking σ^{54} were also analysed by qPCR, in order to confirm the σ^{54}-dependency of the *rtcBA* operon. For this purpose, *E. coli* MG1655 with a scarless deletion of the *rpoN* gene, which encodes σ^{54}, was constructed. During the recombination process, it was observed that MG1655Δ*rpoN* exhibited a slower rate of growth than wild-type MG1655. In order to ensure that this phenotype was due to deletion of *rpoN* and not a side-effect of recombination, σ^{54} was expressed from plasmid pGH25rpoN and the growth of the strains measured by optical density (Figure 11A). Complementation recovered the growth rate of MG1655Δ*rpoN* to that of wild-type MG1655, even without IPTG-induction of *rpoN* transcription from pGH25rpoN. Leaky transcription of *rtcBA* from plasmid

pGH25rtcBA was unable to complement the deletion of *rpoN*, indicating that the phenotype observed was unrelated to *rtcBA* transcription. This was also the case when *rtcBA* transcription was induced from pGH25rtcBA with 1 mM IPTG (data not shown). However, supplementation of LB with increasing concentrations of L-glutamine resulted in a concomitant increase in the growth rate of MG1655Δ*rpoN*, with a concentration of 20 µg/ml sufficient to fully recover wild-type growth (Figure 11B). This suggests that insufficient glutamine synthetase is produced in the absence of σ^{54} in the media used (Schaefer et al., 2015). Therefore, LB media was supplemented with L-glutamine in experiments involving MG1655Δ*rpoN* where relevant, to allow comparable growth with wild-type MG1655.

Using the MG1655Δ*rpoN*, the σ^{54}-dependency of *rtcBA* transcription was determined. As for wild-type MG1655, RtcR and RtcRΔN expression was induced from pGH322R and pGH322RΔN respectively. Whereas MG1655 showed a marked increase in the level of *rtcBA* transcription during RtcRΔN expression, qPCR analysis detected no change in *rtcBA* mRNA in any *rpoN* deletion strain (Figure 10B). This confirms that RtcR-dependent *rtcBA* transcription in *E. coli* occurs from a σ^{54}-regulated promoter as previously reported (Genschik et al., 1998). As expected, *rpoN* transcripts were only detectable in MG1655 and not in MG1655Δ*rpoN*, confirming the specific deletion of *rpoN*.

3.2.2 Identification of an RtcR binding region in the *rtcBA* promoter

3.2.2.1 Identification of minimal promoter requirements for *rtcBA* transcription

To further investigate *rtcBA* promoter activity, plasmid-based *rtcBA* transcriptional fusions with the *lacZ* reporter gene were constructed. The *lacZYA* transcriptional fusion plasmid pOU254 available in the laboratory collection was found to have high levels of β-galactosidase activity even without a promoter. Therefore, Rho-independent terminators of the ribosomal *rrnB* operon were inserted upstream of the *lacZ* gene in pOU254, generating pGH254 (Figure 12A). Comparison of the β-galactosidase activity of *E. coli* strain TB28 (MG1655Δ*lacIZYA*) carrying pOU254 and pGH254 revealed that the terminators were sufficient to reduce β-galactosidase activity from approximately 110 to 5 Miller units (Figure 12B). This suggests that the higher β-galactosidase activity exhibited by pOU254 is due to spurious transcription of *lacZ* from upstream sequences.

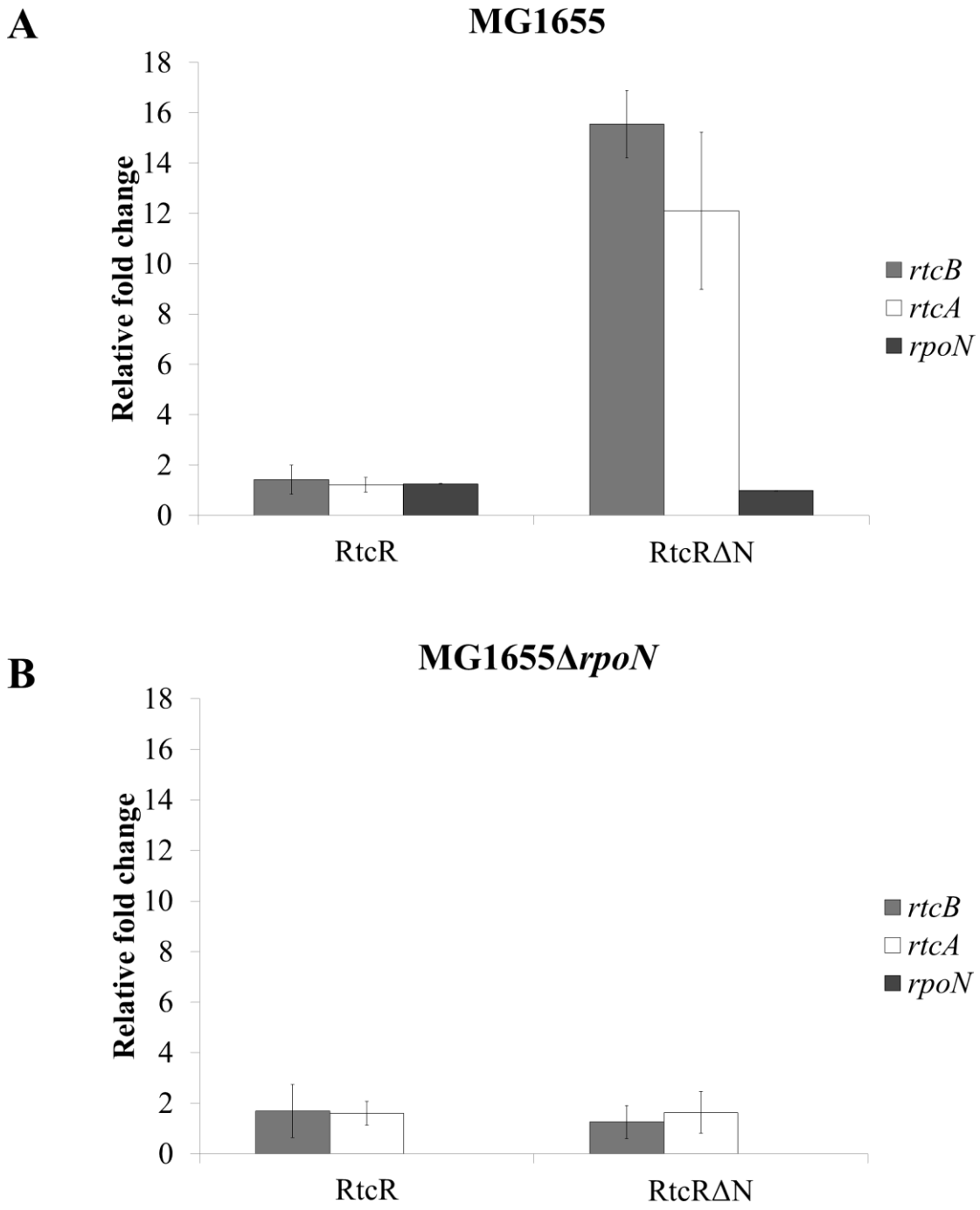


Figure 10 Expression of a mutant form of RtcR, RtcRΔN, activates *rtcBA* transcription.

E. coli MG1655 (A) and MG1655Δ*rpoN* (B) carrying pGH332R (RtcR), pGH322RΔN (RtcRΔN) or pBAD322K (corresponding empty plasmid) were grown exponentially in LB medium containing 25 µg/ml kanamycin and expression of RtcR and RtcRΔN was induced for 30 minutes with 0.2% arabinose. RNA was extracted from each sample, reverse transcribed and the fold change in the transcript levels of *rtcB, rtcA* and *rpoN* in strains expressing RtcR or RtcRΔN relative to strains carrying the empty plasmid were determined by qPCR. qPCR data were analysed using the Pfaffl method and normalised against the *tatA* gene. Data from three independent experiments are shown as the mean relative fold change ± standard deviation.

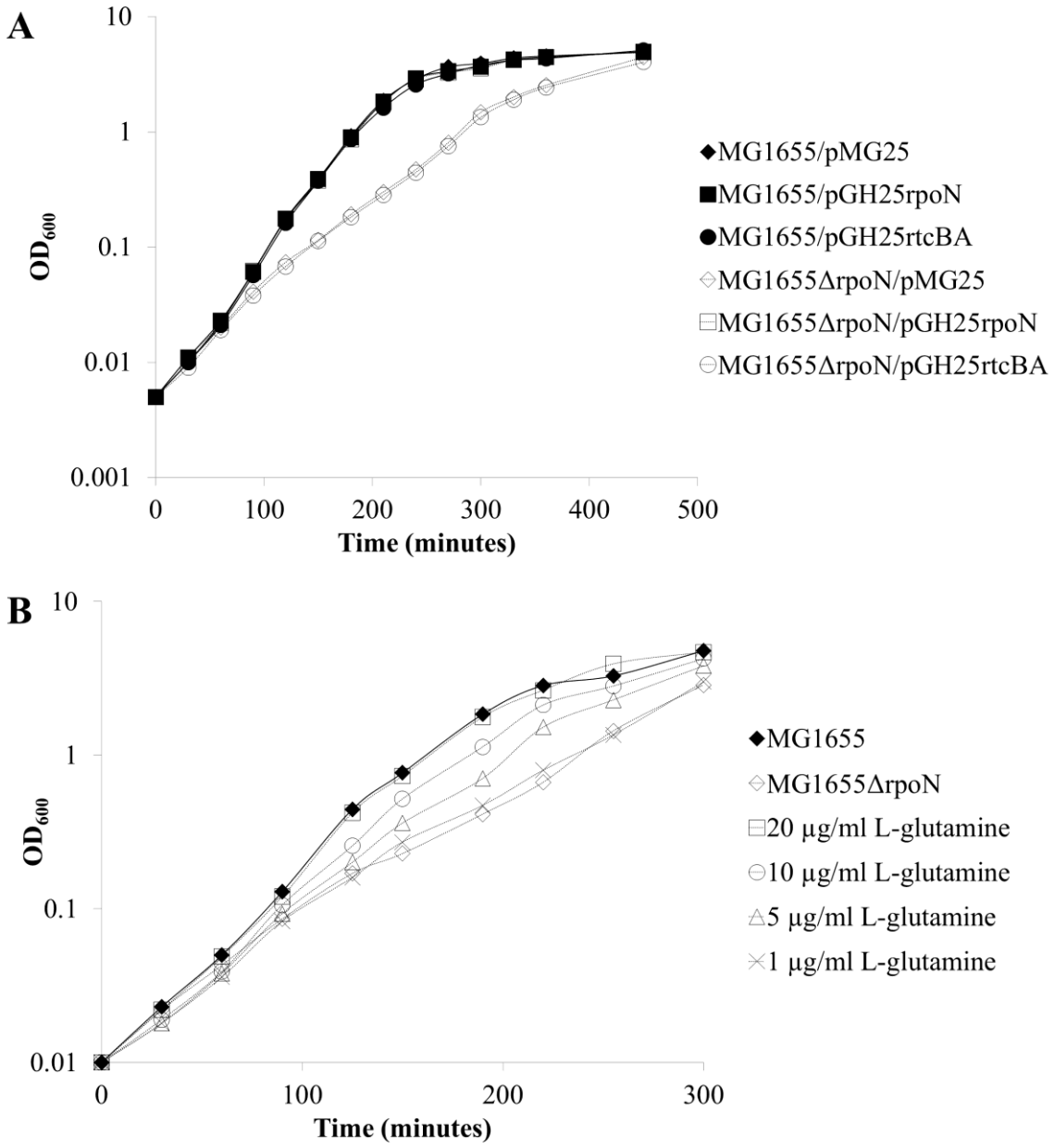


Figure 11 Deletion of *rpoN* reduces the growth rate of *E. coli* MG1655.

(A) *E. coli* MG1655 (black line) and MG1655Δ*rpoN* (dashed line) carrying pGH25rpoN (σ^{54} expression), pGH25rtcBA (RtcB and RtcA expression) or pMG25 (corresponding empty plasmid) were cultured in LB medium containing 100 µg/ml ampicillin and OD_{600} measured at regular intervals. (B) MG1655Δ*rpoN* was also cultured in LB medium supplemented with increasing concentrations of L-glutamine (1-20 µg/ml) and OD_{600} measured at regular intervals.

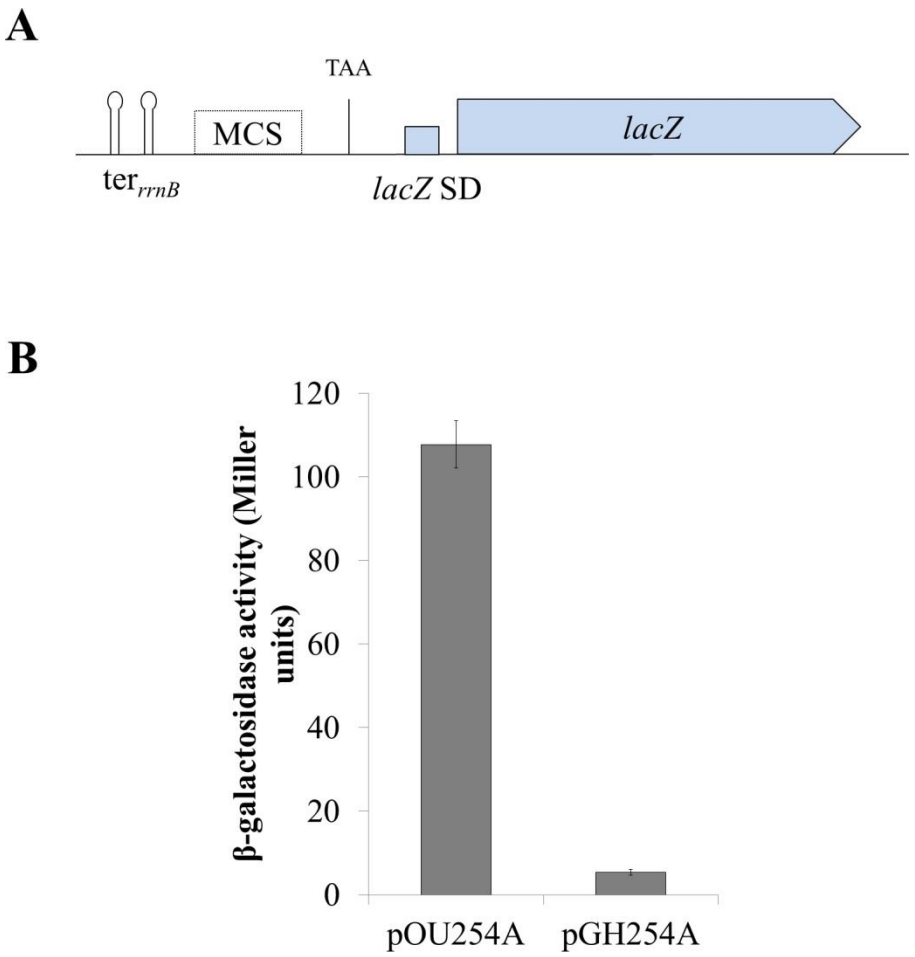


Figure 12 Construction of transcriptional *lacZ* fusions.

(A) Schematic organisation of pGH254 at the region used to generate transcriptional *lacZ* fusions (not to scale). A DNA fragment containing the promoter of interest is amplified and cloned into the multiple cloning site (MCS) of pGH254 between the *rrnB* terminators (ter$_{rrnB}$) and the *lacZ* Shine-Dalgarno (SD) and gene. The *lacZ* gene is thus transcribed under the control of the promoter of interest and translationally regulated by the *lacZ* SD. (B) *E. coli* TB28 (MG1655Δ*lacZYA*) carrying pGH254 or original plasmid pOU254 lacking *rrnB* terminators were cultured exponentially in LB medium supplemented with 30 μg/ml ampicillin. Samples were taken and β-galactosidase activity was measured. β-galactosidase activity is expressed in Miller units and the values are an average of three independent experiments ± standard deviation.

Having established the use of RtcR and RtcRΔN expression as a method for assaying *rtcBA* promoter activity (Section 3.2.1), a series of transcriptional *rtcBA-lacZ* fusions (BA0-3) were made using pGH254 with the aim of determining the minimal promoter requirement for activation of *rtcBA* transcription by RtcRΔN (Figure 13A). The levels of *lacZ* expression as a reporter for *rtcBA* transcriptional activation from each fusion was then monitored by β-galactosidase assay in *E. coli* TB28 carrying pGH322R (RtcR), pGH322RΔN (RtcRΔN) and empty pBAD322K. Following arabinose-induced expression from the pBAD322K variants, fusion BA3 (spanning from -262 bp relative to the *rtcBA* transcription start site (+1) to 30 bp into the *rtcB* open reading frame (ORF), with a transcriptional fusion to *lacZ*) was active in the presence of RtcRΔN but not RtcR (Figure 13B). Further *rtcBA-lacZ* fusions were constructed by successive deletion of 20-30 bp from the -262 position (including BA2 and BA0, data for the rest not shown). BA2, which starts at -164 bp relative to the transcription start site, was found to be the shortest fusion made that showed significant increase in β-galactosidase activity in the presence of RtcRΔN, whereas BA0, a 31 bp truncation of BA2, showed no increase in activity with RtcRΔN compared to RtcR and empty pBAD322K.

This indicated that a regulatory DNA element necessary for RtcR-dependent transcription of *rtcBA* was located between the start of BA2 and the start of BA0 (-164 and -133 relative to +1). As bEBPs usually bind around this region in σ^{54}-dependent promoters, the palindrome application on EMBOSS (Rice et al., 2000) was used to search for candidate IRs for RtcR binding. The longest perfect IRs identified were ATATC-N_{12}-GATAT positioned at -155 to -133 and CGATA-N_6-TATCG at -139 to -123. Thus, a further *rtcBA-lacZ* fusion starting at -142 was constructed (BA1) that truncates the upstream IR while the downstream remains intact. Similarly to BA0, RtcRΔN failed to activate *lacZ* transcription from BA1 (Figure 13B), which supports the upstream IR as the putative RtcR UAS, known as *rtcO*.

3.2.2.2 Mutagenesis reveals an inverted repeat important for RtcR-dependent *rtcBA* transcription

Following the identification of candidate UAS *rtcO* (ATATC-N_{12}-GATAT) in the *rtcBA* promoter, mutations in the transcriptional BA3 *rtcBA-lacZ* fusion were made to disrupt the the IR of *rtcO*. Figure 14A illustrates the organisation of the BA3 *rtcBA-lacZ* fusion (pGH254BA3) and the substitutions made to *rtcO* in ten mutated versions of the fusion, BA3-IR_{M1} to BA3-IR_{M10} (carried on plasmids pGH254BA3-IR_{M1-10}). The effect of the mutations on *rtcBA* transcription was assessed by β-galactosidase activity on solid NA supplemented with X-gal and arabinose in *E. coli* TB28 carrying pGH322R (RtcR), pGH322RΔN (RtcRΔN)

and empty pBAD322K (Figure 14B). As observed with fusions BA0 to BA3, fusions BA3-IR_{M1} to BA3-IR_{M10} demonstrated very low β-galactosidase activity in strains carrying pBAD322K and pGH332R, as indicated by the presence of white colonies only (Figure 14B; top and middle rows of each panel respectively). However, there were observable differences between the activity of the fusions during expression of RtcRΔN (Figure 14B; bottom row of each panel). Despite substitutions of two or more bases to disrupt *rtcO*, strains carrying fusions BA3-IR_{M1}, BA3-IR_{M2}, BA3-IR_{M3}, BA3-IR_{M8} and BA3-IR_{M10} produced blue colonies similar to those of BA3 when grown on X-gal-containing plates, indicative of transcriptional activation of the *rtcBA* promoter. In other cases, including BA3-IR_{M4}, BA3-IR_{M5}, BA3-IR_{M6} and BA3-IR_{M9}, expression of RtcRΔN produced pale blue colonies, indicative of reduced β-galactosidase activity from these strains, while BA3-IR_{M7} colonies appeared white, similar to BA0, in which *rtcO* is absent.

To gain further insight into the activity of each of the *rtcBA-lacZ* fusions carrying mutations in *rtcO*, β-galactosidase assays in liquid culture were performed in *E. coli* TB28 carrying pGH322RΔN, following 30 minutes of arabinose-induced RtcRΔN expression (Figure 14C). Consistent with the results observed on solid medium, fusion BA3-IR_{M7} showed the most significant reduction in β-galactosidase activity compared to BA3, from 168 to 10.6 Miller units. This fusion carried three substitutions in each half of the IR that changed the sequence but maintained the palindrome. Further, fusion BA3-IR_{M9}, with two substitutions in each half site, also demonstrated a significant reduction in β-galactosidase activity to 50.0 units. Fusions where mutations were made in the second half site of the IR only, exhibited reduced activity to varying extents compared to BA3. Mutation of GATAT to CAAAA in fusion BA3-IR_{M6} produced the greatest decrease to 68.4 units. Intriguingly, BA3-IR_{M4}, in which β-galactosidase activity was 120 units, had substitutions in all five bases and thus may have been predicted to be the most affected. However, in accordance with observations on solid media, all fusions with mutations in the first half of the IR (BA3-IR_{M1}, BA3-IR_{M2} and BA3-IR_{M3}) showed similar β-galactosidase activity to BA3 during RtcRΔN expression (Figure 14C). In addition, fusions BA3-IR_{M8} and BA3-IR_{M10}, which carried three and two substitutions in each half site of the IR respectively, both demonstrated β-galactosidase activity comparable to that of BA3.

3.2.2.3 Purification of RtcR and *in vitro* analysis of *rtcBA* promoter binding

In order to further investigate the location and sequence of the RtcR binding site and particularly the role of *rtcO*, purification of RtcR was performed to allow *in vitro* analyses.

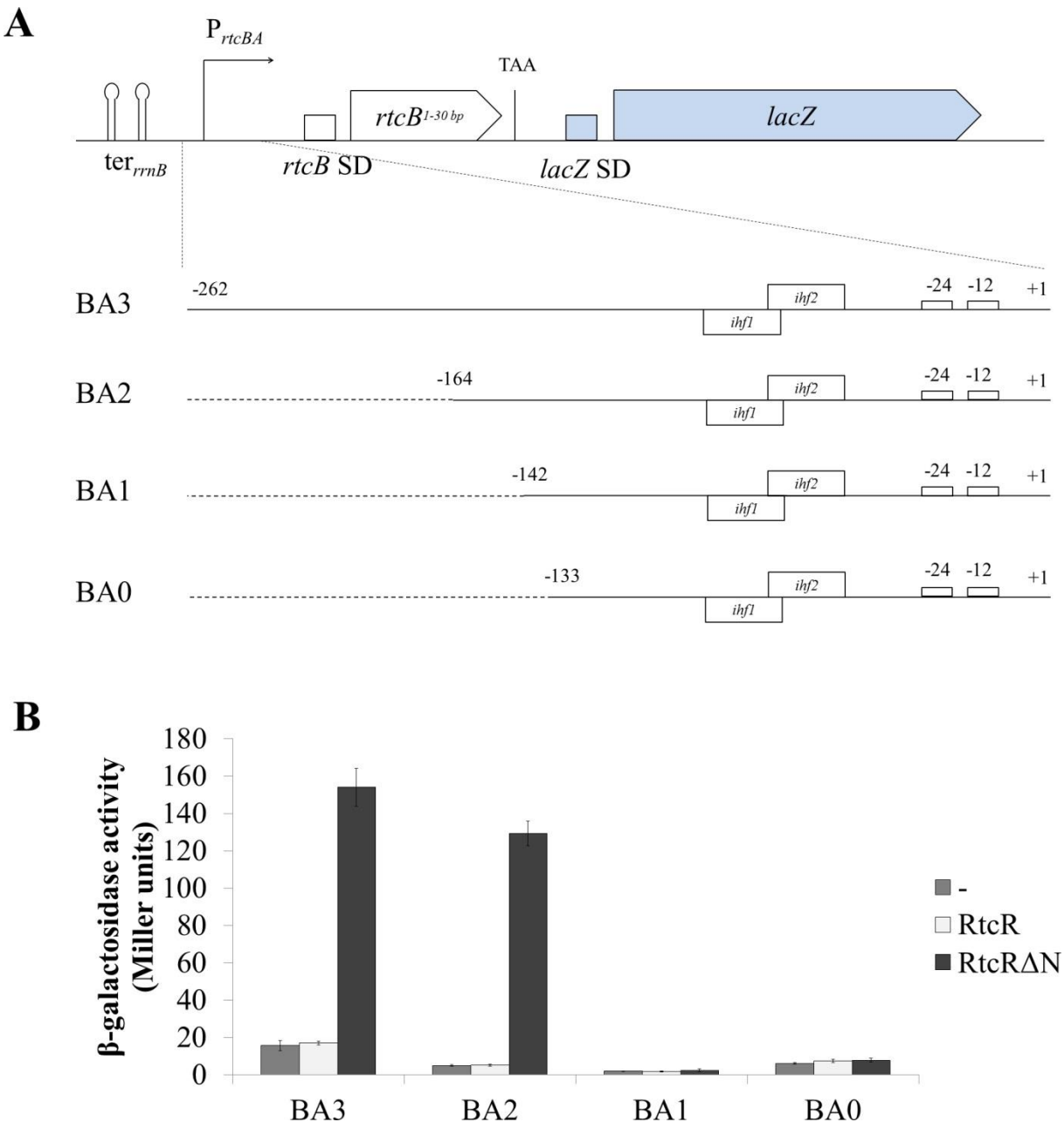


Figure 13 Promoter requirements for *rtcBA* transcription.

(A) Schematic organisation of transcriptional *rtcBA-lacZ* fusions carried on pGH254BA3 (BA3), pGH254BA2 (BA2), pGH254BA1 (BA1) and pGH254BA0 (BA0) (not to scale). Fragments consisting of the upstream sequences of *rtcB* were amplified and cloned into pGH254 to generate transcriptional fusions with the *lacZ* reporter gene. The starting point for the fragments are as follows: the BA3 fusion begins at -262 bp relative to the +1 transcription start site of *rtcB*, BA2 begins at -164, BA1 at -142 and BA0 at -133. All fragments end 30 bp (10 amino acids) into the *rtcB* open reading frame (ORF) and include the -12/-24 elements for σ^{54} recognition and the predicted IHF binding sites (*ihf1* and *ihf2*). (B) *E. coli* TB28 carrying pGH332R (RtcR), pGH322RΔN (RtcRΔN) or pBAD322K (corresponding empty plasmid) with pGH254BA3, pGH254BA2, pGH254BA1, pGH254BA0 or empty pGH254, in all combinations, were cultured exponentially in LB medium supplemented with 30 µg/ml ampicillin and 25 µg/ml kanamycin. Expression of RtcR and RtcRΔN was induced with 0.2% arabinose and samples taken 30 minutes later. β-galactosidase activity of each sample was measured and the values for empty pGH254 with the pBAD322K variants were deducted. β-galactosidase activity is expressed in Miller units and the values are an average of three independent experiments ± standard deviation.


A
-262
P rtcBA
-160 -129
ihf2
ihf1
-24 -12
+1
rtcB 1-30 bp
TAA
lacZ
ter rrnB
rtcB SD
lacZ SD
B
C
β-galactosidase assay
BA3 (WT) CTTAGATATCCTTATAAAAGACGATATTTCAT
168 ± 20.8
BA3-IR M1 -GCGCT----------GATAT-
174 ± 26.7
BA3-IR M2 -GTGTT----------GATAT-
160 ± 20.7
BA3-IR M3 -TTTTG----------GATAT-
194 ± 23.0
BA3-IR M4 -ATATC----------AGCGC-
120 ± 31.0
BA3-IR M5 -ATATC----------AACAC-
100 ± 29.3
BA3-IR M6 -ATATC----------CAAAA-
68.4 ± 10.7
BA3-IR M7 -GTGTT----------AACAC-
10.6 ± 3.85
BA3-IR M8 -GTGTT----------TAGAG-
157 ± 26.1
BA3-IR M9 -GTAGC----------GGTAG-
50.0 ± 8.45
BA3-IR M10 -ATGGC----------GGGAT-
154 ± 17.9
BA0 TTCAT
4.35 ± 1.91
-
-
-
RtcR
RtcRΔN

Figure 14 Mutations in *rtcO* of the *rtcBA* promoter reduce transcription.

(A) Schematic organisation of the transcriptional *rtcBA-lacZ* fusion carried on pGH254BA3 (BA3), as described in Figure 13, including the sequence around the inverted repeat (IR) *rtcO* (marked with arrows) predicted to be involved with RtcR binding. A series of pGH254BA3 variants (pGH254BA3-IR_{M1-10}) with mutations in the inverted repeat (red), generating *rtcBA-lacZ* transcriptional fusions BA3-IR_{M1-10} respectively, are shown. Plasmids pGH254BA0 (BA0, described in Figure 13) and pGH254 (corresponding empty plasmid) are also shown. (B) *E. coli* TB28 carrying each fusion along with pBAD322K (top row of each panel), pGH332R (RtcR; middle row of each panel) or pGH322RΔN (RtcRΔN; bottom row of each panel) were plated in 100-fold serial dilutions on NA plates containing 30 µg/ml ampicillin, 25 µg/ml kanamycin, 40 µg/ml X-gal and 0.2% arabinose and incubated 12 hours at 37°C. (C) *E. coli* TB28 carrying each fusion along with pGH322RΔN were cultured exponentially in LB medium supplemented with 30 µg/ml ampicillin and 25 µg/ml kanamycin and 0.2% arabinose was added to induce expression of RtcRΔN for 30 minutes. Samples were taken and β-galactosidase activity from the transcriptional *rtcBA-lacZ* fusions was measured. β-galactosidase activity is expressed in Miller units and the values are an average of three independent experiments ± standard deviation, with the activity of pGH254 with pGH322RΔN deducted.

RtcR was initially cloned onto high-copy-number plasmid pMG25 with an N-terminal His_6-tag, to enable affinity purification on nickel. However, expression of H_6-RtcR from pMG25 in the *E. coli* strain C41 revealed that upon lysis the protein was found in the insoluble fraction (data not shown). A range of culturing conditions were tested, to attempt to increase the amount of soluble protein including variation of temperature (37°C, 30°C and 16°C), variation of induction time (one hour to ON) and variation of amount of inductive agent (100 µM to 1 mM IPTG). These alterations produced no detectable increase in the amount of soluble H_6-RtcR.

Next, H_6-RtcR was cloned in a lower-copy-number plasmid, pBAD33, producing pGH33H_6R. A second plasmid, pGH33H_6RΔN, expressing H_6-RtcRΔN was also generated, to test whether removal of the regulatory domain of RtcR increased solubility. *E. coli* C41 carrying pGH33H_6R or pGH33H_6RΔN was cultured exponentially at 37°C and protein expression induced with 0.2% arabinose for one hour. H_6-RtcR and H_6-RtcRΔN were purified from the cell pellet as described in Chapter 2. Cellular proteins detectable in the cultures, lysates and elutions during the purification process were resolved by SDS-PAGE and the stained gels are shown in Figure 15A (H_6-RtcR) and Figure 15B (H_6-RtcRΔN). In both cases, although high levels of protein of the expected size, 60.3 kDa for RtcR and 40.7 kDa for RtcRΔN, were expressed (compare lanes 1 and 2 for each), the majority of overexpressed protein was insoluble (compare lanes 3 and 4). However, weak bands corresponding to the approximate size of H_6-RtcR and H_6-RtcRΔN were detectable in the elution fractions (upper bands, lanes 5-9), along with some unknown contaminants or degradation products (lower bands).

Therefore, elution fractions containing the most and purest H_6-RtcR and H_6-RtcRΔN were pooled for use in *in vitro* experiments.

To attempt to increase the amount of soluble protein, the C-terminal DNA binding domain of RtcR was cloned onto pMG25 with an N-terminal His_6-tag. *E. coli* C41 carrying the resulting plasmid, pGH25H_6RC, was cultured exponentially at 37°C and protein expression induced with 1 mM IPTG for four hours. H_6-$RtcR_{C\text{-}term}$ was purified and samples from cultures, lysates and elutions were resolved by SDS-PAGE (Figure 15C). Clear bands corresponding to the size of H_6-$RtcR_{C\text{-}term}$ at 13.2 kDa were visible after IPTG induction (lane 2 onwards) and indicate that, unlike H_6-RtcR and H_6-RtcRΔN, the DNA binding domain of RtcR is soluble in the absence of the other domains. The elution fractions were pooled and used along with H_6-RtcR and H_6-RtcRΔN.

Purified proteins were incubated with a DNA fragment containing the *rtcBA* promoter from -203 bp to +8 relative to the *rtcBA* transcription start site +1 (DNA_{RtcR}) and analysed by EMSA (Figure 16). Two-fold serial dilutions of H_6-RtcR (lanes 1 to 3) and H_6-RtcRΔN (lanes 4 to 6) and ten-fold serial dilutions of H_6-$RtcR_{C\text{-}term}$ (lanes 7 to 9) were used in the binding reactions, however no DNA shift indicative of the formation of an RtcR-DNA_{RtcR} complex was detectable compared to the DNA_{RtcR} fragment alone (lane 10). The range of protein concentrations used was increased and other parameters were changed, including the ratio of acrylamide used in the gels, the gel running time and the use of an alternative binding buffer (Pedersen and Valentin-Hansen, 1997). However, none of the changes elicited a shift in the DNA fragment (data not shown).

As a result of the inability to detect interaction between *rtcBA* promoter DNA and purified RtcR proteins, further EMSAs to examine the ability of RtcR to bind and shift DNA fragments with mutations in *rtcO* and DNase I footprinting analysis to determine which DNA sequences are protected by RtcR *in vitro* were not conducted.

3.2.3 IHF-dependency of *rtcBA* transcription

3.2.3.1 IHF is required for RtcR-dependent *rtcBA* transcription

In order to ascertain the role of IHF in the transcription of *rtcBA*, strains carrying *ihf* deletions were constructed using the Keio collection of mutants, in which single genes are replaced by

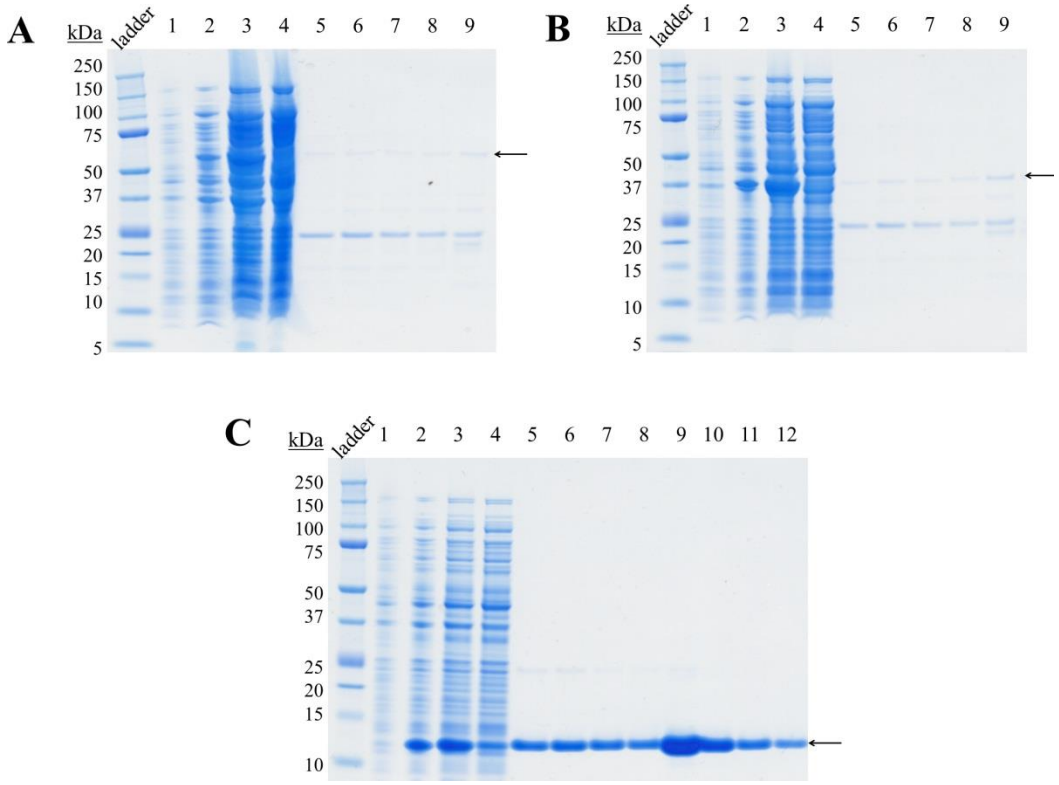


Figure 15 Protein purification of RtcR.

E. coli C41 carrying pGH33H_6R, pGH33H_6RΔN or pGH25H_6RC were cultured exponentially and induced for 1 hour with 0.2% arabinose or 4 hours with 1 mM IPTG as indicated in the text and H_6-RtcR (A), H_6-RtcRΔN (B) and H_6-$RtcR_{C\text{-}term}$ (C) respectively were purified as described in Chapter 2. Samples from cultures before (lane 1) and after induction (lane 2), from lysates (lane 3) and lysate supernatants (lane 4) and from elution fractions (lanes 5 to 12) were resolved by SDS-PAGE and stained for analysis. A protein ladder was included as a marker of protein size (indicated in kDa). Arrows (←) indicate the expected size of the protein purified.

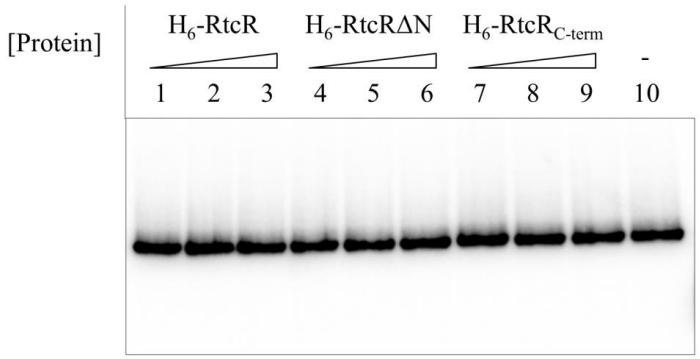


Figure 16 Purified RtcR is unable to bind promoter *rtcBA* DNA *in vitro*.

Purified H_6-RtcR, H_6-RtcRΔN and H_6-$RtcR_{C\text{-}term}$ were added to [γ-^{32}P]ATP-labelled DNA_{RtcR} probe containing the *rtcBA* promoter (-203 to +8 relative to the *rtcBA* transcription start site +1) in the following concentrations: 150 nM, 75 nM and 37.5 nM H_6-RtcR (lanes 1-3), 4 nM, 2 nM and 1 nM H_6-RtcRΔN (lanes 4-6) and 400 nM, 40 nM and 4 nM H_6-$RtcR_{C\text{-}term}$ (lanes 7-9). A no protein control was included (lane 10). An EMSA was performed on a 6% native polyacrylamide gel.

the kanamycin resistance gene (Baba et al., 2006). IHF is a heterodimer of IHFα and IHFβ, encoded by *ihfA* and *ihfB* respectively at different loci of the *E. coli* genome. Initially, the kanamycin resistance gene replacing *ihfA* was transduced from the relevant Keio strain into *E. coli* TB28. In order to combine the Keio *ihfB* deletion into TB28Δ*ihfA::kan*, creating a complete *ihf* deletion, the kanamycin resistance cassette needed to be removed from the *ihfA* locus. However, TB28Δ*ihfA::kan* was unable to tolerate transformation with pCP20, which encodes the *flp* gene required to remove the kanamycin gene. Transformation with pGH254 for the measurement of transcriptional *rtcBA-lacZ* activity was also not possible (data not shown). This suggests that functional IHF plays a role in the maintenance or replication of some plasmids in *E. coli.*

As a result, a different approach was used to investigate the involvement of IHF in *rtcBA* transcription. *ihfA* and *ihfB* deletions from the KEIO collection were transduced separately into *E. coli* TB28 with a transcriptional *rtcBA-lacZ* fusion on the chromosome (TB28*rtcBA-lacZ*; Figure 17A), creating TB28*rtcBA-lacZ*Δ*ihfA::kan* and TB28*rtcBA-lacZ*Δ*ihfB::kan*. Chloramphenicol resistant versions of the plasmids used in Sections 3.2.1 and 3.2.2 for RtcR and RtcRΔN expression were transformed successfully into TB28*rtcBA-lacZ* and the *ihfA* and *ihfB* deletion strains and each strain was plated onto NA containing X-gal and arabinose. As illustrated in Figure 17B, the only strain with blue colonies, indicative of transcriptional activation of the *rtcBA* promoter, was in the wild-type strain expressing RtcRΔN. Both *ihf* deletion strains expressing RtcRΔN had white colonies, as did all strains expressing RtcR and those carrying the empty plasmid. This result was confirmed in liquid culture using a β-galactosidase assay to assess the activity of the transcriptional *rtcBA-lacZ* fusion (Figure 17C). The lack of β-galactosidase activity in the *ihfA* and *ihfB* deletion strains demonstrates a requirement for functional IHF in the activation of RtcR-dependent *rtcBA* transcription.

3.2.3.2 IHF binding sites are functional in mediating *rtcBA* transcription

Global regulator IHF modifies DNA architecture by binding to specific DNA sequences and inducing bending (Xu and Hoover, 2001). In the case of σ^{54}-dependent transcription, this mediates interaction between σ^{54} and the bEBP bound upstream. Two sites with high similarity to the asymmetric IHF consensus sequence were previously identified in the *rtcBA* promoter (Genschik et al., 1998), *ihf1* (on the opposite strand to the *rtcB* coding sequence and orientated away from *rtcB*) and *ihf2* (on the same strand and orientated towards *rtcB*) as shown in Figure 5.

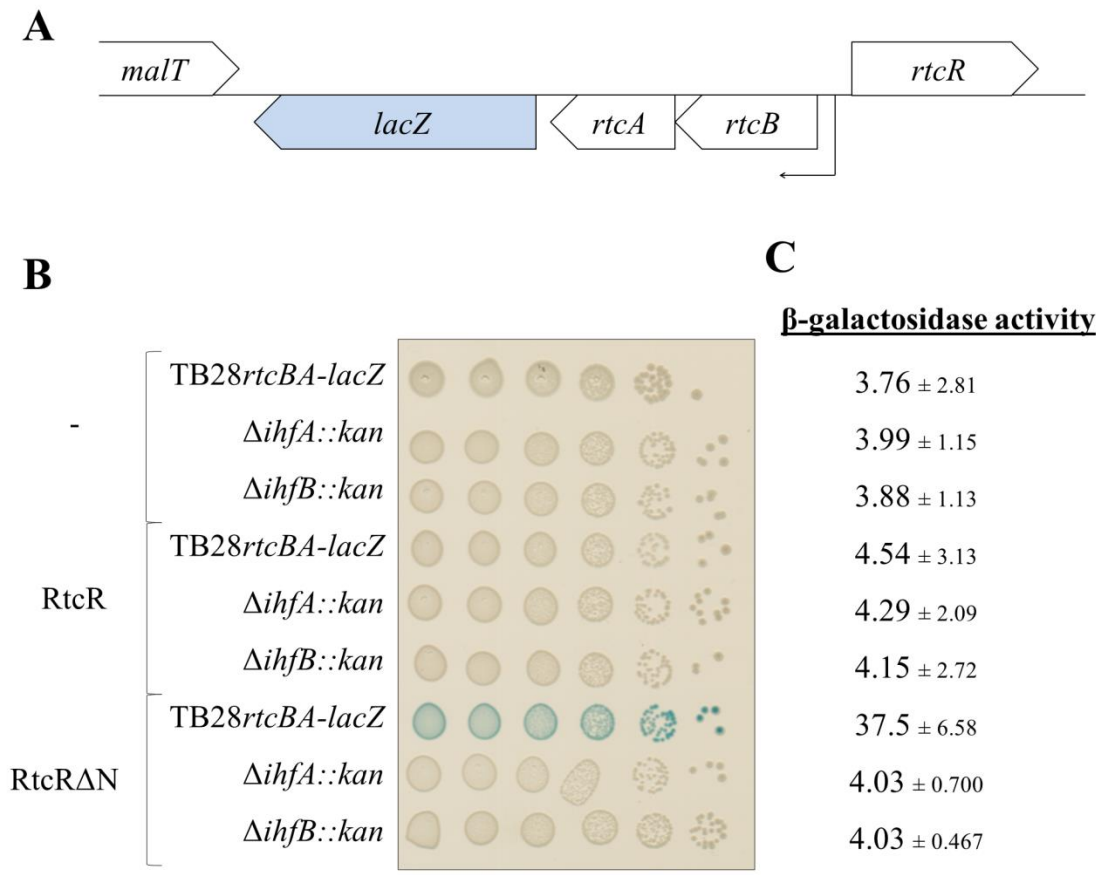


Figure 17 Transcription of *rtcBA* is dependent on the presence of functional IHF.

(A) Genetic organisation at the *rtcBA* locus in the *E. coli* strain TB28*rtcBA-lacZ*, which carries a transcriptional fusion of *lacZ* as part of the *rtcBA* operon. (B) TB28*rtcBA-lacZ*, TB28*rtcBA-lacZ*Δ*ihfA::kan* and TB28*rtcBA-lacZ*Δ*ihfB::kan* carrying pGH322CR (RtcR), pGH322CRΔN (RtcRΔN) or pBAD322C (the corresponding empty plasmid, -) were plated in 10-fold serial dilutions on NA plates containing 50 µg/ml chloramphenicol, 40 µg/ml X-gal and 0.2% arabinose. (C) The same strains were cultured exponentially in LB medium supplemented with 50 µg/ml chloramphenicol and 0.2% arabinose was added to induce expression of RtcR and RtcRΔN for 30 minutes. Samples were taken and β-galactosidase activity from the chromosomal *rtcBA-lacZ* transcriptional fusion was measured. β-galactosidase activity is expressed in Miller units and the values are an average of two independent experiments ± standard deviation, with the activity of TB28 with the pBAD322K variants deducted.

Having established IHF-dependency of *rtcBA* transcription, the two putative IHF sites were mutated in the transcriptional BA3 *rtcBA-lacZ* fusion (pGH254BA3) to determine whether these DNA elements are required for *rtcBA* transcription (Figure 18A). Mutations at highly conserved positions of the IHF consensus sequence were made in *ihf1*, *ihf2* or both according to Hales et al. (1994), where the severity of different substitutions in the IHF binding sites of bacteriophage λ were assessed. The resulting fusions, BA3-IHF$_{M1}$ to BA3-IHF$_{M7}$ (carried on plasmids pGH254BA3-IHF$_{M1}$ to pGH254BA3-IHF$_{M7}$), were evaluated for the effect on *rtcBA* transcription by β-galactosidase activity on NA supplemented with X-gal and arabinose in *E. coli* TB28 carrying pGH322R (RtcR), pGH322RΔN (RtcRΔN) and empty pBAD322K (Figure 18B). Mutations in *ihf1* (BA3-IHF$_{M1}$) or *ihf2* (BA3-IHF$_{M2}$ and BA3-IHF$_{M3}$) alone were insufficient to disrupt *rtcBA* transcription, as indicated by the presence of blue colonies on X-gal-supplemented plates during induction of RtcRΔN expression (bottom row of each panel). Further, fusions with substitutions in one or two positions to mutate both IHF binding sites simultaneously (BA3-IHF$_{M4}$ and BA3-IHF$_{M5}$ respectively) also generated blue colonies. Fusion BA3-IHF$_{M6}$, which combines mutations of BA3-IHF$_{M4}$ and BA3-IHF$_{M5}$, showed an intermediate phenotype when combined with RtcRΔN expression, while BA3-IHF$_{M7}$, in which 12 conserved bases of *ihf1* and *ihf2* were substituted for alternative bases, produced white colonies, indicating that β-galactosidase activity was significantly reduced. As observed for all other transcriptional *rtcBA-lacZ* fusions, BA3-IHF$_{M1}$ to BA3-IHF$_{M7}$ carrying pBAD322K and pGH322R produced white colonies (Figure 18B; top and middle rows of each panel respectively).

The β-galactosidase activity of the transcriptional BA3 *rtcBA-lacZ* fusion was also compared to those with IHF binding site mutations in liquid culture following 30 minutes of RtcRΔN expression from pGH322RΔN by β-galactosidase assay (Figure 18C). Consistent with observations on solid media, fusions BA3-IHF$_{M1}$, BA3-IHF$_{M2}$ and BA3-IHF$_{M3}$, with mutations in *ihf1* or *ihf2* alone, retained high β-galactosidase activity. Fusions BA3-IHF$_{M4}$ to BA3-IHF$_{M7}$ carrying mutations in both *ihf1* and *ihf2* showed a progressive decrease in β-galactosidase activity compared to wild-type BA3 concomitant with increasing number of substitutions. For example, BA3-IHF$_{M4}$, which has a single substitution of T to C in the overlap between *ihf1* and *ihf2*, had a moderate reduction in activity from 168 to 145 Miller units, whereas BA3-IHF$_{M7}$, in which 12 conserved bases of *ihf1* and *ihf2* were substituted, displayed the greatest reduction to 18.4 Miller units, as indicated by the results on X-gal plates. Therefore, the data suggests that the putative IHF binding sites are functional and

necessary regulatory elements of the *rtcBA* promoter that support RtcR-dependent transcription of *rtcBA*, likely as a result of IHF binding.

3.2.3.3 IHF binds to *rtcBA* promoter DNA *in vitro*

The findings of Sections 3.2.3.1 and 3.2.3.2 strongly imply that IHF mediates its effect on *rtcBA* transcription from the two putative IHF binding sites in the *rtcBA* promoter. Therefore, to establish whether IHF can bind to *ihf1* and *ihf2 in vitro*, purification of IHF was conducted.

Nash et al. (1987) determined that IHFα was unstable and IHFβ was insoluble when overexpressed separately, but that co-expression of both subunits produced stable and soluble IHF. Therefore, the *ihfA* and *ihfB* genes were sequentially cloned into pBAD33 to allow stoichiometric expression of the individual subunits. *E. coli* C41 carrying the resulting plasmid, pGH33H_6IHF, was cultured exponentially at 37°C and protein expression induced with 0.2% arabinose for four hours. IHFα was His_6-tagged at the N-terminal to enable co-purification with untagged IHFβ by affinity chromatography, as described in Chapter 2. Samples of cellular proteins from the culture, lysate and elutions during purification were resolved by SDS-PAGE and the stained gel is shown in Figure 19. Clear bands corresponding to the size of H_6-IHFα and IHFβ, at 12.2 kDa and 10.7 kDa respectively, were visible following induction of protein expression (compare lane 1 with lane 2 onwards). The presence of IHFβ in the elution fractions demonstrated that the subunits were purified in complex as IHF (lane 5 onwards). The presence of contaminating protein in the elution fractions of the same size as observed during RtcR purifications (Figure 16), indicates that these bands are a general contaminant. Elution fractions containing high amounts of IHF subunits were pooled and used for *in vitro* DNA binding experiments.

Initially, ten-fold serial dilutions of purified H_6-IHF were incubated with a DNA fragment containing the *rtcBA* promoter from -142 bp to +8 relative to the *rtcBA* transcription start site +1 (DNA_{IHF}; data not shown). Having established that H_6-IHF could successfully form a complex with DNA_{IHF}, further DNA_{IHF} fragments were generated to ascertain whether binding was mediated via the putative IHF binding sites: $DNA_{IHF\text{-}M7}$ had mutations in 12 conserved sites of *ihf1* and *ihf2*, as were made in the transcriptional BA3-IHF_{M7} *rtcBA-lacZ* fusion, and $DNA_{IHF\text{-}M6}$ had 3 mutations spanning *ihf1* and *ihf2*, as for fusion BA3-IHF_{M6} (Figure 18A). A DNA fragment of the same size from an unrelated part of the *E. coli* genome was also included as a control to assay for any non-specific DNA binding by H_6-IHF

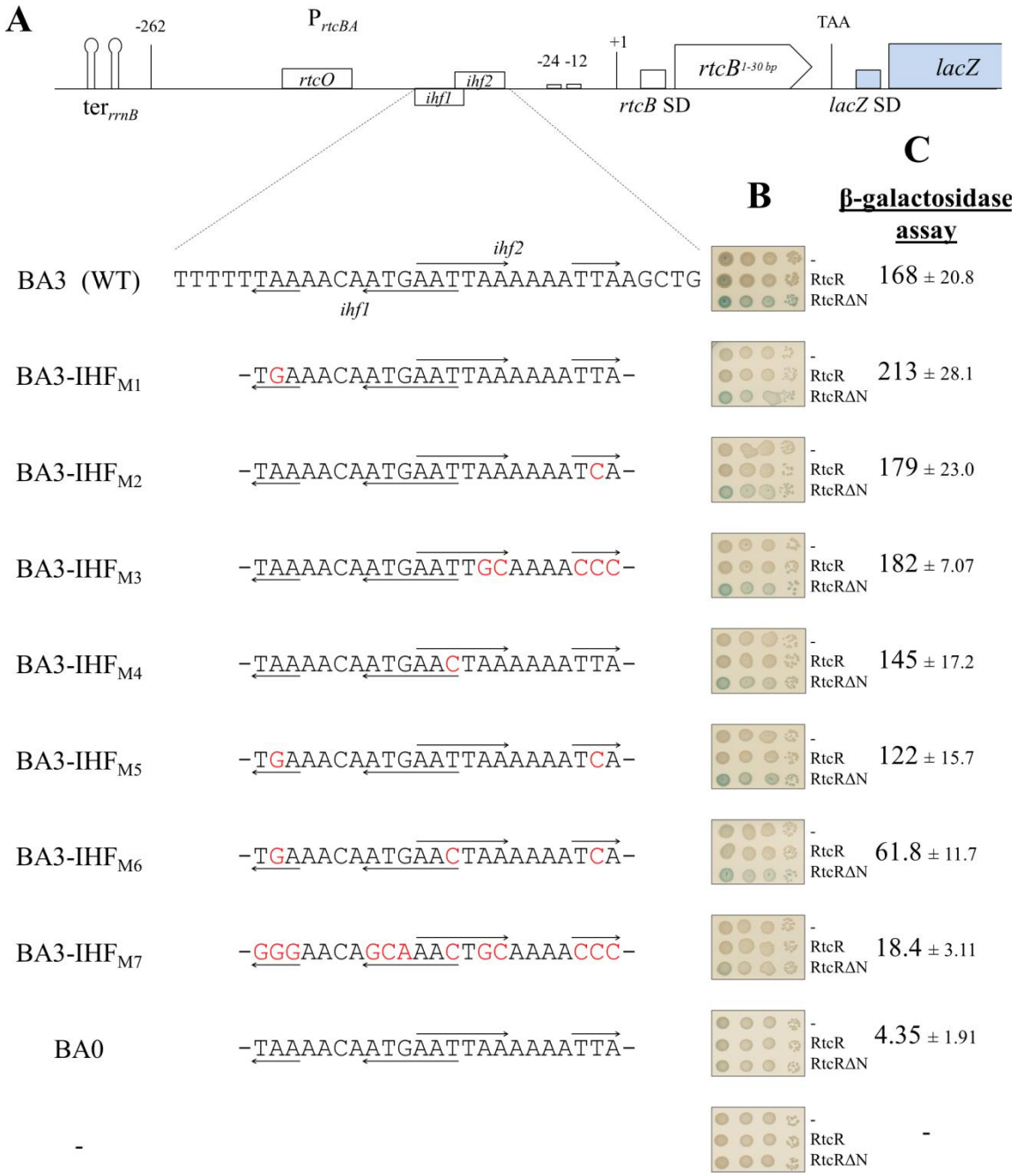


Figure 18 Mutations in IHF binding sites reduce *rtcBA* transcription.

(A) Schematic organisation of the transcriptional *rtcBA-lacZ* fusion carried on pGH254BA3 (BA3), including the sequence around two putative IHF binding sites (*ihf1* and *ihf2*; underlined). Variants of pGH254BA3 (pGH254BA3-IHF$_{M1-7}$) with mutations in *ihf1* and/or *ihf2* (red), generating *rtcBA-lacZ* transcriptional fusions BA3-IHF$_{M1-7}$, are shown. Plasmids pGH254BA0 (BA0) and pGH254 (corresponding empty plasmid) are included as negative controls. (B) *E. coli* TB28 carrying each fusion along with pBAD322K (top row of each panel), pGH332R (RtcR; middle row of each panel) or pGH322RΔN (RtcRΔN; bottom row of each panel) were plated in 100-fold serial dilutions on NA plates containing 30 µg/ml ampicillin, 25 µg/ml kanamycin, 40 µg/ml X-gal and 0.2% arabinose and incubated 12 hours at 37°C. (C) *E. coli* TB28 carrying each fusion along with pGH322RΔN were cultured exponentially in LB medium supplemented with 30 µg/ml ampicillin and 25 µg/ml kanamycin and 0.2% arabinose was added to induce expression of RtcRΔN for 30 minutes. Samples were taken and β-galactosidase activity from the transcriptional *rtcBA-lacZ* fusions was measured. β-galactosidase activity is expressed in Miller units as an average of three independent experiments ± standard deviation, with the activity of TB28 carrying pGH254 and pGH322RΔN deducted.

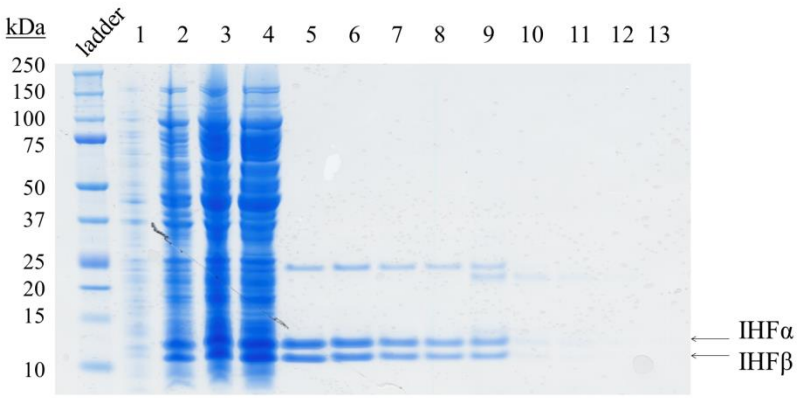


Figure 19 Protein purification of IHF.

E. coli C41 carrying pGH33H_6IHF was cultured exponentially and expression of H_6-IHF induced for 4 hours with 0.2% arabinose then purified as described in Chapter 2. Samples from cultures before (lane 1) and after induction (lane 2), from cell lysate (lane 3) and lysate supernatant (lane 4) and from elution fractions (lanes 5 to 13) were resolved by SDS-PAGE and stained for analysis. A protein ladder was included as a marker of protein size (indicated in kDa). Arrows (←) indicate the expected size of the IHFα and IHFβ subunits purified.

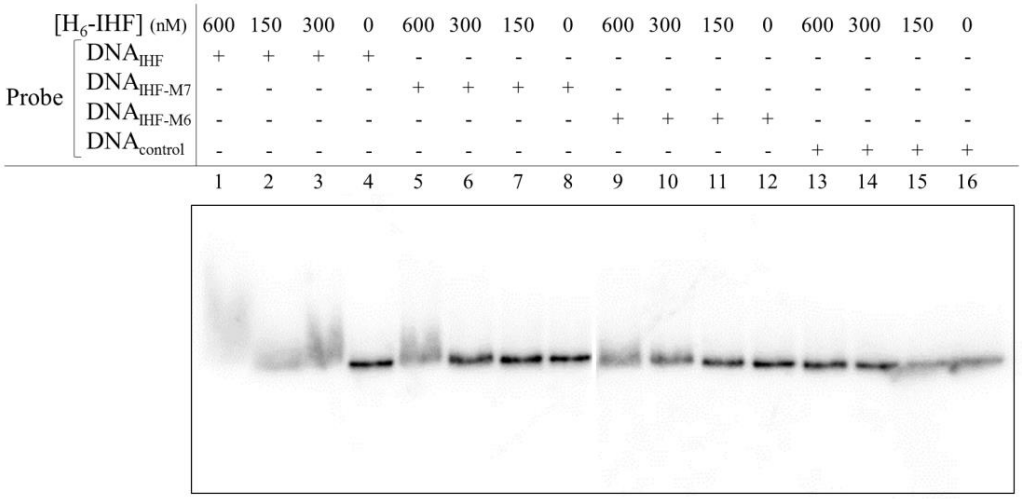


Figure 20 Purified IHF binds to *rtcBA* promoter DNA *in vitro*.

Purified H_6-IHF was added to [γ-^{32}P]ATP-labelled DNA probes in the following concentrations: 600 mM (lanes 1, 5, 9 and 13), 300 nM (lanes 3, 6, 10 and 14) and 150 nM (lanes 2, 7, 11 and 15). No protein controls were included (lanes 4, 8, 12 and 16). Four probes were used as indicated: DNA_{IHF} (-142 to +8 relative to the *rtcBA* transcription start site +1), $DNA_{IHF\text{-}M7}$ (as DNA_{IHF} but with mutations in the putative IHF binding sites as described in Figure 18A for IHF_{M7}), $DNA_{IHF\text{-}M5}$ (as DNA_{IHF} but with mutations in the putative IHF binding sites as described in Figure 18A for IHF_{M5}) and $DNA_{control}$ (DNA from another part of the *E. coli* MG1655 genome). An EMSA was performed on a 6% native polyacrylamide gel.

($DNA_{control}$). The fragments were then incubated with two-fold serial dilutions of H_6-IHF (Figure 20). At a concentration of 600 nM H_6-IHF, a shift indicative of DNA-protein interaction was detectable for DNA_{IHF} (lane 1 compared to lane 4). Smaller shifts were also observed with 300 nM and 150 nM H_6-IHF (compare lanes 3 and 2 with 4 respectively). However, with DNA_{IHF-M7} and DNA_{IHF-M6}, smaller migrations were detectable at 600 nM and 300 nM H_6-IHF and neither fragment was shifted by 150 nM of protein (lanes 5 to 8 and 9 to 12). No interaction was detected between H_6-IHF and $DNA_{control}$ (lanes 13 to 16), indicating that the shifts detected with the *rtcBA* promoter DNA were specific.

This demonstrates that the affinity of H_6-IHF for *rtcBA* promoter DNA is reduced when *ihf1* and *ihf2* are mutated *in vitro*, therefore indicating that IHF binds to these sequences in order to mediate its effect on *rtcBA* transcription during RtcR activation, as indicated by Figure 17 and Figure 18.

3.2.4 Regulation of *rtcR* transcription

3.2.4.1 Mapping the *rtcR* promoter

With the aim of investigating the regulation of *rtcR* transcription, plasmid-based transcriptional *rtcR* fusions were made with the *lacZ* reporter gene using pGH254. As the location of the *rtcR* promoter is unknown, a series of *rtcR-lacZ* fusions were constructed containing sequentially less of the sequence upstream of *rtcR* and including the first 30 bp of the *rtcR* ORF (R3-0), as depicted in Figure 21A. The β-galactosidase activity of each fusion as a reporter of *rtcR* transcription was assessed in *E. coli* TB28 on NA plates containing X-gal. As Figure 21B shows, strains carrying fusions R3 to R1 produced blue colonies, while fusion R0 produced white colonies. This indicates that an essential element of the *rtcR* promoter is located in the 19 bp region between fusion R1, which begins 49 bp upstream of the first codon of *rtcR*, and R0, which starts 30 bp upstream of *rtcR*.

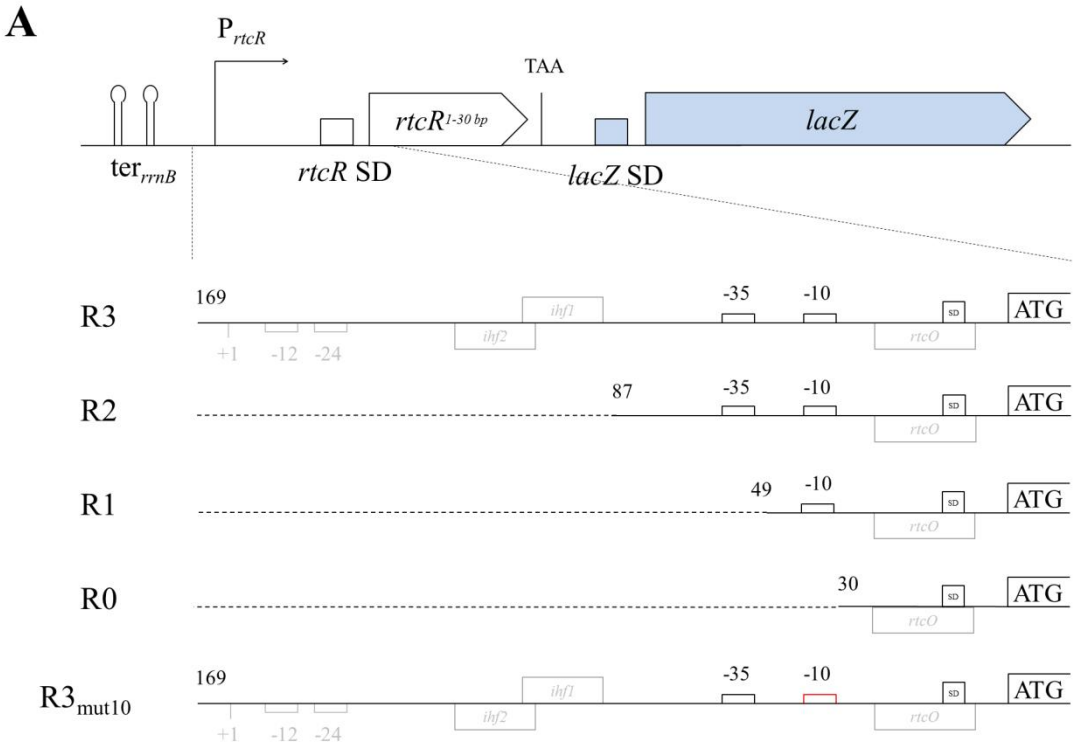
The majority of bEBPs are constitutively transcribed at low levels under the control of the housekeeping sigma factor, σ^{70}, however some are reported to also have a σ^{54}-controlled promoter, for example the two component systems ZraRS and NtrBC (Leonhartsberger et al., 2001; Magasanik, 1988). To this end, the β-galactosidase activity of fusion R3 was compared in *E. coli* TB28 and an isogenic *rpoN* deletion strain (Figure 21C). Colonies from TB28Δ*rpoN* were blue demonstrating that the presence or absence of σ^{54} does not affect *rtcR* transcription.

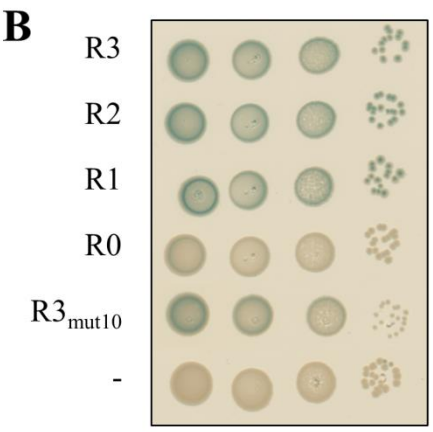
Unlike σ^{54}, σ^{70} family-associated RNAP recognises and binds to DNA elements centred around -10 and -35 relative to the transcription start site, which is unknown for *rtcR*. A search of the sequence upstream of *rtcR* identified sequences between 30 and 58 bp upstream of *rtcR* with similarity to and appropriate spacing of the σ^{70} consensus sequence (Shultzaberger et al., 2006). The R3 *rtcR-lacZ* fusion was mutated at the -10 element of the hypothetical *rtcR* promoter identified, as indicated in Figure 21D. The resulting fusion, $R3_{mut10}$, displayed reduced β-galactosidase activity on plates supplemented with X-gal compared to R3 to R1 (Figure 21B). However, the activity of $R3_{mut10}$ did not decrease to the same extent as R0, suggesting that the substitutions made to the hypothetical -10 do not completely eliminate *rtcR* transcription.

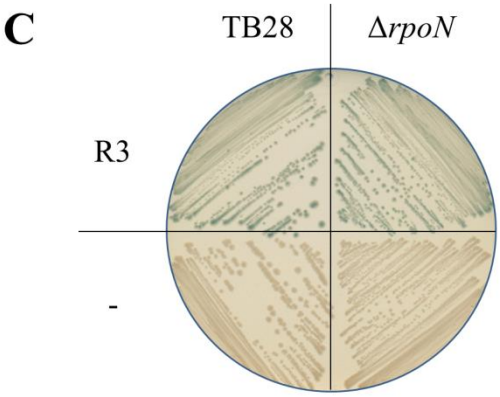
3.2.4.2 Evidence of negative auto-regulation of *rtcR* transcription

Owing to the close proximity of the *rtcR* binding region *rtcO* to the *rtcR* promoter (Figure 21A), it was hypothesised that RtcR may have a role in regulating its own expression. To investigate this possibility, β-galactosidase activity of the transcriptional R3 *rtcR-lacZ* fusion was monitored in *E. coli* TB28 carrying pGH332R for arabinose-inducible expression of RtcR and empty pBAD322K (Figure 22A). As observed in Figure 22A, strains carrying fusion R3 produced blue colonies on NA plates supplemented with X-gal (left panel, top two rows). However, in the presence of arabinose, colonies became white in the strain expressing RtcR (right panel, compare first and second rows), suggesting that overexpression of RtcR initiates a negative feedback to reduce *rtcR* transcription. *E. coli* TB28 carrying empty fusion plasmid pGH254 produced white colonies as expected (bottom two rows of each panel).

To determine whether the IR of *rtcO* identified as a potential RtcR binding site for *rtcBA* transcriptional activation also mediates negative autoregulation by RtcR binding, a system was established to sequester RtcR from the *rtcR* promoter using the high-copy-number plasmid pMG25. Three pMG25 variants were used for this purpose: pGH251, which contained the region upstream of *rtcR* including *rtcO,* pGH252, which was the same as pGH251 but lacked *rtcO* and empty pMG25. If RtcR directly regulated *rtcR* transcription by binding to an element in the *rtcR* promoter, then the presence of pGH251 may titrate RtcR and enable *rtcR* transcription to resume, producing a blue colony phenotype in cells carrying transcriptional *rtcR-lacZ* fusions. If RtcR binding were mediated via *rtcO* in this case, pGH252 would be unable to titrate RtcR and colonies would remain white, along with pMG25. To test this hypothesis, *E. coli* TB28 carrying pGH251, pGH252 or pMG25 with a


A
P rtcR
TAA
rtcR 1-30 bp
lacZ
ter rrnB
rtcR SD
lacZ SD
R3
R2
R1
R0
R3 mut10
169
87
49
30
-35
-10
ATG
SD
ihf1
ihf2
rtcO
+1
-12
-24


B
R3
R2
R1
R0
R3 mut10
-


C
TB28
ΔrpoN
R3
-

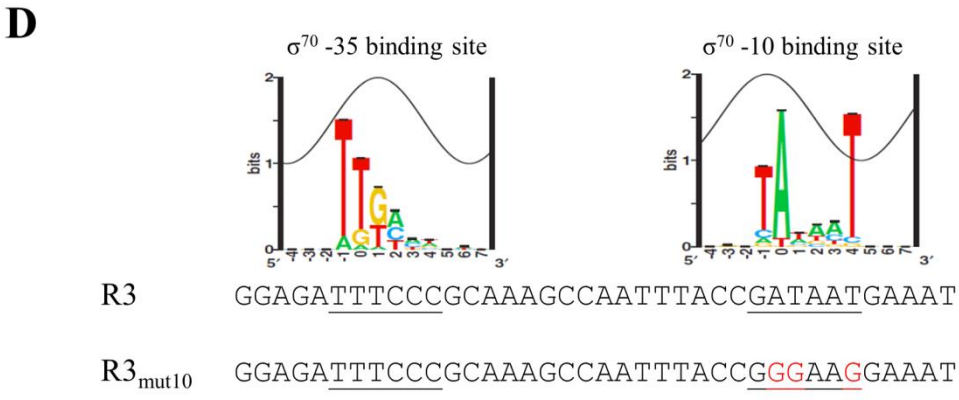

D
σ70 -35 binding site
σ70 -10 binding site
bits
R3 GGAGATTTCCCGCAAAGCCAATTTACCGATAATGAAAT
R3 mut10 GGAGATTTCCCGCAAAGCCAATTTACCGGGAAGGAAAT

Figure 21 Promoter requirements for *rtcR* transcription.

(A) Schematic organisation of transcriptional *rtcR-lacZ* fusions carried on pGH254R3 (R3), pGH254R2 (R2), pGH254R1 (R1), pGH254R0 (R0) and pGH254R3$_{mut10}$ (R3$_{mut10}$) (not to scale). Fragments consisting of the sequence upstream of *rtcR* were amplified and cloned into pGH254 to generate transcriptional fusions with the *lacZ* reporter gene. The R3 and R3$_{mut10}$ fusions begin 169 bp upstream of the *rtcR* start codon (ATG), R2 begins at 87 bp upstream, R1 at 49 bp and R0 at 30 bp, as indicated. All fragments end 30 bp (10 amino acids) into the *rtcR* ORF. The position of hypothetical -10/-35 elements for σ^{70} recognition are indicated (red indicates mutation of the -10 element). Features of the *rtcBA* promoter are shown in grey. (B) *E. coli* TB28 carrying fusions R0 to R3, mutated fusion R3$_{mut10}$ or empty pGH254 (-) were plated in 100-fold serial dilutions on NA plates containing 30 µg/ml ampicillin and 40 µg/ml X-gal, with or without 0.2% arabinose as indicated, and incubated at 37°C. (C) *E. coli* TB28 and TB28Δ*rpoN* carrying pGH254R3 (R3) or pGH254 (-) were plated on NA plates containing 30 µg/ml ampicillin, 20 µg/ml L-glutamine and 40 µg/ml X-gal and incubated at 37°C. (D) A hypothetical -35/-10 element upstream of *rtcR* compared to a sequence logo for σ^{70} binding (Shultzaberger et al., 2006). Mutations were made in the -10 site (red), generating R3 variant R3$_{mut10}$.

kanamycin-resistant version of fusion R3 and pGH322CR for expression of RtcR were grown on X-gal plates (Figure 22B). However, all strains maintained the white colony phenotype in the presence of arabinose. Furthermore, a comparison of the β-galactosidase activity of R3 in wild-type TB28 and a scarless *rtcR* deletion strain by β-galactosidase assay revealed that there was no difference in activity of the fusion in the presence or absence of chromosomal *rtcR* (Figure 22C).

3.3 Discussion

In this chapter, the regulation of *rtcBA* transcription in *E. coli* was investigated. Transcription of *rtcBA* was activated upon expression of RtcRΔN in a σ^{54}-dependent manner, verifying the role of RtcR as the bEBP of the *rtcBA* operon as well as the function of the N-terminal regulatory domain in negatively regulating RtcR activity in the absence of an activating signal, as previously described by Genschik et al. (1998). This property of RtcRΔN was used as a means to investigate DNA elements involved in *rtcBA* transcription including putative IHF binding sites and the unknown binding site of RtcR.

A series of progressively truncated transcriptional *rtcBA-lacZ* fusions identified a region upstream of the *rtcBA* promoter that, when deleted, prevented RtcRΔN-mediated transcriptional activation of *rtcBA.* Inspection of this region identified a candidate UAS of RtcR, *rtcO,* which was further investigated by mutagenising the IR and assessing the effect on *rtcBA-lacZ* transcription by RtcRΔN. β-galactosidase activity on both solid and liquid media

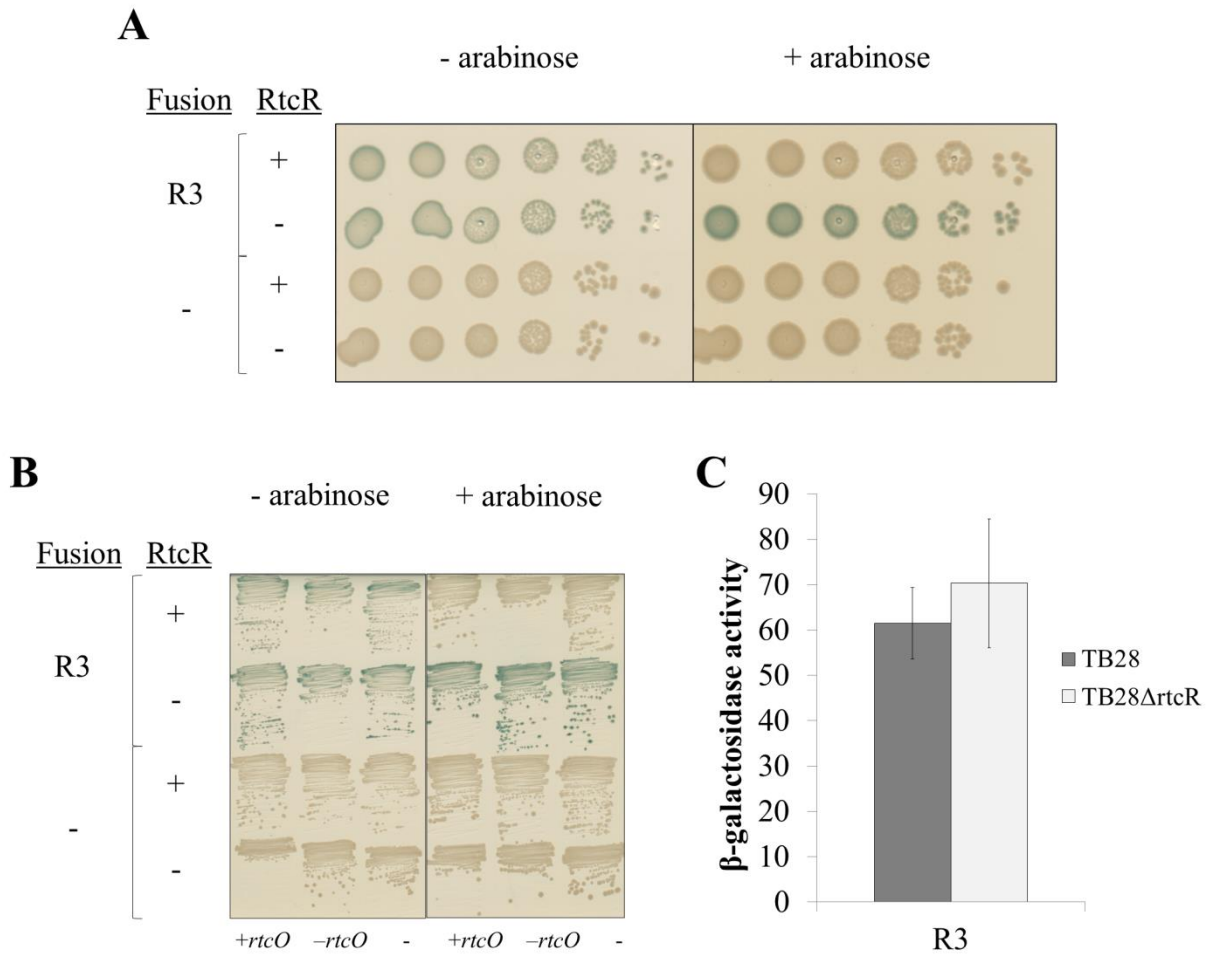


Figure 22 Negative autoregulation of *rtcR* transcription.

(A) *E. coli* TB28 carrying pGH254R3 (R3) or empty pGH254 (-) with pGH322R or pBAD322K (corresponding empty plasmid) for RtcR expression as indicated, were plated in 10-fold serial dilutions on NA plates containing 30 µg/ml ampicillin, 25 µg/ml kanamycin and 40 µg/ml X-gal, with or without 0.2% arabinose and incubated at 37°C. (B) *E. coli* TB28 carrying pGH254KR3 (R3) or empty pGH254K (-) with pGH322CR or pBAD322C (corresponding empty plasmid), for RtcR expression as indicated, and pGH251 (+*rtcO*), pGH252 (-*rtcO*) or pMG25 (-) were plated in all combinations on NA plates containing 25 µg/ml kanamycin, 50 µg/ml chloramphenicol, 100 µg/ml ampicillin and 40 µg/ml X-gal, with or without 0.2% arabinose as indicated and incubated at 37°C. (C) *E. coli* TB28 and TB28Δ*rtcR* carrying pGH254R3 (R3) or pGH254 were cultured in LB medium supplemented with 30 µg/ml ampicillin and sampled during exponential growth. β-galactosidase activity of each sample was measured and the values for empty pGH254 were deducted. β-galactosidase activity is expressed in Miller units and the values are an average of two independent experiments ± standard deviation.

revealed that substitution of three bases in each half of the IR in fusion BA3-$IR_{M7,}$ which changed the sequence but maintained its palindromic nature, abolished transcription of *rtcBA-lacZ.* However, BA3-IR_{M2} and BA3-IR_{M5}, which carried the same mutations in either the first or second half-site of the IR only, did not show similar reduction in β-galactosidase activity. Furthermore, substitution of all five bases in either the first or second half-site (fusions BA3-IR_{M1} and BA3-IR_{M4} respectively) failed to reduce β-galactosidase activity to the same extent, implying that maintenance of one IR half-site is sufficient to retain *rtcBA-lacZ* transcription. In particular, mutation of the first half-site alone had no impact on β-galactosidase activity, which is inconsistent with the finding that fusion BA1, which carries the downstream half-site of the IR only, shows no detectable induction of *rtcBA* transcription by RtcRΔN.

Taken together, the data suggests that *rtcO* is involved in mediating RtcR-dependent *rtcBA* transcription. However, it is also notable that some of the mutations made in *rtcO* failed to affect β-galactosidase activity of the *rtcBA-lacZ* fusions, suggesting that the ability of RtcR to initiate *rtcBA* transcription is not compromised. Moreover, based on the substitutions made in the IR, there is no clear pattern to indicate which bases would be likely to participate in an interaction with RtcR. It is possible that other bases surrounding the IR of *rtcO* could contribute to an interaction with RtcR, creating an imperfect IR, which could explain why some substitutions failed to affect *rtcBA* transcription. Another possibility is that RtcR has an additional UAS. In this scenario, the extra UAS would be unable to activate *rtcBA* transcription when *rtcO* is completely abolished, as is the case with BA3-IR_{M7}, but would be competent to mediate some degree of transcription when mutations in *rtcO* reduce but don't eliminate RtcR binding at this site. Therefore, while a potential UAS for RtcR has been identified, it is clear that further investigation is necessary to establish the extent of the role of the IR, whether other bases in the region also contribute to an extended *rtcO* and whether additional UASs are present.

Owing to the inconclusive results obtained from mutagenesis of the candidate *rtcO* sequence, we attempted to analyse the location and sequence of the UAS of RtcR *in vitro.* However, purified proteins RtcR, RtcRΔN and $RtcR_{C\text{-}term}$ were unable to form a complex with *rtcBA* promoter DNA including *rtcO* and the surrounding region. RtcR and RtcRΔN were largely insoluble so may not be sufficiently concentrated in the native form to produce a shift in the assay. The inability of $RtcR_{C\text{-}term}$ to elicit complex formation could indicate that the DNA binding domain of RtcR alone cannot form a stable dimer for DNA binding and that the major dimerisation determinants may be located in the regulatory or ATPase domains, as is the case for some bEBPs including NtrC1 and DctD (Sallai and Tucker, 2005; Vidangos et al., 2013).

In order to conduct *in vitro* analyses with RtcR, attempts to purify increased amounts of RtcR or RtcRΔN could be made by introducing denaturing and refolding steps into the purification protocol (Singh and Panda, 2005). In addition, *in vivo* DNA footprinting could represent an alternative method to identify regions of DNA protected by RtcR.

Transcriptional *rtcBA-lacZ* fusions were also used to investigate the role of IHF in *rtcBA* transcription. Independent deletion of the α and β subunits of IHF abolished *rtcBA-lacZ* transcription during RtcRΔN expression, demonstrating that IHF is required for RtcR-dependent *rtcBA* transcription. As two putative IHF binding sites, *ihf1* and *ihf2*, were previously identified upstream of the *rtcBA* operon (Genschik et al., 1998), we investigated whether one or both of these sites were necessary to mediate the IHF-dependency of *rtcBA* transcription during RtcRΔN expression. Fusions carrying mutations in either *ihf1* or *ihf2* alone remained competent for *rtcBA-lacZ* transcription. However, simultaneous mutation of *ihf1* and *ihf2* with an increasing number of substitutions was accompanied by a gradual decrease in β-galactosidase activity. In particular, fusions BA3-IHF$_{M1}$ and BA3-IHF$_{M2}$ had single base pair substitutions in *ihf1* or *ihf2* respectively and retained high β-galactosidase activity, while combination of both substitutions into BA3-IHF$_{M5}$ reduced activity to approximately 73% of wild-type fusion BA3. This suggests that one functional IHF site is sufficient for efficient *rtcBA* transcription, but that either *ihf1* or *ihf2* may be able to fulfil this role. In support of this hypothesis, fusion BA3-IHF$_{M3,}$ which introduced five substitutions in *ihf2*, had similar activity to BA3, but in combination with mutation of *ihf1* in fusion BA3-IHF$_{M7}$ produced a severe reduction in *rtcBA* transcription. To verify this, a similar fusion with mutations in *ihf1* only is required to confirm that *ihf2* is sufficient for *rtcBA* transcription in the absence of *ihf1*.

The greatest reduction in β-galactosidase activity during RtcRΔN expression was observed with BA3-IHF$_{M7}$, which carried 12 substitutions across both *ihf1* and *ihf2*. Nevertheless, the activity was still higher than that of fusion BA0, which lacks *rtcO*, indicating that even when a substantial number of substitutions are introduced into *ihf1* and *ihf2*, *rtcBA* transcription is not completely inhibited. Accordingly, purified IHF was able to form a complex with *rtcBA* promoter DNA containing the same mutations as BA3-IHF$_{M7}$ *in vitro*. However, the complex formed between IHF and DNA containing wild-type *ihf1* and *ihf2* at equivalent protein concentrations appeared to be more stable in comparison and was maintained at lower concentrations of IHF. This demonstrates that, while some IHF binding is still able to occur and mediate the low levels of *rtcBA* transcription observed for BA3-IHF$_{M7}$, IHF has a greater affinity for wild-type *ihf1* and *ihf2* sequences.

Together, the data shows that RtcR-dependent *rtcBA* transcription requires functional IHF, which confers its effect in a DNA-dependent manner by binding to *ihf1* and/or *ihf2*. It would be of interest to ascertain whether the role of IHF in *rtcBA* transcription is solely to mediate correct interfacing between RtcR and the σ^{54}-RNAP holoenzyme or if it has additional functions, as described for some other σ^{54}-dependent promoters (Chapter 1.3.3). For example, IHF is known to aid binding of the bEBP PspF to its UASs and in a reciprocal arrangement, PspF also facilitates IHF binding to DNA (Jovanovic and Model, 1997). Cooperative binding of RtcR and IHF could provide another possible explanation for the lack of complex formation between $RtcR_{C\text{-}term}$ and *rtcBA* promoter DNA and could also explain the high concentrations of IHF necessary to detect complex formation here, compared to previous studies (Jovanovic and Model, 1997; Sze et al., 2001).

Transcriptional regulation of *rtcR* was studied using a similar approach to that used to identify the RtcR binding region in the *rtcBA* promoter. The transcriptional *rtcR-lacZ* fusion R3 was found to have moderate β-galactosidase activity in standard conditions, indicating that transcription may be regulated by the housekeeping σ^{70} factor. As the fusion retained the same activity in an *rpoN* deletion strain, we were able to ascertain that the transcription observed from R3 was not σ^{54}–dependent. Upon inspection of the sequence upstream of the *rtcR* ORF, potential -35/-10 promoter elements were identified and, accordingly, mutation of the -10 motif was found to decrease β-galactosidase activity from the R3 variant $R3_{mut10}$. However, *rtcR-lacZ* fusions containing successively less of the upstream sequence revealed that the shortest fusion retaining similar β-galactosidase activity to R3 and thus representing the minimal *rtcR* promoter, lacked the -35 element of the hypothetical promoter. Combined with the fact that fusion $R3_{mut10}$ does not reduce *rtcR-lacZ* transcription to the same extent as R0, the longest fusion lacking β-galactosidase activity, despite the introduction of three substitutions in the hypothetical -10, this suggests that the potential σ^{70} promoter identified here may be incorrect. This could be resolved by using primer extension analysis to map the transcription start site of *rtcR.*

Expression of RtcR from a plasmid was found to decrease the β-galactosidase activity of fusion R3, indicating that transcription of *rtcR* is negatively autoregulated. This negative feedback mechanism of regulation has also been described for other bEBPs, including NtrC, PspF and CrbB, and is mediated via binding of the bEBPs to DNA sequences overlapping their σ^{70} promoter elements, enabling constant and relatively low levels to be maintained (Abdou et al., 2011; Jovanovic et al., 1997; Magasanik, 1988). As *rtcO* is in close proximity to the start site of *rtcR,* we reasoned that a similar mechanism of transcriptional regulation

might exist for *rtcR*. However, no difference was observed in the β-galactosidase activity of fusion R3 between wild-type *E. coli* and an *rtcR* deletion, indicating that the absence of RtcR does not result in derepression of the *rtcR* promoter. Moreover, attempts to titrate RtcR using a high copy number plasmid carrying the sequence upstream of *rtcR* (including *rtcO*) did not recover the activity of fusion R3 during RtcR expression. This suggests that either RtcR was insufficiently titrated by the plasmid, not able to be titrated by the sequence carried on the plasmid or that promoter DNA binding by RtcR is not involved in regulating *rtcR* transcription. Therefore, at present it is unclear how the negative autoregulation of *rtcR* is mediated. In order to establish whether the binding of RtcR to promoter DNA is involved in this regulatory feedback mechanism, truncated RtcR lacking the DNA binding domain could be expressed. In the event that *rtcR* transcription was not inhibited by this RtcR variant, this would support the hypothesis that RtcR negatively regulates *rtcR* transcription by binding to *rtcO* (or another UAS of RtcR). Alternatively, transcription of *rtcR* could be regulated indirectly by another protein, however the titration experiment would imply that this is also not mediated by DNA binding.

In summary, the work presented in this chapter has further elucidated the transcriptional regulation of the *rtcBA* operon of *E. coli* K-12. The minimal promoter required for RtcR-dependent *rtcBA* transcription by the σ^{54}-RNAP holoenzyme has been identified and putative RtcR binding region, *rtcO* has been characterised by mutagenesis. Furthermore, an essential role for IHF in the activation of *rtcBA* transcription by RtcR has been confirmed and found to be dependent on interaction with at least one of the IHF binding sites in the *rtcBA* promoter. Finally, the minimal promoter required for *rtcR* transcription has been identified and is likely to be regulated by σ^{70}-associated RNAP and may be subject to an autoregulatory feedback mechanism.

Chapter 4 Activation of *rtcBA* transcription in *E. coli*

4.1 Introduction

In the previous chapter, it was established that an N-terminally truncated version of RtcR (RtcRΔN) constitutively activates transcription from the *rtcBA* promoter, while wild-type RtcR does not, which is consistent with the findings of Genschik et al. (1998). This suggests that, in accordance with other bEBPs, the N-terminal regulatory domain of RtcR has a sensory capacity and confers a level of control over *rtcBA* transcription that restricts the activity of RtcB to specific conditions.

The regulatory domain of RtcR has no similarity with those of other bEBPs (Genschik et al., 1998), suggesting that it detects a novel type of ligand or receives sensory information in a distinct way. Instead, the regulatory domain was found to have similarity to the CARF domains of a group of CRISPR-cas-associated proteins, which are predicted to bind nucleotide or nucleotide-derived molecules (Makarova et al., 2014). Although the CARF-like domain of RtcR has diverged from those associated with CRISPR-cas, this raises the possibility that the signal detected by RtcR could be single nucleotides or perhaps oligonucleotides released from damaged or cleaved RNA. However, the conditions in which these or other activating signals might be created are unknown.

In this chapter, we use genetic screening and quantitative PCR to identify conditions in which transcription of *rtcBA* is activated in *E. coli.*

4.2 Results

4.2.1 The *rtc* genes are non-essential for growth in laboratory conditions

To establish whether the *rtc* genes are dispensible in *E. coli* under standard laboratory conditions, the growth of *rtcB, rtcA, rtcR, rtcBA* and *rtcBAR* deletion strains were compared to that of wild-type MG1655 in LB media at 37°C. In contrast to a recent publication, which showed that deletion of *rtcB* or *rtcA* decreased the growth rate of *E. coli* (Engl et al., 2016), no difference in growth was observed between any of the strains (Figure 23). This outcome is

supported by Temmel et al. (2016) and Genschik et al. (1998), who reported no difference in the growth rates of *rtcB* and *rtcA* null strains respectively in laboratory conditions. The *rtc* deletion strains were also cultured in minimal media with glucose or glycerol as carbon sources and in the presence and absence of casamino acids. In all cases, there was no difference in growth (data not shown). In combination with the low levels of *rtcBA* transcription observed in the absence of RtcRΔN in standard laboratory conditions (Chapter 3), this suggested a tightly-regulated function for RtcB activated in response to stress in *E. coli*, such as repair of stress-induced RNA damage.

4.2.2 Genetic screening for regulators of *rtcBA* expression

4.2.2.1 Screening for inhibitors of *rtcBA* transcription

In order to investigate whether the disruption of any single genes in *E. coli* activate *rtcBA* transcription, transposon mutagenesis was performed using a kanamycin-resistant mini-Tn10 transposon carried on suicide plasmid pNKBOR (Rossignol et al., 2001). Initially, pNKBOR was electroporated into *E. coli* TB28 carrying pGH254BA2, encoding transcriptional *rtcBA-lacZ* fusion BA2. As shown in Chapter 3, β-galactosidase activity of BA2 was increased during RtcRΔN expression, suggesting it has all essential *rtcBA* promoter elements, but had low activity in its absence, suitable for use in genetic screening. Thus, upon spreading of electroporated cells on X-gal-supplemented plates containing kanamycin for selection of transposition events, blue colonies were indicative of genetic disruption caused by mini-Tn10 that activated the *rtcBA-lacZ* fusion. Approximately 10,000 colonies were screened, of which twenty were blue. However, plasmids purified from these strains were kanamycin resistant and remained blue when re-transformed into strain TB28, indicating that in all cases the transposon had integrated into pGH254BA2.

To avoid the use of a reporter plasmid in genetic screening and circumvent associated problems including transposition into plasmid DNA and identification of genes involved in plasmid regulation as opposed to candidate *rtcBA* regulators, a transcriptional *rtcBA-lacZ* fusion on the *E. coli* chromosome was generated. Having established that the resulting strain, TB28*rtcBA-lacZ*, was functional by expressing RtcRΔN (Chapter 3.2.3), a library of single gene mutants was generated and screened for blue colonies as described above. Of approximately 35,000 colonies screened, 12 were blue. P1 phage transduction was used to transfer the transposon from the 12 blue strains into clean TB28*rtcBA-lacZ*. Four strains

retained the blue-colony phenotype, while the other eight produced white colonies, suggesting that the transposon from pNKBOR produced unstable phenotypes at high frequency.

In order to map the site of transposon insertion in the four strains with stable phenotypes, arbitrary PCR was employed. For those strains where arbitrary PCR was unsuccessful, the self-cloning properties of the transposon were exploited (Rossignol et al., 2001). The PCR products or plasmids generated respectively were purified and sequenced with transposon-specific primers. All the transposition sites that stably activated the transcriptional *rtcBA-lacZ* fusion were in ORFs and are listed in Table 5.

For confirmation of the phenotypes observed with transposon mutagenesis, the relevant strains from the Keio collection of mutants (Baba et al., 2006) were used to generate single gene deletions of the four candidate genes in *E. coli* TB28 carrying pGH254BA2. Deletions in *ycfT* and *lpp* produced white colonies on plates containing X-gal, comparable to the wild-type strain carrying fusion BA2. The only strains producing blue colonies were those with deletions of *tolC* and *acrA*. However, when these deletions were combined with an *rtcR* deletion, colonies from TB28Δ*rtcR-tolC* and TB28Δ*rtcR-acrA* carrying pGH254BA2 remained blue, indicating that transcription of the fusion was independent of RtcR (data not shown).

As the genes identified in the screen were demonstrated to have either weak phenotypes (*ycfT* and *lpp*) or RtcR-independency (*tolC* and *acrA*) and their functions indicated no clear link with the activity of an RNA ligase (summarised in Table 5), they were not considered good candidates to further investigate the activation of *rtcBA* transcription. To account for any transposition bias of the mini-Tn10 from pNKBOR and to ensure coverage of the genome was sufficiently saturated, a second transposon carried on suicide plasmid pMiniHimar (Rubin et al., 1999) was used to generate and screen a library of single gene mutants in the same way. However, no blue colonies were observed in this screen.

4.2.2.2 Screening for activators of *rtcBA* transcription

To search for proteins that induce *rtcBA* transcription, genetic screening of a genomic library of *E. coli* was performed. The library was constructed by purifying genomic DNA from strain TB28 and partially digesting with restriction enzyme Sau3AI, which has approximately 19,000 restriction sites in the *E. coli* genome. DNA fragments of the appropriate size (2 to 6 Kb) were cloned into BamHI-digested pBR322, as BamHI and Sau3AI produce compatible

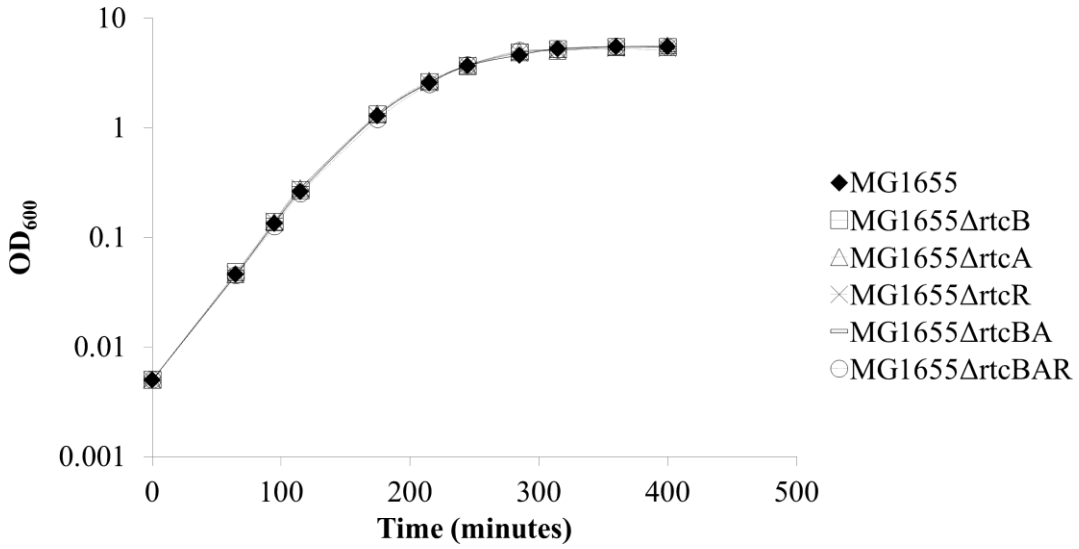


Figure 23 Deletion of the *rtc* genes has no effect on growth of *E. coli* MG1655.

E. coli strains MG1655, MG1655Δ*rtcB*, MG1655Δ*rtcA*, MG1655Δ*rtcR*, MG1655Δ*rtcBA* and MG1655Δ*rtcBAR* were cultured in LB medium and OD_{600} measured at regular intervals.

Table 5 Candidate genes for inhibiting *rtcBA* transcription.

Site of mini-Tn10 insertion	Gene function	Colony phenotype	Confirmation with Keio strain	RtcR-dependency
tolC	Outer membrane channel	5	Yes	No
acrA	Periplasmic protein	3	Yes	No
ycfT	Inner membrane protein	2	No	-
lpp	Outer membrane lipoprotein	2	No	-

('Gene functions' were obtained from the EcoCyc database (Keseler et al., 2011). 'Colony phenotype' was ranked on a scale of 1 to 10 where 1 is the β-galactosidase activity of TB28*rtcBA-lacZ* alone and 10 is the activity of TB28*rtcBA-lacZ* during overexpression of RtcRΔN.)

overhangs. Strain TB28*rtcBA-lacZ* was transformed with the ligation mixture and cultured on plates containing ampicillin, for selection of the pBR322 library, and X-gal. The resulting colonies were screened for multicopy activation of *rtcBA-lacZ* transcription. Of the approximately 12,000 colonies screened, 27 were blue. Plasmids were purified from the blue colonies (Untergasser, 2006), transformed into fresh TB28*rtcBA-lacZ* cells and plated onto X-gal containing NA to assess the stability of the phenotypes. Thirteen strains retained the blue-colony phenotype and the plasmids from these strains (named pGH01 to pGH13) were sequenced with pBR322-specific primers to identify the genomic DNA fragments carried by each. All the complete ORFs carried in pGH01 to pGH13 are listed in Table 6.

To confirm the phenotypes observed from TB28*rtcBA-lacZ* carrying pGH01 to pGH13 and, in the case of plasmids carrying multiple ORFs, to attempt to identify the relevant gene responsible, the ASKA collection of plasmids was used (Kitagawa et al., 2005). However, when expressed individually from the relevant ASKA plasmid, none of the candidate genes were able to activate the chromosomal *rtcBA-lacZ* fusion (data not shown). The reason for this is unclear. It may be that co-expression of some of the genes identified is required to activate the transcriptional fusion or the function of some proteins is altered by the presence of an N-terminal His_6 tag from the ASKA plasmid.

In order to determine whether the activation of the chromosomal *rtcBA-lacZ* fusion by pGH01 to pGH13 was RtcR-dependent, TB28 and TB28Δ*rtcR* carrying pGH254KBA2 (a kanamycin-resistant plasmid encoding transcriptional *rtcBA-lacZ* fusion BA2) in combination with pGH01 to pGH13 were plated on NA supplemented with X-gal. In the majority of cases, both wild-type and *rtcR* deletion strains produced blue colonies, suggesting that transcription of the fusion was independent of RtcR. However, deletion of *rtcR* in strains carrying pGH10, pGH12 or pGH13 changed the colony phenotype from blue to white, indicating that activation of the transcriptional BA2 fusion was RtcR-dependent (Figure 24B). This was further confirmed with an *rtcR* null strain, MG1655*rtcR*^{M1X}, in which the start codon of *rtcR* was mutated to a stop codon (Figure 24B). As plasmids pGH10, pGH12 and pGH13 all contained a region of the *E. coli* genome encompassing the *yaeP, rof* and *tilS* genes (Figure 24A), these were considered promising candidates for further investigation.

4.2.3 Rof overexpression activates *rtcBA* transcription

Plasmids pGH10, pGH12 and pGH13 all contained complete ORFs of the *yaeP*, *rof* and *tilS* genes. Because the relevant ASKA plasmids were not able to reproduce the phenotype

observed in the screen, *yaeP*, *rof* and *tilS* were cloned individually onto pBAD33, generating plasmids pGH33yaeP, pGH33rof, pGH33rof$_{ext}$ and pGH33tilS respectively. As annotation of yaeP and rof on the EcoCyc database (Keseler et al., 2011) indicates they are co-transcribed from the same promoter, a plasmid containing the *yaeP-rof* operon, pGH33rof-yaeP, was also constructed. In order to establish which gene or combination of genes were responsible for activating transcription of the *rtcBA-lacZ* fusion, TB28*rtcBA-lacZ* carrying each plasmid was plated on NA supplemented with X-gal and arabinose to induce expression of the respective genes. As Figure 25A shows, the plasmids encoding *rof* alone and *yaeP-rof* produced blue colonies with the same unusual appearance as produced from pGH10, pGH12 and pGH13 (compare to Figure 24B). The stronger phenotype observed from Rof expression from pGH33rof is due to the presence of an optimised SD sequence, whereas pGH33rof$_{ext}$ and pGH33yaeP-rof expressed Rof from its native SD. As the strains expressing TilS or YaeP alone were white on X-gal plates, this demonstrated that Rof was able to activate transcription of the *rtcBA-lacZ* chromosomal fusion.

To further confirm this, the effect of Rof expression was compared in strain TB28 carrying transcriptional *rtcBA-lacZ* fusions BA2 and BA0 on plates supplemented with X-gal and arabinose. Rof expression in the strain carrying fusion BA2 produced blue colonies, while BA0 produced colonies sick in appearance that remained white, indicating that the unusual colony appearance was associated with Rof rather than *rtcBA* transcription (Figure 25B). The empty pBAD33 controls produced white colonies of normal appearance. As fusion BA2 included *rtcO*, the region of the *rtcBA* promoter necessary for RtcR responsiveness, whereas BA0 lacked *rtcO*, this established that Rof expression activated *rtcBA* transcription in an RtcR-dependent manner.

Transcriptional regulation of the *rtcBA* operon by Rof was also investigated in liquid culture: *E. coli* TB28*rtcBA-lacZ* carrying pGH33rof or pBAD33 was cultured in the presence of arabinose and β-galactosidase activity was assayed over a period of five hours (Figure 26A). No difference in β-galactosidase activity from the *rtcBA-lacZ* fusion was observed after one hour of Rof expression compared to activity just prior to induction. However, β-galactosidase activity progressively increased from 8.12 to 66.1 Miller units between two and five hours post-induction, when the cultures were entering stationary phase. As a small increase in the activity of *TB28rtcBA-lacZ* carrying pBAD33 was also observed over five hours, β-galactosidase activity was measured in a *rof* deletion strain to ascertain whether this was due to an increase in expression of Rof from the chromosome. However, there was no difference detected between TB28*rtcBA-lacZ* and TB28*rtcBA-lacZ*Δ*rof* (data not shown).

Table 6 Candidate genes for inducing *rtcBA* transcription.

Plasmid	Gene(s)	Gene function	Colony phenotype	Confirmation with ASKA plasmid	RtcR-dependency
pGH01	*yraI*	Predicted fimbrial chaperone	2	No	No
pGH02	*yraK*	Predicted fimbrial protein	3	No	No
pGH03	*yhdV*	Predicted outer membrane lipoprotein	3	No	No
pGH04	*ygiV*	Transcriptional repressor	3	No	No
	mqsR	Toxin (mRNase)			
	mqsA	Antitoxin			
pGH05	*mreC*	Cell shape maintenance	4	No	No
	mreD	Cell shape maintenance			
	yhdE	NTP pyrophosphatase			
pGH06	*tsgA*	Putative transporter	4	No	No
	ppiA	Prolyl isomerase			
	ficA	Antitoxin			
pGH07	*yhdH*	Acrylyl-CoA reductase	6	No	No
pGH08	*panF*	Pantothenate/Na^+ symporter	5	No	No
	yhdT	Inner membrane protein			
pGH09	*yihV*	Sulfoquinovose kinase	2	No	No
	yihW	Transcriptional regulator			
pGH10	*yaeP*	Unknown	7	No	Yes
	rof	Antiterminator protein			
	tilS	$tRNA^{Ile}$ synthetase			
pGH11	*yfaL*	Predicted adhesin	6	No	No
pGH12	*yaeP*	Unknown	8	No	Yes
	rof	Antiterminator protein			
	tilS	$tRNA^{Ile}$ synthetase			
	yaeR	Predicted lyase			
pGH13	*yaeP*	Unknown	8	No	Yes
	rof	Antiterminator protein			
	tilS	$tRNA^{Ile}$ synthetase			

('Gene functions' were obtained from the EcoCyc database (Keseler et al., 2011). 'Colony phenotype' was ranked on a scale of 1 to 10 where 1 is the β-galactosidase activity of TB28*rtcBA-lacZ* alone and 10 is the activity of TB28*rtcBA-lacZ* during overexpression of RtcRΔN.)

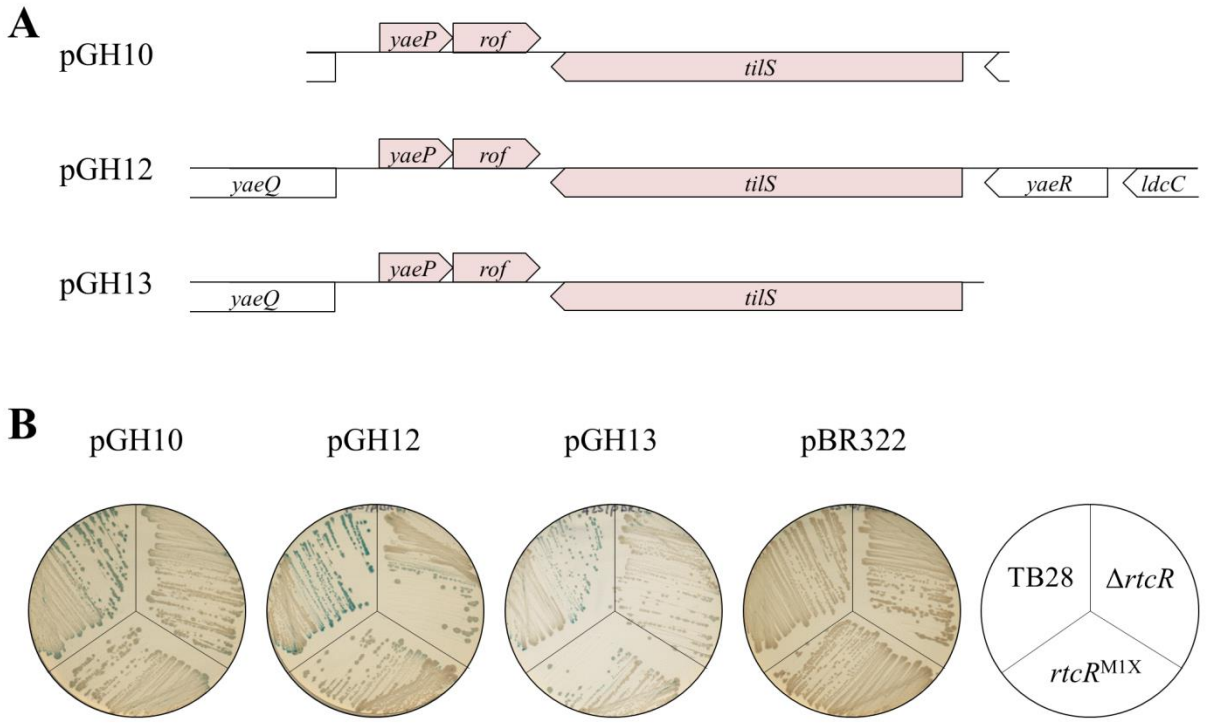


Figure 24 Activation of the transcriptional *rtcBA-lacZ* fusion by the *yaeP-rof-tilS* locus is RtcR dependent.

(A) Schematic organisation of the regions of the *E. coli* genome carried on pGH10, pGH12 and pGH13. Complete genes in all three plasmids are shaded. (B) *E. coli* TB28 and TB28Δ*rtcR* carrying reporter plasmid pGH254KBA2 (transcriptional *rtcBA-lacZ* fusion BA2) with pBR322 or variants pGH10, pGH12 or pGH13 identified by genetic screening were plated on NA containing 25 µg/ml kanamycin, 100 µg/ml ampicillin and 40 µg/ml X-gal and incubated at 37°C for 24 hours.

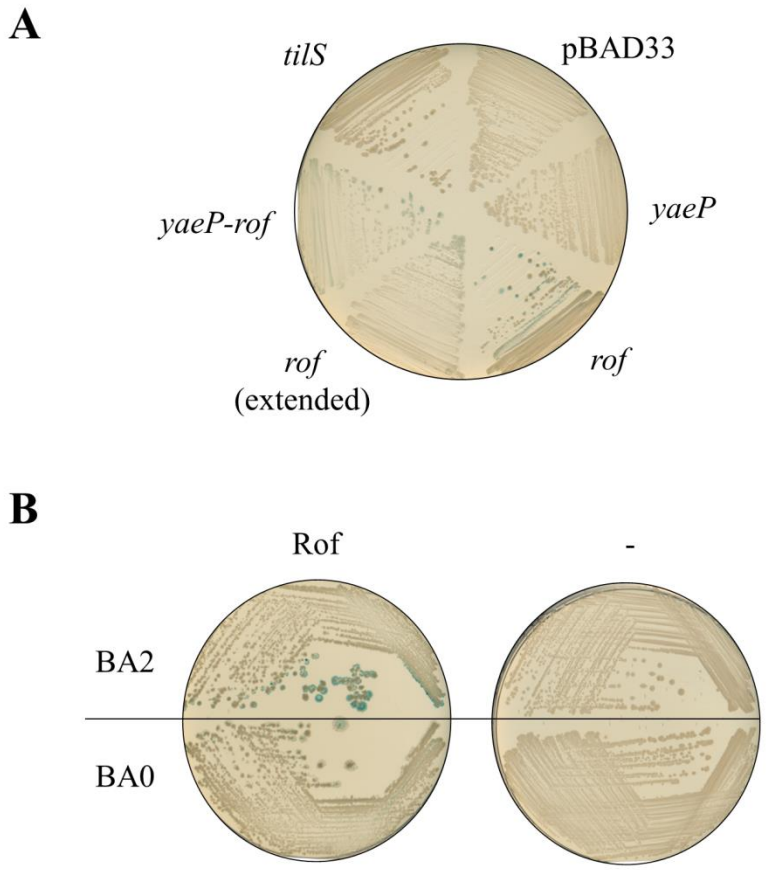


Figure 25 Expression of Rof induces transcription of the *rtcBA-lacZ* fusion in an RtcR-dependent manner.

(A) *E. coli* TB28*rtcBA-lacZ* carrying pGH33yaeP (YaeP expression), pGH33rof (Rof expression), pGH33rof$_{ext}$ (Rof and *rof-tilS* intergenic region), pGH33yaeP-rof (YaeP and Rof expression) pGH33tilS (TilS expression) or pBAD33 (corresponding empty plasmid) were plated on NA supplemented with 50 µg/ml chloramphenicol, 40µg/ml X-gal and 0.2% arabinose. Plates were incubated at 37°C for 24 hours. (B) *E. coli* TB28 carrying pGH254BA2 or pGH254BA0 (containing transcriptional *rtcBA-lacZ* fusions BA2 and BA0 respectively) with pGH33rof (Rof) or pBAD33 (-) in all combinations were plated on NA supplemented with 30 µg/ml ampicillin, 50 µg/ml chloramphenicol, 40µg/ml X-gal and 0.2% arabinose. Plates were incubated at 37°C for 24 hours.

The growth of *E. coli* strains MG1655, MG1655Δ*rtcBA* and MG1655Δ*rtcB* carrying pGH33rof or pBAD33 was monitored over the same timescale as the β-galactosidase assay (Figure 26B). This showed that, while no difference in growth was observed between the wild-type strain and the *rtcBA* and *rtcB* deletions, Rof expression reduced the rate of growth and led to early entry into stationary phase. Furthermore, the onset of the gradual decline in growth occurred approximately two hours after induction of Rof, coinciding with the activation of *rtcBA* transcription by Rof in Figure 26A.

To investigate whether RtcB was involved in mediating recovery following Rof expression, cells were reseeded in fresh media without the inductive agent after 120 and 240 minutes of Rof expression. Although the cells previously exposed to Rof expression had an extended lag phase compared to empty pBAD33 controls, wild-type, *rtcB* or *rtcBA* deletion strains resumed growth at the same time (data not shown). When Rof was expressed from a higher-copy-number plasmid pGH25rof (a variant of pMG25) the growth rate decreased more rapidly than observed from pGH33rof and again coincided with an increase in *rtcBA-lacZ* transcription. However, even following exposure to higher levels of Rof, the outgrowth of strains in fresh media failed to reveal any difference in recovery between MG1655 and the isogenic *rtcB* and *rtcBA* deletions (data not shown).

4.2.4 Agents that target translation induce *rtcB* transcription

4.2.4.1 Chloramphenicol treatment activates *rtcB* transcription

In the search for conditions where RtcB is expressed and thereby possibly functional in *E. coli*, strain MG1655 was treated with a range of antibiotics for 30 minutes and relative changes in expression of the *rtcBA* operon were assessed by qPCR using *rtcB*-specific primers. Figure 27A shows that antibiotics targeting translation more than doubled the levels of *rtcB* mRNA, compared to untreated cells. The greatest increase of approximately 6.5-fold was observed following chloramphenicol treatment, which inhibits translation by binding to 23S rRNA of the 50S ribosomal subunit and preventing peptide bond formation (Schlunzen et al., 2001). Tetracycline and kanamycin, which both target 16S rRNA of the 30S subunit (Brodersen et al.; François et al., 2005), resulted in an approximate 4-fold and 2.5-fold increase in *rtcB* transcript levels respectively. By comparison, treatment with ampicillin, which inhibits cell wall synthesis (Drawz and Bonomo, 2010), did not significantly change *rtcB* transcription.

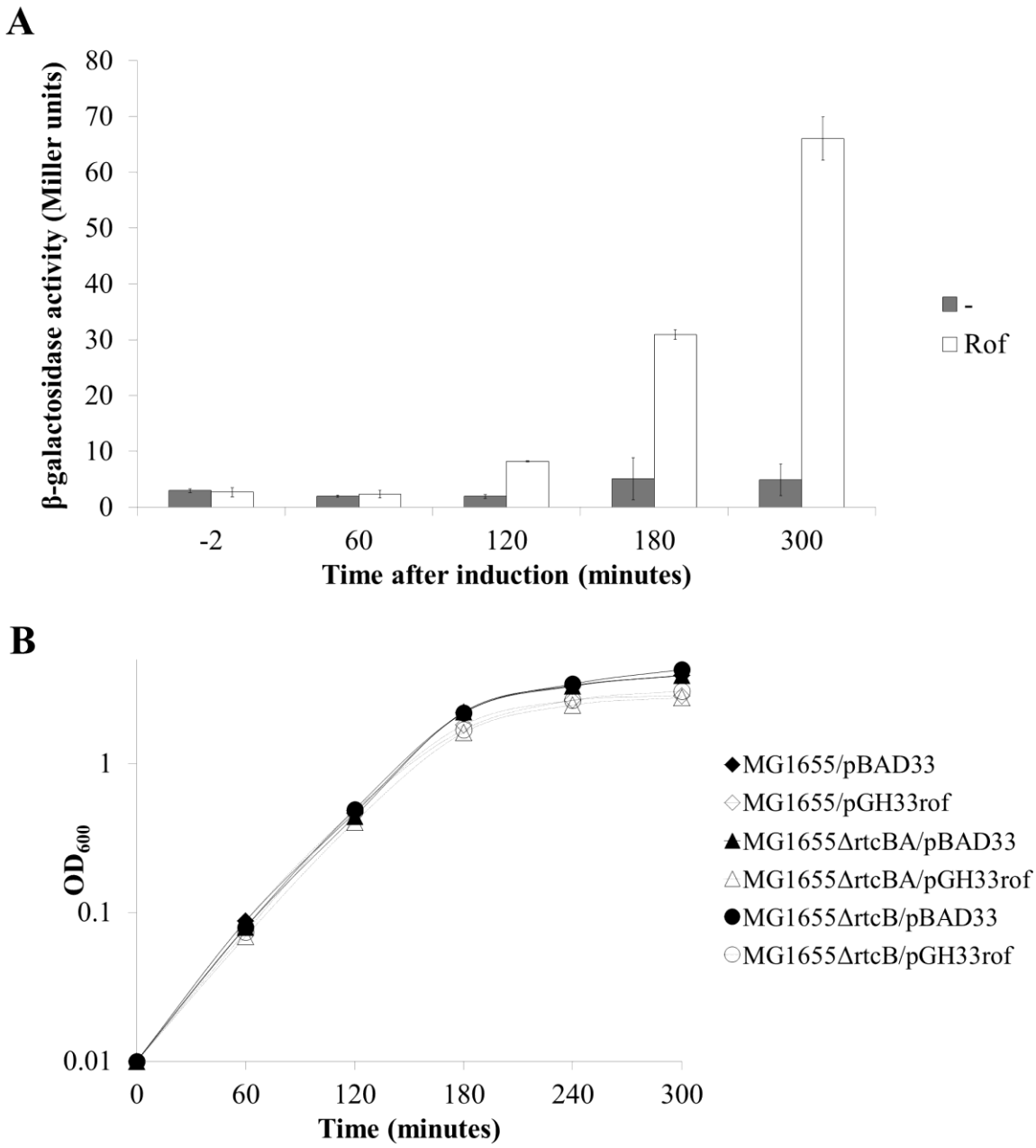


Figure 26 Transcription of the *rtcBA-lacZ* fusion increases concurrently with a reduction in growth of *E. coli* strains during Rof expression.

(A) *E. coli* TB28*rtcBA-lacZ* carrying pGH33rof (Rof) or pBAD33 (corresponding empty plasmid; -) were cultured in LB medium containing 50 µg/ml chloramphenicol and 0.2% arabinose. Samples were taken just prior to inoculation of cultures in medium containing arabinose (-2 minutes) and over a period of five hours (60, 120, 180 and 300 minutes) and β-galactosidase activity was measured. β-galactosidase activity is expressed in Miller units and the values are an average of three independent experiments ± standard deviation. (B) *E. coli* MG1655, MG1655Δ*rtcBA* and MG1655Δ*rtcB* carrying pGH33rof or pBAD33 were cultured in LB medium containing 50 µg/ml chloramphenicol and 0.2% arabinose and OD_{600} measured at regular intervals. The values are an average of three independent experiments.

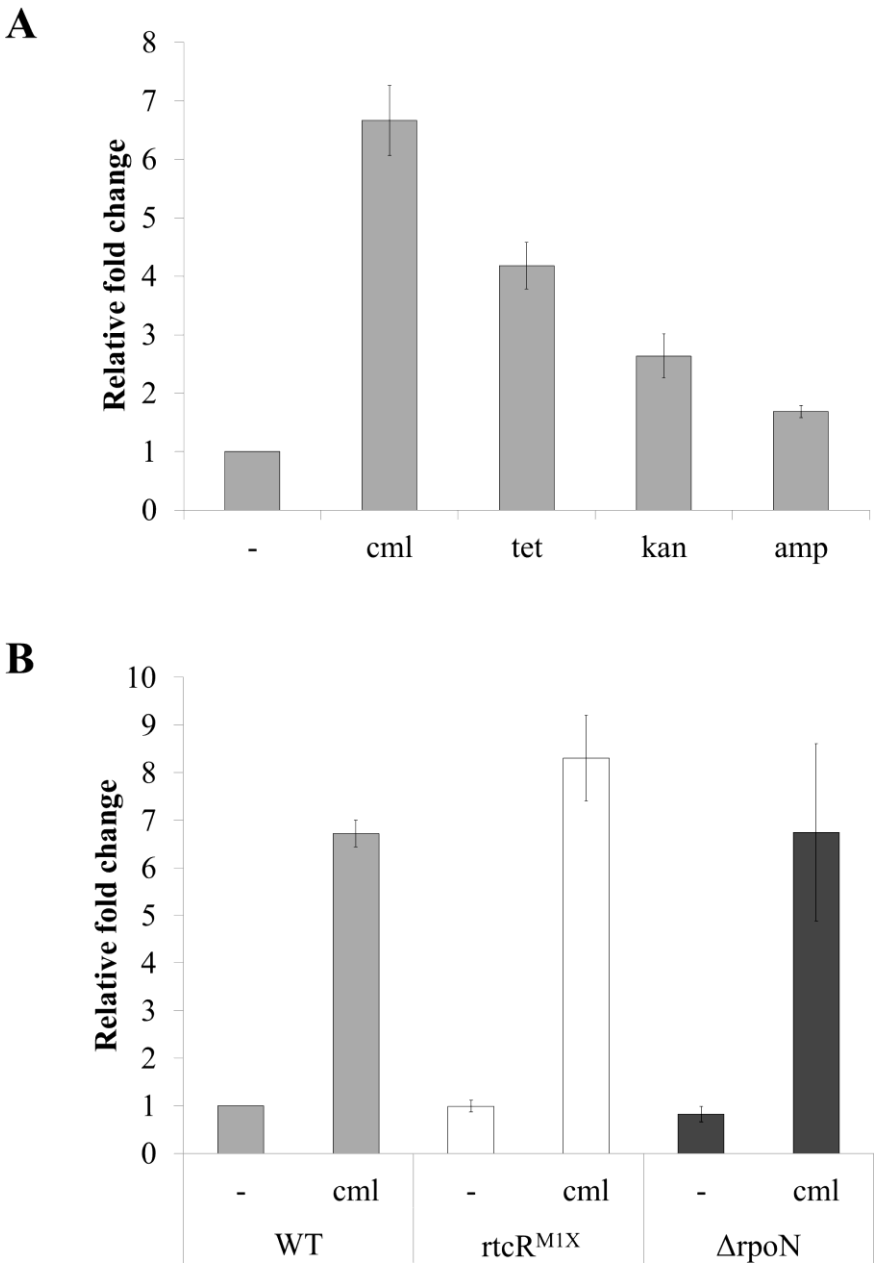


Figure 27 Levels of *rtcB* mRNA increase during treatment with chloramphenicol.

(A) *E. coli* MG1655 was grown exponentially in LB medium then treated with 50 µg/ml chloramphenicol (cml), 10 µg/ml tetracycline (tet), 25 µg/ml kanamycin (kan) or 30 µg/ml ampicillin (amp) for 30 minutes. RNA was extracted from each sample, reverse transcribed and the fold change in transcript levels of *rtcB* in MG1655 treated with antibiotic relative to untreated MG1655 (-) was determined by qPCR. qPCR data were analysed using the Pfaffl method and normalised against the *tatA* gene. Data from three independent experiments are shown as the mean relative fold change ± standard deviation. (B) Exponentially growing MG1655 (WT), MG1655*rtcR*^{M1X} (rtcR^{M1X}) and MG1655Δ*rpoN* (ΔrpoN) were treated with 50 µg/ml chloramphenicol (cml) for 30 minutes and *rtcB* transcription was analysed by qPCR as described for (A).

Having been confirmed with two additional reference genes for normalisation of qPCR data (Appendix A), further analysis of the effect of chloramphenicol treatment on RtcB was undertaken. Treatment of an *rtcR* null strain MG1655*rtcR*^{M1X} revealed similar increases in *rtcB* mRNA to that observed in wild-type MG1655, indicating that transcription of *rtcB* is independent of RtcR (Figure 27B). Furthermore, levels also increased in MG1655Δ*rpoN* during chloramphenicol treatment, suggesting that it was not transcribed by σ^{54}-associated RNAP. In the absence of chloramphenicol, the *rtcR* null and *rpoN* deletion strains had similar levels of *rtcB* mRNA to wild-type as expected. qPCR analysis with *rtcA*-specific primers showed that *rtcA* mRNA levels also increased during chloramphenicol treatment in wild-type, *rtcR* null and *rpoN* deletion strains (data not shown).

To ascertain whether RtcB produced an observable growth phenotype during chloramphenicol-mediated inhibition of translation, recovery was monitored following 30 and 120 minutes of chloramphenicol treatment, in a similar manner to that described in Section 4.2.3. No difference in growth was observed between *rtcBA* deletion strains and wild-type strains, either during chloramphenicol treatment or recovery (data not shown). In addition, counting the colony forming units (CFU) of wild-type and *rtcB* deletion strains during five hours of chloramphenicol treatment indicated that cell viability was unaffected by RtcB in these conditions (data not shown).

4.2.4.2 MazF activates *rtcB* transcription

The effect of various toxin components of TA modules on *rtcB* transcription was also assessed. *E. coli* MG1655 carrying pSC3326, pKW82, pKP3035, pKP3065, pMCD8 or pBAD33 (empty plasmid control) were cultured exponentially and expression of toxins MazF, VapC, RelE, YafQ or ParE respectively were induced for 30 minutes upon addition of arabinose. Samples were analysed by qPCR with *rtcB*-specific primers. As shown in Figure 28, expression of MazF, a sequence-specific endoribonuclease that cleaves mRNA independently of the ribosome and thereby inhibits translation (Zhang et al., 2003), increased *rtcB* mRNA levels approximately five-fold relative to the plasmid control. Other mRNA interferases RelE and YafQ, which are both ribosome-dependent endoribonucleases (Pedersen et al.; Prysak et al., 2009), did not show similar increases in *rtcB* transcription. Expression of VapC, a site-specific endoribonuclease of *Shigella flexneri* 2a virulence plasmid pMYSH6000 that targets initiator tRNA^{fMet} (Winther and Gerdes, 2011), produced an approximate 3.5-fold increase in *rtcB* mRNA. Meanwhile, ParE, a toxin that targets DNA gyrase to inhibit DNA

replication (Jiang et al., 2002), did not change the levels of *rtcB* compared to the plasmid control.

The increase in *rtcB* transcription in the presence of MazF was confirmed by qPCR analysis using additional reference genes (Appendix B). Comparison of strains MG1655*rtcR*^{M1X} and MG1655Δ*rpoN* with MG1655 revealed that, like chloramphenicol treatment, MazF expression produced similar increases of around five-fold in *rtcB* mRNA levels in all strains relative to wild-type cells carrying empty pBAD33, regardless of the presence or absence of RtcR or σ^{54} (Figure 28B). This was also observed when samples were analysed with *rtcA*-specific primers (data not shown).

To determine whether RtcB had a functional role during or after MazF expression in *E. coli*, growth of strains MG1655 and MG1655Δ*rtcB* were compared. As expected, Figure 29A shows induction of *mazF* from pSC3326 rapidly reduced growth as a result of translation inhibition. However, there was no difference between the wild-type and *rtcB* deletion strains during MazF expression. Following 30 and 120 minutes of MazF-induced translation inhibition, cells were inoculated in fresh medium to remove arabinose and prevent further production of MazF, in order to monitor the recovery of growth in the two strains. Strains carrying empty pBAD33 and thus not experiencing inhibition of translation in the original cultures grew normally upon reseeding, while those exposed to MazF had a longer lag phase. Following 30 minutes of MazF expression, cells began to resume growth after approximately 90 minutes, with those with *rtcB* deleted displaying a slightly slower recovery compare to wild-type (Figure 29B). The delayed recovery of the *rtcB* deletion strain was more pronounced following 120 minutes of MazF expression, with cells resuming growth around 150 minutes after reseeding, compared to 120 minutes for the wild-type strain (Figure 29C). However, the extent of the difference between the wild-type and *rtcB* strain in recovery was not consistent, as indicated by the large error bars in Figure 29B and C. In the three experiments conducted, a clear difference was observed in one, while less defined in the other two.

To investigate whether RtcB mediated cell survival during growth inhibition as a result of MazF expression, cell viability was assessed. For this purpose, *E. coli* strain SC28, which carries a deletion of the *mazEF* TA locus, and an isogenic *rtcB* deletion strain were used to allow controlled expression of MazF from pSC3326 and antitoxin MazE from pSC228, as described by Pedersen et al. (2002). SC28 and SC28Δ*rtcB* carrying pSC228 and pSC3326 or pBAD33 were cultured exponentially and MazF expression induced with arabinose for a

period of five hours. Every hour, samples from each culture were plated on NA supplemented with glucose to prevent further MazF production and IPTG to induce the cognate antitoxin MazE and subsequently, the CFU counted. As previously demonstrated, MazE was able to resuscitate MazF-inhibited cells and allow colonies to form, while lack of MazE induction reduced the viable counts of MazF-inhibited cells (Figure 30). However, no difference was observed between wild-type and *rtcB* deletion strains, indicating that while RtcB may function in mediating recovery following MazF activity, it does not appear to influence cell survival during MazF inhibition of growth.

4.3 Discussion

In this chapter, we investigated the conditions under which *rtcBA* is transcribed in *E. coli* K-12. We started from the premise that, as *rtcBA* was non-essential for growth in standard laboratory conditions and transcription was low in the absence of induction with constitutively active RtcRΔN (this study; Genschik et al., 1998), *rtcBA* was likely to be activated in response to some type of cellular stress. Therefore, we used genome-wide screening approaches to search for genes that, when disrupted by transposon mutagenesis or expressed from a multi-copy plasmid, upregulated *rtcBA* transcription, as indicated by increased β-galactosidase activity from the chromosomal *rtcBA-lacZ* fusion.

Genetic screening led to the identification of Rof as a candidate activator of RtcR-dependent *rtcBA* transcription. A *rof*-carrying plasmid was found three times in the screen, indicating that adequate coverage of the genome was achieved, and independent expression of Rof from other plasmids was confirmed to induce *rtcBA* transcription. The RtcR-dependency of the phenotype was verified in both *rtcR* deletion and null strains and using an *rtcBA* fusion lacking the putative RtcR binding region *rtcO*. Further analysis of the phenotype revealed that *rtcBA* transcription increased concomitantly with a Rof-mediated reduction in growth rate at late exponential/early stationary phase, linking changes in *rtcBA* transcription to a growth phenotype associated with Rof expression in *E. coli*. However, as deletion of the *rtcBA* operon had no effect on growth, either during Rof expression or upon outgrowth in fresh media (not shown), its physiological function under these conditions remains unclear.

The only known function of Rof is to non-specifically inhibit the activity of the transcription terminator Rho by competitive inhibition of RNA binding (Gutierrrez et al., 2007; Pichoff et al., 1998). As a global regulator of transcription, Rho has roles in a variety of processes

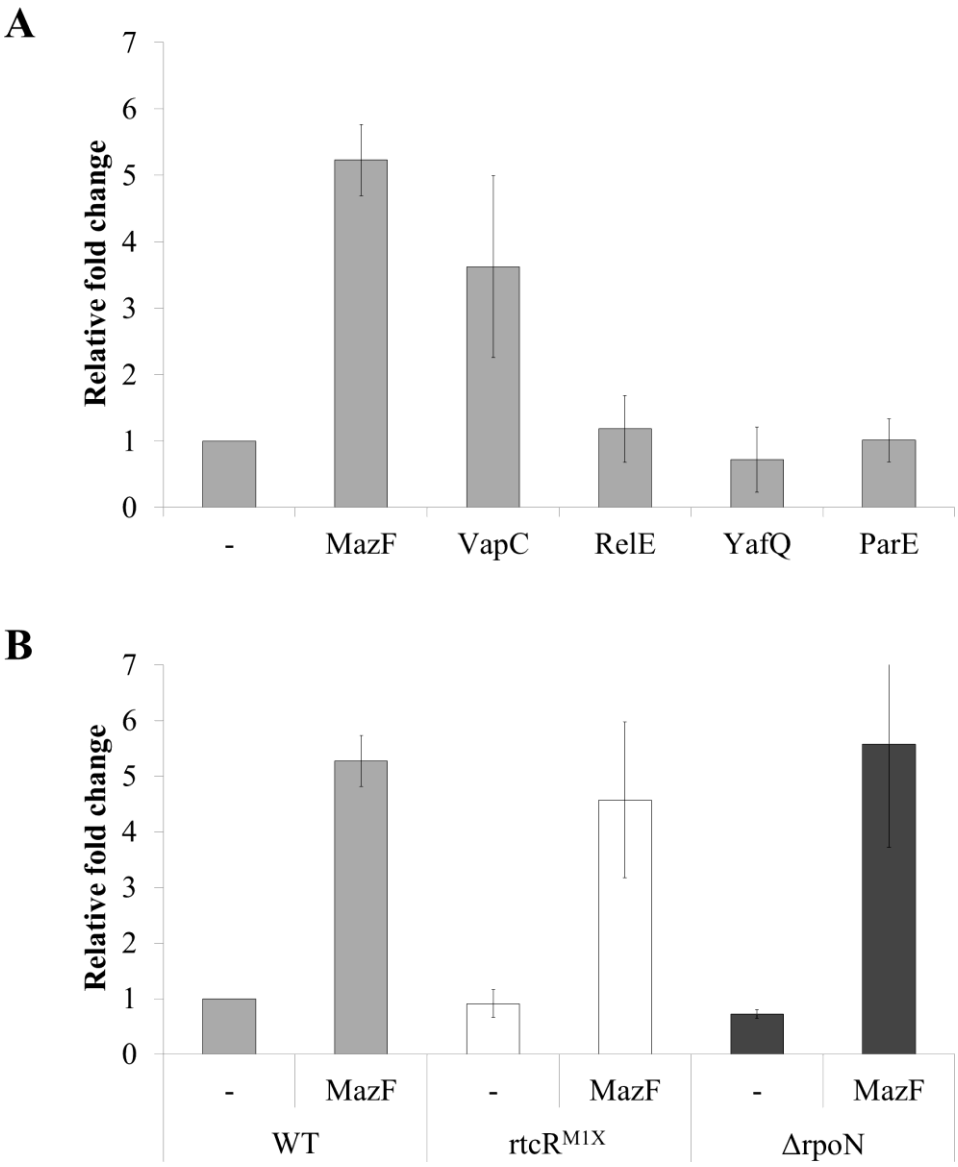


Figure 28 Levels of *rtcB* mRNA increase during MazF expression.

(A) *E. coli* MG1655 carrying pSC3326 (MazF), pKW82 (VapC), pKP3035 (RelE), pKP3065 (YafQ) and pMCD8 (ParE) and pBAD33 (-; corresponding empty plasmid) were grown exponentially in LB medium supplemented with 50 µg/ml chloramphenicol and toxin expression was induced for 30 minutes with 0.2% arabinose. RNA was extracted from each sample, reverse transcribed and the fold change in transcript levels of *rtcB* in the strains expressing toxins relative to the strain carrying the empty plasmid were determined by qPCR. qPCR data were analysed using the Pfaffl method and normalised against the *tatA* gene. Data from three independent experiments are shown as the mean relative fold change ± standard deviation. (B) Exponentially growing MG1655 (WT), MG1655*rtcR*^{M1X} (rtcR^{M1X}) and MG1655Δ*rpoN* (ΔrpoN) carrying pSC3326 were treated with 0.2% arabinose to induce MazF expression for 30 minutes and *rtcB* transcription was analysed by qPCR as described for (A).

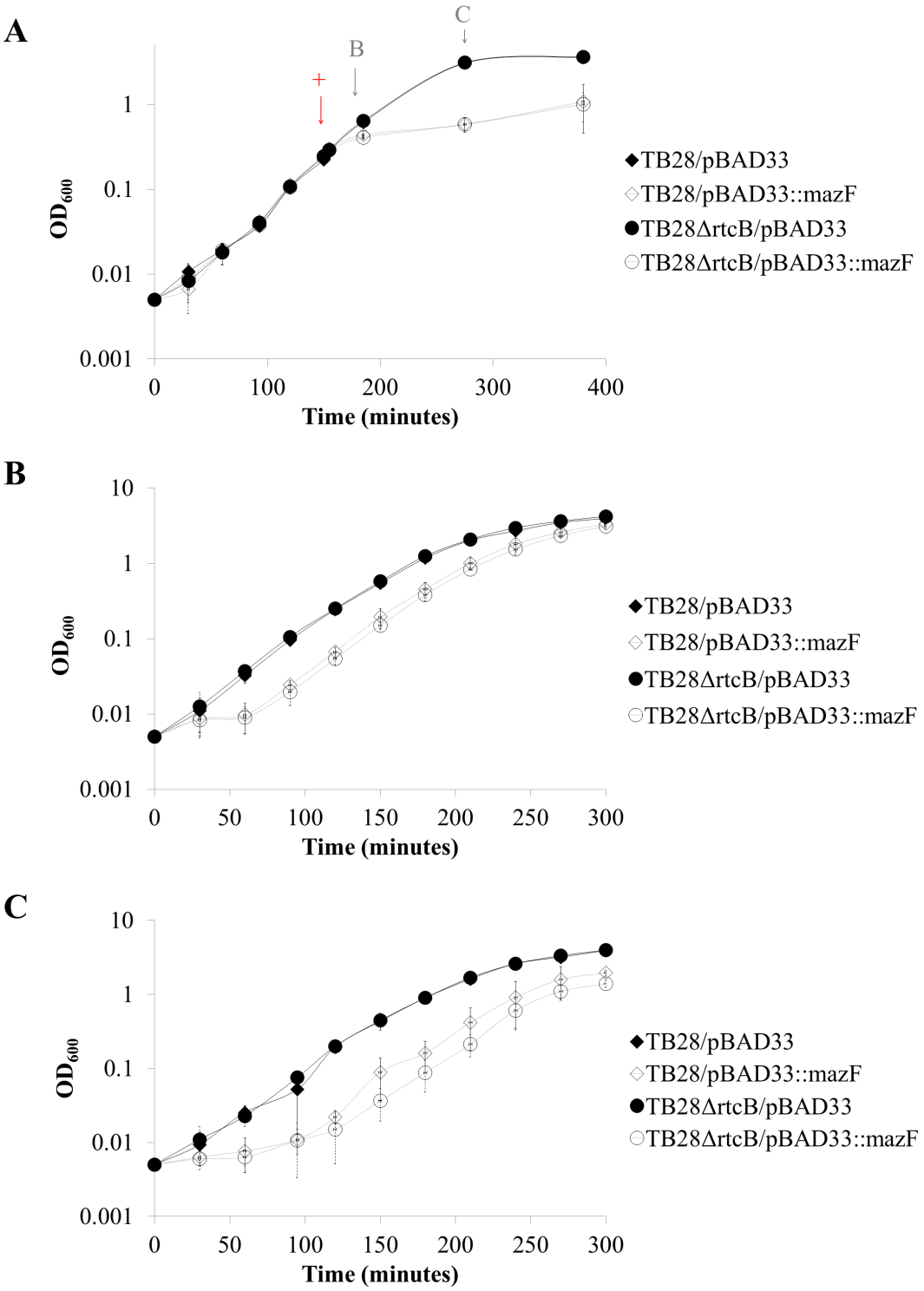


Figure 29 RtcB mediates recovery of growth following MazF expression.

(A) *E. coli* MG1655 and MG1655Δ*rtcB* carrying pSC3326 (MazF) or pBAD33 (corresponding empty plasmid) were cultured exponentially in LB medium supplemented with 50 µg/ml chloramphenicol and 0.2% arabinose was added to induce MazF expression as indicated (+, red arrow). OD_{600} was measured at regular intervals. After 30 mins (B) and 120 mins (C) of MazF expression as indicated (grey arrows), cells from each culture were washed to remove the inductive agent and used to inoculate fresh LB medium containing 50 µg/ml chloramphenicol and 0.2% glucose to repress MazF expression from pSC3326 and recovery of growth was monitored. The values are an average of three independent experiments.

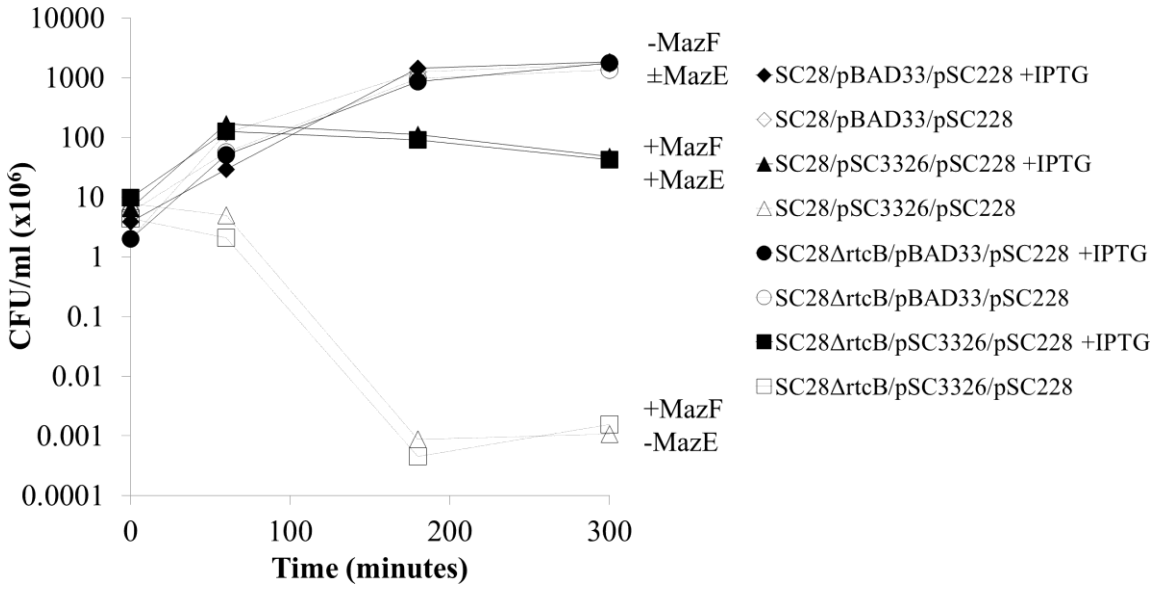


Figure 30 Deletion of *rtcB* has no effect on cell viability during MazF expression.

E. coli SC28 (MC1000Δ*mazEF*) and SC28Δ*rtcB* carrying pSC228 (pNDM220::*mazE*) and pSC3326 (pBAD33::*mazF*) or pBAD33 were cultured exponentially in LB medium containing 30 µg/ml ampicillin and 50 µg/ml chloramphenicol. At 0 minutes 0.2% arabinose was added to induce MazF expression. Samples were taken from each culture at time 0, 60, 180 and 300 minutes and spread on NA plates supplemented with 30 µg/ml ampicillin, 50 µg/ml chloramphenicol and 0.2% glucose to repress MazF expression, with (closed symbols) and without (open symbols) 1 mM IPTG to induce MazE expression.

including formation of transcript 3′-ends, suppression of antisense transcription and preventing the formation of RNA/DNA hybrids (Grylak-Mielnicka et al., 2016). Rho mediates transcription termination by binding to regions of transcribing RNA that are C-rich, largely unstructured and around 80 nucleotides long and then hydrolyses ATP to stimulate translocation along RNA, causing the dissociation of RNAP from RNA and DNA (Peters et al., 2011). Therefore, the increase in *rtcBA* levels observed upon expression of Rof could be a result of decreased premature transcription termination of *rtcBA* when the activity of Rho is inhibited. Alternatively, it could imply that RtcB and/or RtcA have functional roles in conditions where Rho activity is disrupted. Preliminary data (not shown) suggests that deletion of *rof* had no effect on the growth of *E. coli* in standard conditions (supported by (Pichoff et al., 1998)) and does not affect levels of β-galactosidase activity from the chromosomal *rtcBA-lacZ* fusion, indicating that investigation into the regulation of *rof* expression may be prudent in order to further characterise a role for RtcB or RtcA during Rof expression. Furthermore, treatment of cells with bicyclomycin, which also inhibits Rho but via a different mechanism (Gutierrrez et al., 2007), could establish whether the observed increase in *rtcBA* transcription is specific to Rof-mediated Rho inhibition or a general effect of dysregulated transcription termination.

A number of other conditions were found to increase the levels of *rtcBA* by qPCR analysis, all of which have the common function of targeting translation, either at the level of mRNA, tRNA or rRNA. These include several antibiotics, namely chloramphenicol, tetracycline and kanamycin and the endoribonucleases MazF and VapC. This finding corroborates a recent study by Engl et al. (2016), who also observed *rtcBA* induction upon treatment with a variety of agents that target the translation apparatus, which included chloramphenicol, tetracycline and VapC. Interestingly, they were not able to detect healing of VapC-induced breaks in $tRNA^{fMet}$ by RtcB, leading to the proposal that RtcB may have other functions possibly linked to secondary effects of damaged tRNA on the process of translation. We did observe some specificity of the activation of *rtcBA* when targeting translation, as toxins RelE and YafQ were not able to induce *rtcBA* transcription. Therefore, it appears that *rtcBA* responds to a subset of challenges to the translation apparatus.

We further investigated this idea with chloramphenicol and MazF, which induced the greatest increases in *rtcBA* transcription of approximately 6.5-fold and 5-fold respectively. During MazF expression, the growth rate and viability were found to be the same in wild-type and RtcB deficient strains. A small difference was observed in the recovery of growth following MazF exposure, with cells lacking *rtcB* lagging slightly compared to wild-type. This is in

contrast to a recent publication from Temmel et al. (2016), who also reported a 4- to 5-fold increase in *rtcBA* levels but observed a much clearer difference between wild-type and *rtcB* deletion strains in the recovery following MazF expression. The reason for this is unclear, but could be due to technical differences in the experiment, such as the plasmid used to express MazF. When treated with chloramphenicol, no difference was observed between wild-type and *rtcBA* deletion strains in growth rate, recovery or viability (not shown). Therefore, the role of RtcB during chloramphenicol treatment and MazF expression remains unclear based on these analyses. However, since MazF cleavage is known to produce 2′,3′-phosphate and 5′-hydroxyl terminals suitable for RtcB ligation (Zhang et al., 2005), we hypothesised that the induction of *rtcB* during MazF expression may be a stress response to mediate RNA repair. This is explored further in Chapter 5.

Surprisingly, the effects of both chloramphenicol and MazF on *rtcBA* transcription were found to be independent of RtcR. To eliminate the possibility that cross-activation from another bEBP in the absence of RtcR was responsible, levels of *rtcBA* were also assessed in an *rpoN* deletion strain. This confirmed that transcription was also independent of σ^{54}, indicating that transcription of *rtcBA* in these conditions was likely to be mediated via another promoter regulated by a member of the σ^{70} family. In order to identify this additional promoter, primer extension could be used to map the 5′ end of the *rtcBA* transcript in the presence of MazF and chloramphenicol. As several *rtcBA* activators identified by genetic screening were also found to be RtcR-independent, some of these may represent potentially interesting candidates for further study.

In summary, using a combination of genetic screening and targeted analysis of various antibiotics and toxins by qPCR, we identified several conditions whereby *rtcBA* transcription is activated in *E. coli*. Expression of Rof, a Rho inhibitor, led to an increase in *rtcBA* transcription, raising the possibility that RtcB functions in situations where Rho-dependent transcription termination is disrupted. This could reasonably involve the RNA ligase activity of RtcB or the lesser studied DNA ligase or DNA capping activities. Moreover, *rtcBA* seems to respond to the translational status of the cell, with challenges at multiple levels of translation resulting in induction of *rtcBA*. These include targeting 23S rRNA (chloramphenicol), 16S rRNA (MazF, tetracycline and kanamycin), initiator $tRNA^{fMet}$ (VapC) and mRNA (MazF). The RNA targets of these agents imply that a functional role for RtcB in these conditions may more likely involve its RNA ligase activity.

Chapter 5 Function of RtcB in *E. coli*

5.1 Introduction

RtcB is a widely conserved protein with homologues in all domains of life (Popow et al., 2012). Biochemical studies of RtcB using a range of RNA substrates revealed a unique GTP- and Mn^{2+}-dependent 3′-5′ RNA ligase activity (Chakravarty and Shuman, 2012; Chakravarty et al., 2012; Tanaka et al., 2011a). RtcB demonstrated specificity for 2′,3′-cyclic phosphate or 3′-phosphate RNA terminals and 5′-hydroxyl terminals, ligating to form a 3′,5′-phosphodiester bond.

Complementation of the normally lethal deletion of Trl1 RNAL in *S. cerevisiae* by recombinant RtcB from *E. coli* provided an indication of the function *in vivo* (Tanaka et al., 2011b). RtcB was able to perform all essential RNA ligations in this organism including sealing of tRNA halves and *XBP-1* mRNA of the UPR following intron excision and protected against site-specific endoribonuclease activity of fungal ribotoxin PaT. Accordingly, the human homologue of RtcB is known to be essential for tRNA maturation (Popow et al., 2011). Moreover, RtcB is required for splicing of *XBP-1* mRNA in metazoans (Jurkin et al., 2014; Kosmaczewski et al., 2014; Lu et al., 2014). An inhibitory role for RtcB in axon regeneration has also been described that is independent of its functions in tRNA maturation and the UPR (Kosmaczewski et al., 2015). This suggests that RtcB can perform RNA ligations in a range of contexts in multicellular organisms.

However, the role of RtcB in bacteria is unclear as the majority of bacterial RNA transcripts do not require splicing or are self-splicing (Reinholdhurek and Shub, 1992) and the UPR is a phenomenon of eukaryotes (Hollien, 2013). The expression of *E. coli* RtcB from a σ^{54}-dependent promoter combined with the fact that it is non-essential for growth of *E. coli* in standard laboratory conditions, as described in Chapters 3 and 4, suggests it may be involved in repair of stress-mediated RNA damage. Consistent with this idea, Temmel et al. (2016) demonstrated that RtcB could repair MazF-cleaved 16S rRNA in *E. coli*. Further, roles for RtcB in rRNA stability and cell motility were recently described by Engl et al. (2016). These reports indicate that the RNA ligase activity of RtcB influences a variety of pathways in bacteria, as is the case in metazoans.

In the following chapter, potential RNA substrates of RtcB were identified by *in vivo* crosslinking in relevant conditions and high throughput sequencing, with the aim of further elucidating the spectrum of biological functions of RtcB in *E. coli*. A role for RtcB in maintaining rRNA stability during amino acid starvation was also considered.

5.2 Results

5.2.1 CRAC of RtcB

5.2.1.1 Experimental procedure and high throughput sequencing

Cellular RNA substrates of RtcB in *E. coli* were investigated using a technique known as crosslinking and analysis of cDNA (CRAC). CRAC detects both stable and relatively transient interactions between RNA and a tagged protein of interest and involves UV-induced protein-RNA crosslinking *in vivo*, followed by purification of the protein-RNA complexes and preparation and sequencing of a cDNA library, as shown in Figure 31A (Granneman et al., 2009; Tree et al., 2014; Winther et al., 2016).

To implement CRAC, the *rtcB* gene was cloned into plasmid pKW220 downstream of the LacI-regulated promoter and in-frame with an N-terminal dual affinity tag consisting of three FLAG tags, a TEV protease cleavage site and a His_6 tag (FTH), which enables stringent protein purification. Resulting plasmid pKW220rtcB expressed FTH-tagged RtcB upon IPTG induction. As treatment with chloramphenicol and expression of MazF were found to increase transcription of *rtcB* (Chapter 4), these conditions were selected to perform CRAC of RtcB to allow identification of potentially biologically relevant targets. Therefore, four strains were used for CRAC: *E. coli* MG1655Δ*rtcB* carrying pKW220rtcB (RtcB); MG1655Δ*rtcB* carrying pKW220rtcB and treated with chloramphenicol (+Cml); MG1655Δ*rtcB* carrying pKW220rtcB and pSC3326 for expression of MazF (+MazF) and MG1655Δ*rtcB* carrying pKW220 (control). Bacteria were cultured exponentially and FTH-tagged RtcB expression was induced for 10 minutes with IPTG, followed by 30 minutes of treatment with chloramphenicol for the +Cml strain or 30 minutes of arabinose-induction of MazF expression for the +MazF strain. Expression of FTH-RtcB was verified in each culture by Western blot using anti-His antibodies. As shown in Figure 31B, RtcB was present in all samples carrying pKW220rtcB and absent in the sample carrying the corresponding empty plasmid.

Bacterial cultures were UV-irradiated, triggering covalent crosslinking between interacting RNAs and FTH-tagged RtcB and the resulting RtcB-RNA complexes were harvested from cell lysates using the FTH tag. RtcB-bound RNAs were trimmed, radioactively labelled and ligated with linkers at the 5′- and 3′-ends. The 5′ linker contained a barcode unique to each of the four samples, to allow multiplex sequencing. RtcB-RNA complexes were separated by SDS-PAGE and visualized by autoradiography. As RtcB was crosslinked to RNAs of varying lengths, the complexes appeared as smears on the autoradiogram, which were size-selected according to the molecular weight of RtcB (45.2 kDa), as indicated in Figure 31C. Following protease digestion of RtcB, the RNAs were reverse transcribed to generate cDNA, amplified using linker-specific primers and resolved by agarose gel electrophoresis (Figure 31D). Fragments between 150 and 300 bp were selected from the gel, as this range was above non-specific amplification products present in the control PCR reaction and purified to generate a cDNA library.

A sample from each of the four cDNA libraries was cloned into plasmid pCR4-TOPO and five clones from each were sent for sequencing with plasmid-specific primers. All four libraries were found to have the correct barcoded 5′ linker and identical 3′ linkers, suggesting that linker ligation steps of CRAC were successful. The size of the fragment after accounting for the linkers and the primers used to amplify the cDNA ranged from 15 to 84 bp with an average of 41.1 bp, which was deemed appropriate (data not shown).

Consequently, the cDNA libraries were pooled and sent for deep sequencing. The raw data was sorted and analysed by Dr Kristoffer Winther (University of Copenhagen) using the PyCRAC software (Webb et al., 2014). First, the reads were trimmed to remove the linkers and any low quality reads or reads without linkers were removed. Next, reads were sorted according to their barcode, separating the RtcB, +Cml, +MazF and control sample reads. The remaining reads were collapsed, removing copies of reads with identical barcode, sequence and length. Finally, reads were normalised to number of reads per million and mapped to the *E. coli* genome.

5.2.1.2 Identification of candidate RNA targets of RtcB

The sorted data from Dr Kristoffer Winther enabled visual comparison of samples. RNAs enriched in the control sample were used to assess the background of the experiment and were considered to arise from RNA binding to the FTH tag expressed from pKW220 or retained on

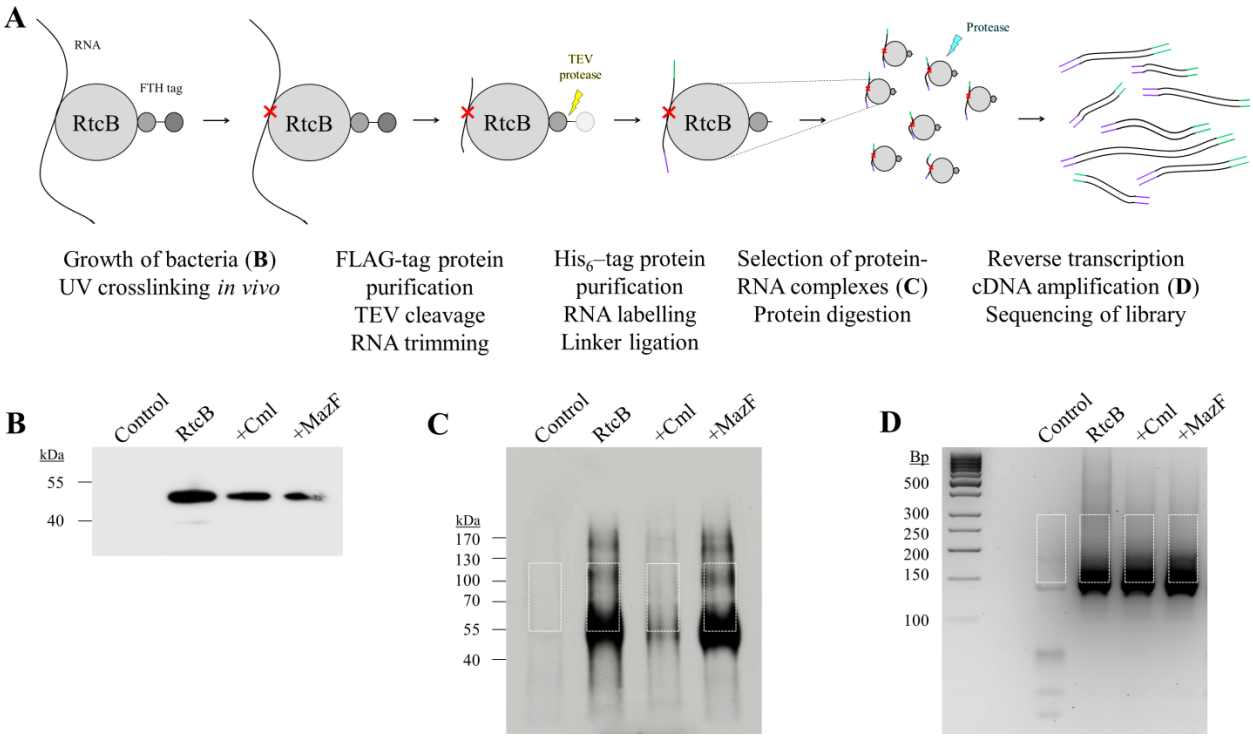


Figure 31 CRAC of RtcB in *E. coli*.

(A) Schematic illustration of CRAC of RtcB. The four bacterial cultures described in the text were exposed to 1800 mJ of UV light *in vivo*, crosslinking RNA in close contact with FLAG-TEV-His_6 (FTH)-tagged RtcB. RtcB-RNA complexes were purified from bacterial lysates using anti-FLAG antibodies, followed by TEV protease cleavage to remove the complexes from the anti-FLAG resin. RNA was trimmed with an RNase cocktail, radioactively labelled and barcoded linkers were ligated. Further His_6-tag purification on nickel beads was conducted and RtcB-RNA complexes were resolved by SDS-PAGE then transferred to a nitrocellulose membrane and visualised by autoradiography. The smear, which comprises RtcB complexed with RNAs of differing length, were cut from the membrane and protease-treated, leaving RNA. cDNA was synthesised by reverse transcription and amplified by PCR, using linker-specific primers. The resulting cDNA library was deep sequenced using the Illumina platform. (B) Western blot analysis of FTH-RtcB expression. Samples were collected following induction of RtcB and treatment with chloramphenicol or MazF expression, resolved by SDS–PAGE, transferred to a PVDF membrane and RtcB was detected with anti-His antibodies. (C) Autoradiogram of purified RtcB-RNA complexes. Regions indicated by white dashed boxes were cut from the membrane and RNA isolated. (D) Agarose gel electrophoresis of amplified cDNA. Regions indicated by white dashed boxes were cut from the gel and purified, creating cDNA libraries. 'RtcB' is MG1655Δ*rtcB*/pKW220rtcB, '+Cml' is MG1655Δ*rtcB*/pKW220rtcB treated with chloramphenicol, '+MazF' is MG1655Δ*rtcB*/pKW220rtcB/pSC3326 and 'Control' is MG1655Δ*rtcB*/pKW220.

the column used during the CRAC protocol. Generally, these represented the most abundant RNAs in the cell, including rRNA, some tRNAs and mRNA encoding highly expressed proteins. Visual inspection of the peak profiles of the RtcB, +Cml and +MazF samples compared with the control identified significant enrichment of 93 RNAs in samples expressing RtcB. Figure 32A illustrates that the majority of these increased in all three samples, while twelve were specific to the +Cml sample, +MazF sample or both and therefore represented targets of potential biological relevance.

To gain an overall impression of the RNAs interacting with RtcB, the RNA types and functional groupings in the +Cml and/or +MazF only samples were compared to the whole RtcB data set. As Figure 32B shows, the 93 RNAs found to be enriched in any RtcB-containing sample were largely tRNAs, followed by mRNAs and rRNAs. Accordingly, over 75% of these RNAs were involved in translation, encoding tRNA, rRNA, ribosomal proteins or global regulators of translation (Figure 32C). In contrast, the majority of +Cml- and/or +MazF-specific RNAs were mRNAs, while the remainder were small (s)RNAs (Figure 32B). Those involved in translation (59%) represented a smaller proportion compared with the total RtcB data set (Figure 32C). A quarter had functional roles in metabolism and the rest were involved in transport or stress response.

Rather than the non-coding RNA species that dominated the data set as a whole, RNAs involved in translation that were increased in the +Cml and/or +MazF samples only were all mRNAs encoding 50S or 30S ribosomal subunit proteins. These accounted for seven of the twelve +Cml- and/or +MazF-specific RNAs interacting with RtcB. The peak profiles of sequencing reads from RtcB CRAC revealed enrichment of *rplL*, *rplD* and *rplM* mRNAs, which encode L12, L4 and L13 proteins of the 50S subunit respectively, in both +Cml and +MazF samples (Figure 33A-C). Enrichment of *rpmA* mRNA, encoding 50S protein L27, was most significant in the +MazF sample, while *rplU* mRNA, encoding 50S protein L21, was most enriched in the +Cml sample (Figure 33D and E). CRAC of RtcB also revealed enrichment of *rpsB* and *rpsJ* mRNA, which encode 30S ribosomal subunit proteins S2 and S10 (also known as NusE) respectively, in the +Cml sample only (Figure 34A and B).

The peak profiles for the remaining five RNAs found in +Cml and/or +MazF samples only are shown in Figure 35 and include two sRNAs, *gcvB* and *glmY*, which have regulatory functions in expression of amino acid transporters and an enzyme involved in hexosamine biosynthesis respectively. Enrichment of *gcvB* sRNA by RtcB was observed in both +Cml and +MazF samples, whereas *glmY* sRNA was specific to +MazF. Further, RtcB enriched

ilvL mRNA, encoding the leader peptide of the *ilv* operon, *rraA* mRNA, encoding an inhibitor of RNase E, and *mazF* mRNA in the +MazF sample only. The results of CRAC of RtcB are summarised in Table 7.

Table 7 Summary of RNAs identified by CRAC of RtcB during chloramphenicol treatment or MazF expression or both.

Gene	RNA type	Functional classification	Function	Sample(s)
rplL	mRNA	Translation	50S ribosomal protein L7/12	+Cml +MazF
rplD	mRNA	Translation	50S ribosomal protein L4	+Cml +MazF
rplM	mRNA	Translation	50S ribosomal protein L13	+Cml +MazF
rpmA	mRNA	Translation	50S ribosomal protein L27	+MazF (+Cml)
rplU	mRNA	Translation	50S ribosomal protein L21	+Cml (+MazF)
rpsB	mRNA	Translation	30S ribosomal protein S2	+Cml
rpsJ	mRNA	Translation	30S ribosomal protein S10 or ribosome-independent transcription termination regulator NusE	+Cml
gcvB	sRNA	Transport	Regulates expression of amino acid transporters DppA, CycA SstT and others	+Cml +MazF
glmY	sRNA	Metabolism	Regulates expression of glucosamine-6-phosphate synthase involved in hexosamine biosynthesis	+MazF
ilvL	mRNA	Metabolism	Leader peptide of *ilv* operon, encoding enzymes for valine and isoleucine biosynthesis	+MazF
rraA	mRNA	Metabolism	Inhibitor of RNase E	+MazF
mazF	mRNA	Stress response	Ribosome-independent sequence-specific endoribonuclease	+MazF

(The 'RNA type', 'Functional classification' and 'function' was obtained from the EcoCyc database (Keseler et al., 2011).)

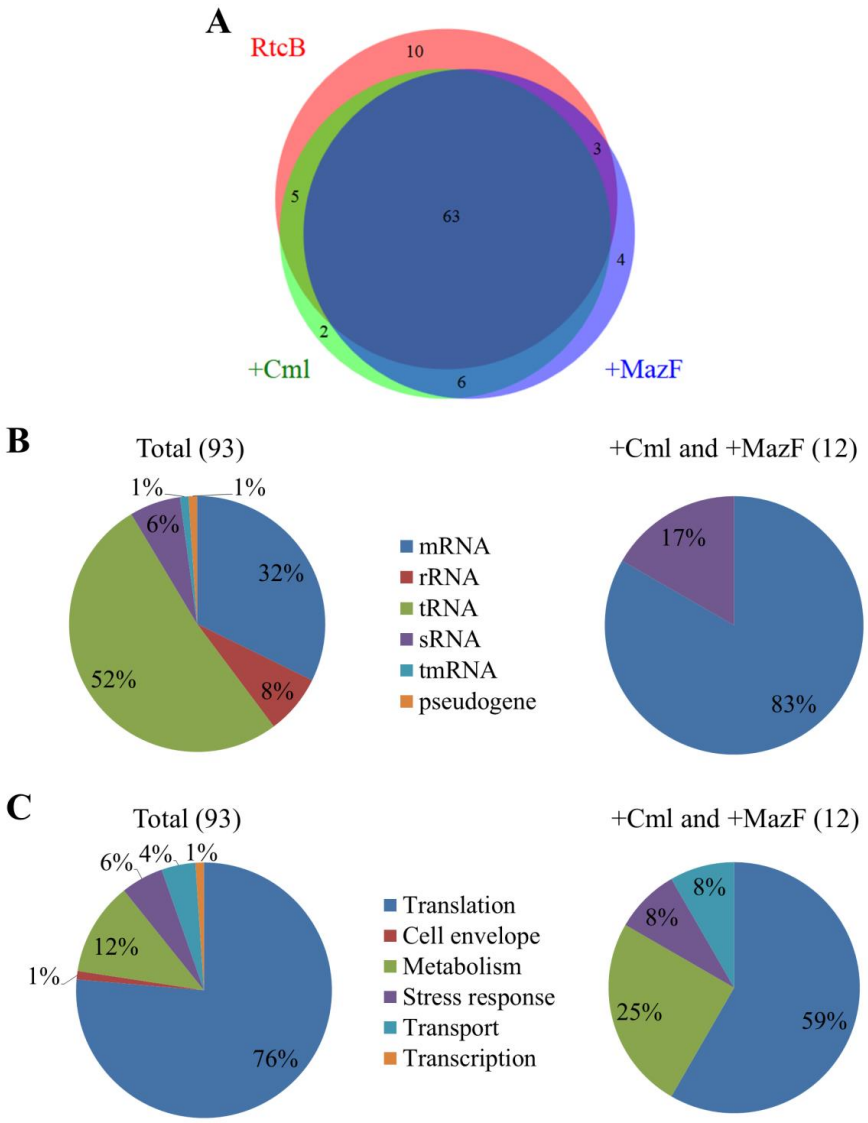


Figure 32 Overview of interactions between RtcB and RNAs identified by CRAC.

(A) Venn diagram showing the number of RNAs significantly enriched in samples from *E. coli* MG1655Δ*rtcB* carrying pKW220rtcB (RtcB), MG1655Δ*rtcB* carrying pKW220rtcB and treated with chloramphenicol (+Cml) and MG1655Δ*rtcB* carrying pKW220rtcB and pSC3326 (+MazF) compared to MG1655Δ*rtcB* carrying pKW220 (corresponding empty plasmid). (B) Pie charts illustrating the type of RNAs enriched in the total RtcB data set (total) and in those specific to chloramphenicol- and/or MazF-treated samples (+Cml and +MazF). (C) Pie charts illustrating the functional classification of RNAs enriched in the total RtcB data set (total) and in those specific to chloramphenicol- and/or MazF-treated samples (+Cml and +MazF). RNA types and functional classifications were obtained from the EcoCyc database (Keseler et al., 2011).

As described in Chapter 4, MazF specifically cleaves RNA at single-stranded ACA sites (Zhang et al., 2003). The position of ACA sites in the RNAs found to crosslink RtcB in +Cml and/or +MazF samples only were mapped and are indicated on the X-axis of the profiles of sequencing reads by red arrows (Figure 33 to Figure 35).

5.2.1.3 RNAs identified by CRAC are unaffected by deletion of *rtcB*

To investigate whether RtcB functions in the repair of crosslinked RNAs specifically identified in the +Cml and/or +MazF samples, the relative levels of each RNA were compared by qPCR analysis in *E. coli* strains MG1655 and MG1655Δ*rtcB* following 30 minutes of chloramphenicol treatment or MazF expression as appropriate.

Figure 36A shows the levels of *rtcB, rtcA* and eight RNAs found to be enriched by RtcB in the +Cml CRAC sample, *rplL, rplD, rplM, rpmA, rplU, rpsB, rpsJ* and *gcvB,* after chloramphenicol treatment. In accordance with data shown in Chapter 4, levels of both *rtcB* and *rtcA* mRNA increased by approximately six-fold and as expected, *rtcB* transcripts were only detectable in MG1655 and not in the isogenic *rtcB* deletion strain, while *rtcA* was unaffected by deletion of *rtcB*. The levels of *rplL, rplD, rplM, rpmA, rplU, rpsB, rpsJ* and *gcvB* RNA also increased in the chloramphenicol-treated samples compared to untreated. Moreover, there was no difference between the wild-type and *rtcB* deletion strains in treated and untreated conditions respectively, indicating that the presence or absence of RtcB had no effect on the stability of the RNAs during chloramphenicol treatment.

qPCR analyses during MazF expression of *rtcB, rtcA* and RNAs enriched in the +MazF CRAC sample, including *rplL, rplD, rplM, rpmA, rplU, gcvB, glmY, ilvL* and *rraA,* are shown in Figure 36B. Consistent with previous results, MazF expression produced a five-fold increase in *rtcB* transcript levels in wild-type cells. Similar levels of *rtcA* were detectable in wild-type and *rtcB* deletion strains. With the exception of *rplD* and *rplM* mRNA, MazF expression resulted in elevated levels of each RNA, suggesting they were not targets of MazF degradation. Levels of *rplD* and *rplM* transcripts decreased in the presence of MazF, indicating that these mRNAs may be cleaved by MazF. However, as observed for the RNAs tested during chloramphenicol treatment, in all cases the RNA levels were unchanged in the *rtcB* deletion strain during MazF expression, which suggests that RtcB is not involved in maintaining the stability of these RNAs.

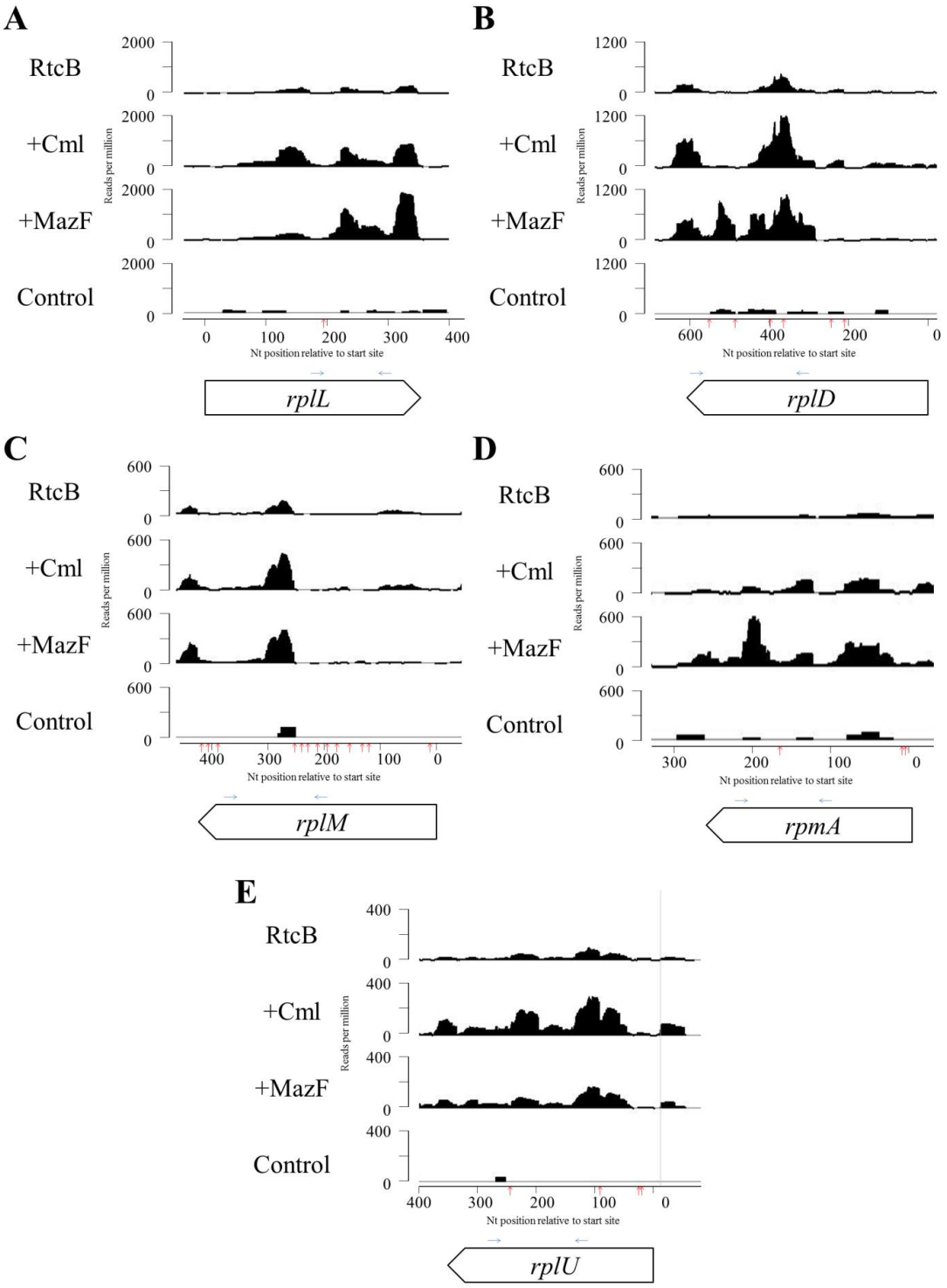


Figure 33 Interactions between RtcB and RNAs encoding 50S ribosome proteins in *E. coli*.

Enrichment of *rplL* (A), *rplD* (B), *rplM* (C), *rplU* (D) and *rpmA* (E) RNA by CRAC analysis of RtcB. Profiles are shown for RNA obtained from: *E. coli* MG1655Δ*rtcB* carrying pKW220rtcB (RtcB); the same strain treated with chloramphenicol (+Cml); *E. coli* MG1655 Δ*rtcB* carrying pKW220rtcB and pSC3326 (+MazF) and *E. coli* MG1655 Δ*rtcB* carrying pKW220 (control). Note the changes in scale of the X- and Y-axes for parts (A) to (E). The location of potential MazF cleavage sites are indicated on the X-axis (red arrow). Primer annealing sites used for qPCR analysis (Figure 36) are indicated where appropriate (blue arrows).

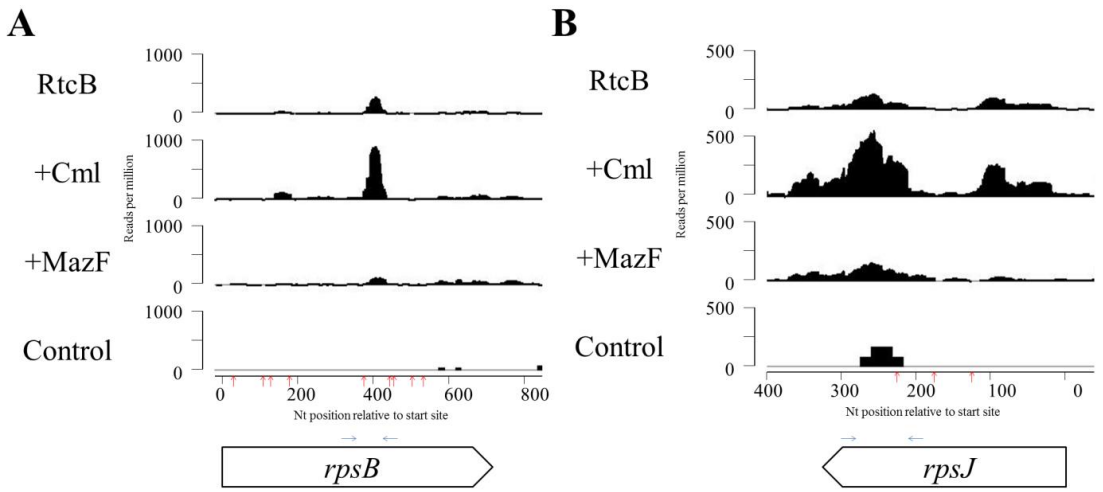


Figure 34 Interactions between RtcB and RNA encoding 30S ribosome proteins in *E. coli*.

Enrichment of *rpsB* (A) and *rpsJ* (B) RNA by CRAC analysis of RtcB. Profiles are shown for *E. coli* MG1655Δ*rtcB* carrying pKW220rtcB (RtcB), the same strain treated with chloramphenicol (+Cml), *E. coli* MG1655 Δ*rtcB* carrying pKW220rtcB and pSC3326 (+MazF) and *E. coli* MG1655 Δ*rtcB* carrying pKW220 (control). Note the changes in scale of the X- and Y-axes for parts (A) and (B). The location of potential MazF cleavage sites are indicated on the X-axis (red arrow). Primer annealing sites used for qPCR analysis (Figure 36) are indicated where appropriate (blue arrows).

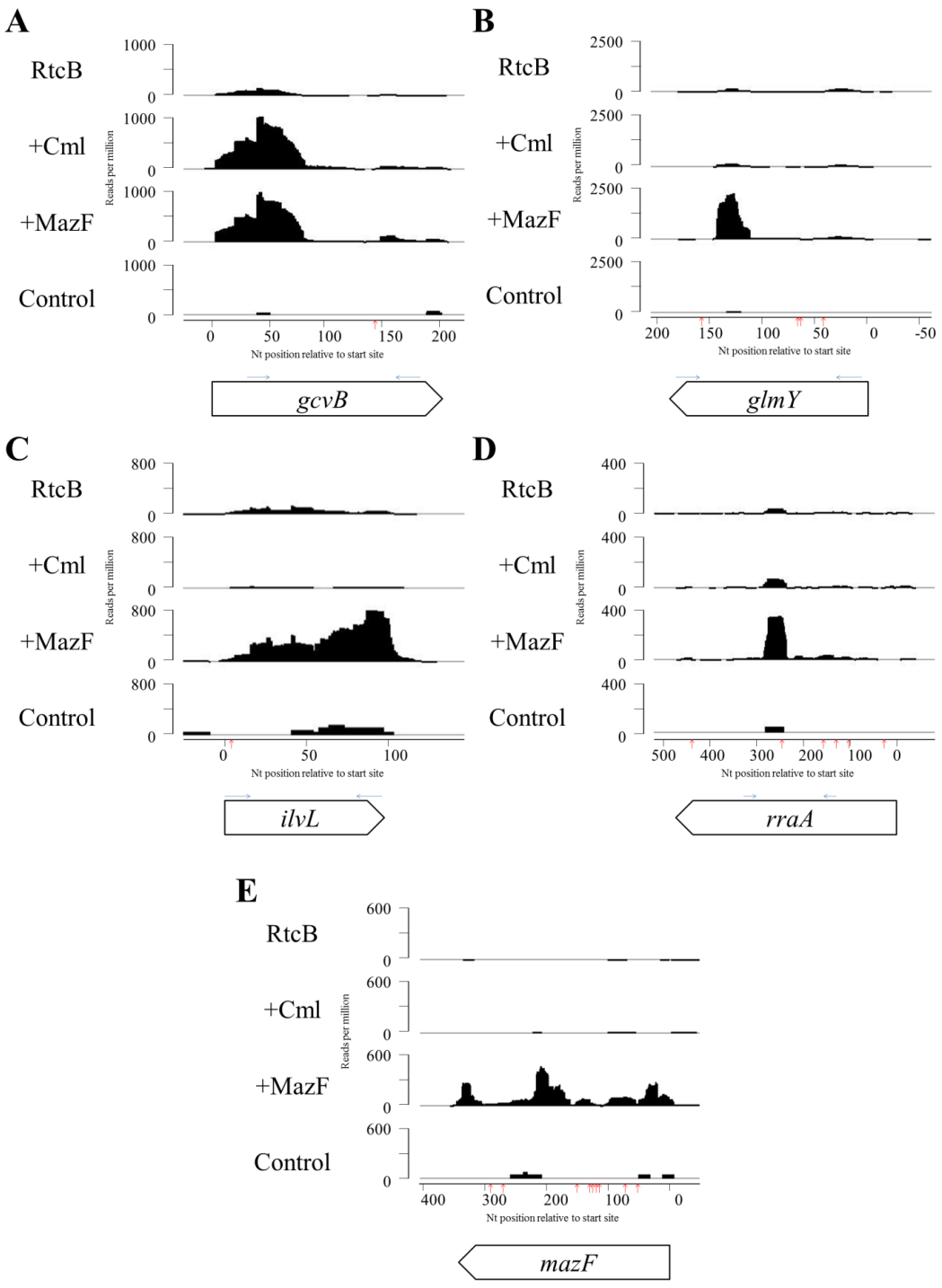


Figure 35 Other interactions between RtcB and RNA in *E. coli*.

Enrichment of *gcvB* (A) and *glmY* (B) regulatory sRNAs and *ilvL* (C), *rraA* (D) and *mazF* (E) mRNAs by CRAC analysis of RtcB. Profiles are shown for *E. coli* MG1655Δ*rtcB* carrying pKW220rtcB (RtcB), the same strain treated with chloramphenicol (+Cml), *E. coli* MG1655 Δ*rtcB* carrying pKW220rtcB and pSC3326 (+MazF) and *E. coli* MG1655 Δ*rtcB* carrying pKW220 (control). Note the changes in scale of the X- and Y-axes for parts (A) to (E). The location of potential MazF cleavage sites are indicated on the X-axis (red arrow). Primer annealing sites used for qPCR analysis (Figure 36) are indicated where appropriate (blue arrows).

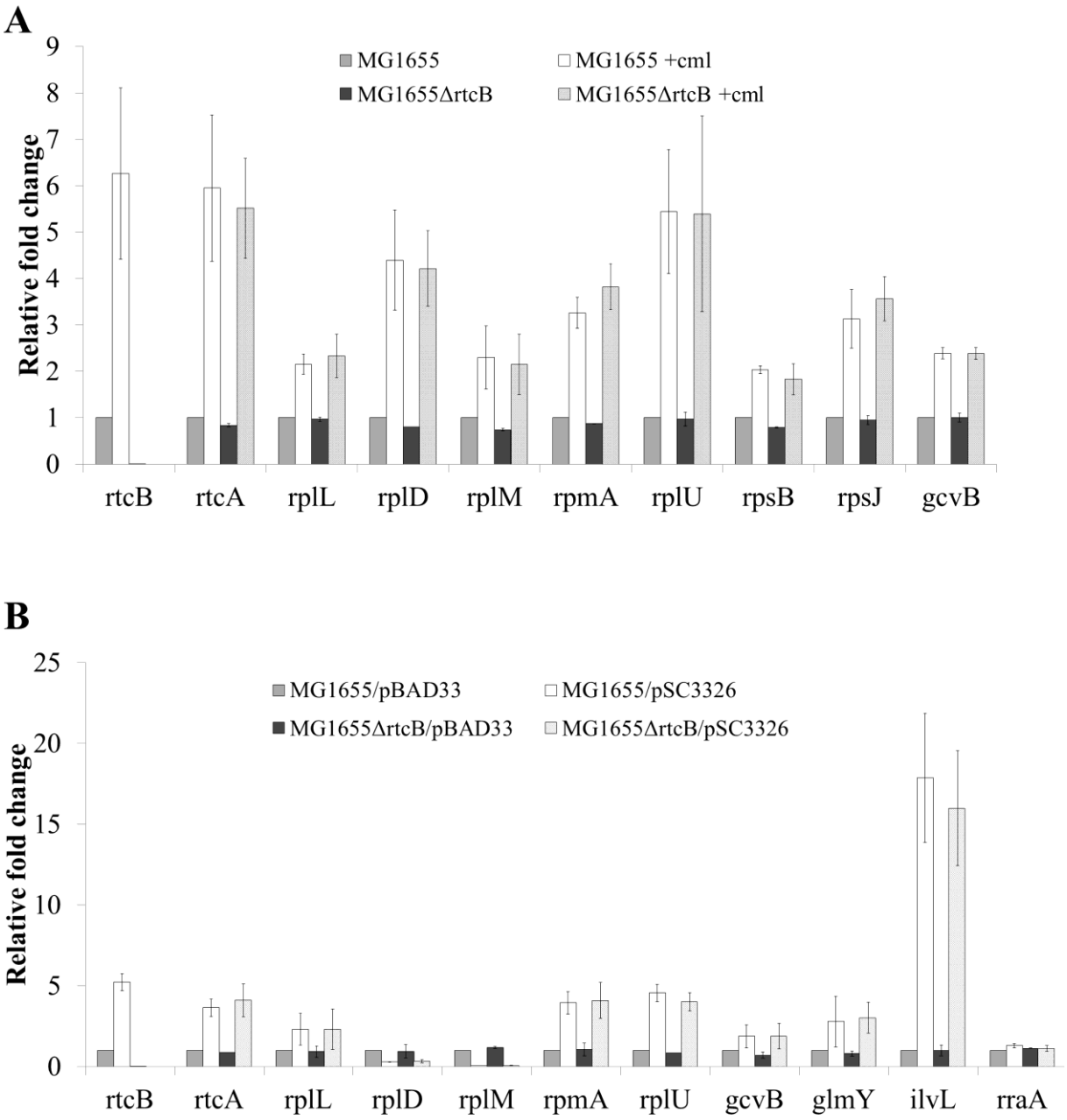


Figure 36 Stability of RNAs crosslinked to RtcB are unchanged upon deletion of *rtcB*.

(A) *E. coli* strains MG1655 and MG1655Δ*rtcB* were grown exponentially in LB medium then treated with 50 µg/ml chloramphenicol (cml) for 30 minutes. RNA was extracted from each sample, reverse transcribed and the fold change in transcript levels of *rtcB*, *rtcA* and eight genes found to be enriched by CRAC of RtcB during chloramphenicol treatment (*rplL, rplD, rplM, rpmA, rplU, rpsB, rpsJ* and *gcvB*) were determined by qPCR, relative to untreated MG1655. qPCR data were analysed using the Pfaffl method and normalised against the *tatA* gene. Data from two independent experiments are shown as the mean relative fold change ± standard deviation. (B) *E. coli* MG1655 and MG1655Δ*rtcB* carrying pSC3326 (MazF) or pBAD33 (corresponding empty plasmid) were grown exponentially in LB medium supplemented with 50 µg/ml chloramphenicol and toxin expression was induced for 30 minutes with 0.2% arabinose. RNA was extracted from each sample, reverse transcribed and the fold change in transcript levels of *rtcB, rtcA* and nine genes found to be enriched by CRAC of RtcB during MazF expression (*rplL, rplD, rplM, rpmA, rplU, gcvB, glmY, ilvL* and *rraA*) were determined by qPCR, relative to MG1655/pBAD33. qPCR data were analysed as described for (A). Primer annealing sites for *rplL, rplD, rplM, rpmA, rplU, rpsB, rpsJ, gcvB, glmY, ilvL* and *rraA* used for qPCR analysis are indicated in Figure 33 to Figure 35 (blue arrows).

5.2.2 No evidence for regulation of ribosomal RNA by RtcB

5.2.2.1 RtcB does not influence rRNA levels in standard conditions

CRAC of RtcB revealed a high coverage of reads across all rRNAs, including in the control sample, which was interpreted as non-specific enrichment of highly prevalent RNA species in the experiment (data not shown). The few enriched areas in rRNAs that did not have equivalent peaks in the control were present in all three RtcB samples (RtcB, +Cml and +MazF). Because they were not specific to samples subjected to stress conditions found to increase *rtcB* transcription, they were considered to arise from non-specific interactions between overexpressed RtcB and RNA and therefore not considered biologically relevant.

However, Engl et al. (2016) recently proposed a function for RtcB in ribosome homeostasis that is operational in standard growth conditions. In view of this, the levels of 23S, 16S and 5S rRNA were assessed by qPCR in exponentially-growing *E. coli* MG1655 and MG1655Δ*rtcB* without any exogenous stress. Contrary to the findings of a similar experiment by Engl et al. (2016), no difference was detected in rRNA stability between wild-type cells and those lacking *rtcB* (Figure 37).

5.2.2.2 RtcB does not repair rRNA following amino acid starvation

Ribosomal RNA is known to be degraded under various conditions of nutrient deprivation (Sulthana et al., 2016; Zundel et al., 2009). In collaboration with the Sørensen lab, University of Copenhagen, the role of RtcB in recovery of bacterial growth after a period of amino acid starvation was investigated. The Sørensen group previously observed a rapid reduction in levels of 23S and 16S rRNA during isoleucine starvation, while 5S rRNA remained stable (unpublished observations). Moreover, once starvation was terminated, they found that the rate of bacterial growth quickly recovered.

Therefore, it was hypothesised that RtcB may repair rRNA in order to allow rapid recovery of growth when isoleucine became plentiful. To explore this idea, *E. coli* MG1655 and MG1655Δ*rtcB* were cultured in minimal media and isoleucine starvation was induced by the addition of valine. After 80 minutes, cultures were re-fed with isoleucine to end amino acid starvation. As expected, the growth of the wild-type strain was inhibited upon isoleucine

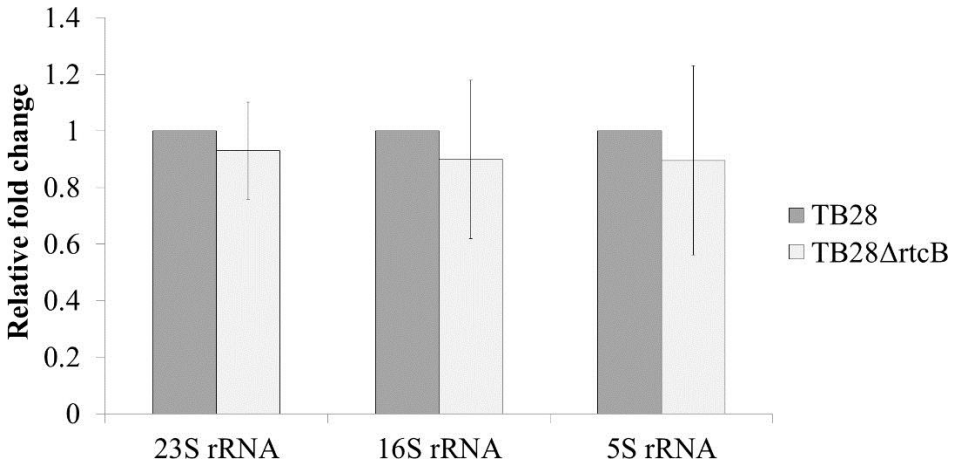


Figure 37 Stability of rRNA is unchanged upon deletion of *rtcB*.

E. coli strains MG1655 and MG1655Δ*rtcB* were grown exponentially in LB medium. RNA was extracted from each sample, reverse transcribed and the fold change in levels of 23S rRNA, 16S rRNA and 5S rRNA in the *rtcB* deletion relative to wild-type MG1655 were determined by qPCR. qPCR data were analysed using the Pfaffl method and normalized against the *tatA* gene. Data from three independent experiments are shown as the mean relative fold change ± standard deviation.

starvation and quickly resumed when starvation was terminated. However, the same pattern of growth was also observed in the strain lacking *rtcB* (data not shown). Further, Northern blot analysis of 23S and 16S rRNA at regular intervals throughout the experiment revealed that levels of rRNA were the same at equivalent time points in wild-type and *rtcB* deletion strains (data not shown; conducted by Dr Michael Sørensen and Mrs Marit Warrer, University of Copenhagen). Thus, RtcB does not appear to mediate rRNA repair during or after amino acid starvation.

5.3 Discussion

In this chapter, we investigated the hypothesis that MazF induction activates RtcB-mediated RNA repair processes in the cell using an RNA-protein crosslinking and deep sequencing approach known as CRAC, with the aim of further elucidating the function of RtcB in *E. coli*. This hypothesis was based on the finding that *rtcB* mRNA levels increased upon MazF expression (Chapter 4) and the knowledge that MazF cleavage yields suitable terminals for RNA ligation by RtcB (Tanaka and Shuman, 2011; Zhang et al., 2005). Chloramphenicol treatment was included in the analysis because it was also found to increase transcription of *rtcB* (Chapter 4). Aside from targeting 23S rRNA (Schlunzen et al., 2001), chloramphenicol is known to have secondary effects on cell physiology as a result of translation inhibition, such as induction of rRNA and tRNA synthesis (Shen and Bremer, 1977; Svenningsen et al., 2017), transcription of genes encoding ribosomal proteins (Dennis, 1976) and stimulation of toxin activity including MazF (Sat et al., 2001). We therefore reasoned that the cellular changes brought about by chloramphenicol may be interesting with respect to possible activities for RtcB.

CRAC of four samples was conducted, namely RtcB expression alone from a plasmid (RtcB), RtcB expression with chloramphenicol treatment (+Cml), RtcB expression with induction of *mazF* (+MazF) and a control lacking RtcB. Overall, there was an abundance of reads in all four samples mapping across the *E. coli* genome. The majority of these reads belonged to highly abundant rRNA and tRNA species and were also highly enriched in the control sample, indicating that a shortcoming of the experiment is bias towards detection of highly transcribed RNAs. Visual analysis led to the identification of 93 RNAs that were significantly enriched in RtcB-containing samples compared to the control. Most were crosslinked to RtcB in the absence of MazF- or chloramphenicol-induced stress, when chromosomal *rtcB* expression is normally very low (Chapters 3 and 4), including many

tRNAs and mRNAs encoding highly prevalent proteins such as elongation factor Tu and acyl carrier protein. This could suggest that RtcB is capable of interacting with many RNAs in the cell and has a broad substrate range, curtailed by the conditions in which it is actively expressed combined with its requirement for a 2′,3′-cyclic phosphate or 3′-phosphate RNA terminal and a 5′-hydroxyl RNA terminal. Another possibility is that abundant RNAs transiently and non-specifically interact with overexpressed RtcB in the experiment and do not represent a genuine physiological interaction. As CRAC analysis of other proteins expressed from plasmids have also revealed crosslinking with highly prevalent RNA species, this supports the latter suggestion (Gerdes group, unpublished observations). Further, the enrichment of *mazF* mRNA in the +MazF sample may also indicate that RNA abundancy influences crosslinking with RtcB.

Aside from *mazF* mRNA, 11 other RNA hits were specific to +MazF and/or +Cml conditions and thus deemed to be of potential biological relevance based on the induction of *rtcB* observed in these conditions. Interestingly, the majority of mRNAs identified encoded ribosomal proteins, supporting a link between RtcB and the translation apparatus, as indicated in Chapter 4. While interactions with non-coding RNAs dominated the data set as a whole, only two were specific to the +MazF and/or +Cml samples, namely the regulatory sRNAs *gcvB* and *glmY*. In addition to a σ^{70}–regulated promoter, transcription of *glmY* is known to occur from a σ^{54}-dependent promoter (Gopel et al., 2011), suggesting a possible regulatory link with *rtcB*. Mapping of ACA sites in the ORFs of the candidate RNAs in the +MazF sample revealed an apparent correlation between a lack of read coverage and the presence of a potential MazF cleavage site, indicating that RtcB may crosslink RNA fragments produced by MazF cleavage, which would support the hypothesis that RtcB ligates MazF-cleaved RNA. However, similar correlations can be observed in the +Cml sample, including those that are not highly enriched in the +MazF sample, and also in the smaller peaks in RtcB only and control samples. As the same peak profiles are largely observed in all samples, this could suggest that these regions may instead be the preferred sites for the RNase cocktail used to trim crosslinked RNAs in the CRAC protocol. In order to clarify which, if any, of the ACA sites are cleaved by MazF and are thus available for RtcB-mediated repair, an examination of the secondary structure of the relevant RNAs would be necessary, as only single-stranded ACA regions are suitable targets for MazF (Zhang et al., 2003), and could be confirmed by primer extension.

To establish whether RtcB influences the stability of the candidate RNA targets identified by CRAC, qPCR analysis was used to assess the levels of each in *rtcB* deletion strains with and

without MazF expression or chloramphenicol treatment relative to untreated wild-type cells. No difference was detected between the levels of the candidate RNAs in wild-type and *rtcB* deletion strains, suggesting that RtcB is not involved in repair of these RNAs in the conditions tested. In addition, in most cases the RNA levels were increased upon MazF expression or chloramphenicol treatment, indicating that these RNA were unlikely to be degraded by MazF cleavage or as a result of chloramphenicol-mediated translation inhibition and like *rtcB* are more prevalent in these conditions. As some of the hits found, specifically *rplL* and *ilvL*, are part of the MazF regulon as reported by Sauert et al. (2016), this provides an explanation for the observed increase in these RNAs in the +MazF sample and could suggest that RtcB repairs these leaderless RNAs to restore canonical translation. Owing to the position of the primers used for qPCR analysis (see Figure 33 to Figure 35), a role for RtcB in this context would likely be missed by this experiment and therefore qPCR analysis using primers that flank the 5′ untranslated region, where MazF reportedly processes mRNAs for selective translation by the stress ribosome (Vesper et al., 2011), may represent a possible avenue for further investigation of these candidates. Alternatively, the increased abundance of these RNAs during MazF expression could indicate a non-specific interaction with RtcB, as observed with abundant RNAs in the RtcB alone sample and indicated by the identification of *mazF* mRNA in the +MazF sample. The same explanation could also apply to the chloramphenicol-treated sample. Accordingly, increased transcription of ribosomal proteins is known to occur during chloramphenicol treatment (Dennis, 1976). Moreover, there was no detectable enrichment of reads in the +MazF sample in the vicinity of the MazF cleavage site in 16S rRNA, which removes the anti-SD sequence and generates the stress ribosome (Vesper et al., 2011). As RtcB has been demonstrated to be capable of repairing this cleavage site and is postulated to mediate recovery as part of a post-MazF stress response (Temmel et al., 2016), this may indicate that the experimental setup is not sufficiently sensitive to detect this interaction or that owing to the enrichment of reads across rRNAs in general, any specific interactions were masked. Therefore, the approach undertaken here to identify and verify candidate RNA substrates of RtcB by CRAC and qPCR is inconclusive with regard to the hypothesis that RtcB mediates repair of MazF-cleaved RNAs and a role for RtcB during chloramphenicol treatment also remains unclear.

We also investigated the levels of rRNA during growth of *E. coli* in standard laboratory conditions, to ascertain whether low levels of *rtcB* transcription in the absence of an activating signal might have a role in maintaining rRNA stability. In contrast to Engl et al. (2016), who recently characterised a role for RtcB in rRNA homeostasis, there was no marked difference

observed in the levels of 23S, 16S or 5S rRNA under standard growth conditions. The reason for this discrepancy is unclear, but supports the finding that the *rtcBA* operon is non-essential for growth in standard conditions (this study; Genschik et al., 1998; Temmel et al., 2016).

A role for RtcB in rRNA repair following degradation during amino acid starvation was also investigated in collaboration with the Sørensen group, University of Copenhagen. *E. coli* MG1655 was found to rapidly resume bacterial growth upon termination of amino acid starvation. However, this was not found to be due to RtcB-mediated repair of rRNA degraded during starvation (not shown).

In summary, we have attempted to learn more about the *in vivo* function of RtcB in *E. coli* by employing CRAC to identify potential RNA substrates of RtcB using conditions found to increase *rtcB* transcription in Chapter 4. However, the outcomes were largely inconclusive, with qPCR analysis of the candidate RNAs specific to the MazF- or chloramphenicol-treated samples revealing there was no difference in stability upon *rtcB* deletion. The high enrichment of reads in rRNA and tRNA genes in the untreated RtcB and control CRAC samples raises the possibility that non-specific interactions may mask genuine interactions in biologically relevant conditions. In addition, we were unable to detect any difference in rRNA levels in the absence of RtcB in standard conditions or during amino acid starvation. Therefore, based on the analyses conducted here, the biological function of RtcB remains elusive.

Chapter 6 Summary and perspectives

In this study, we have confirmed that σ^{54}-dependent transcription of *rtcBA* proceeds in an RtcR-dependent manner in *E. coli* K-12. Using a series of transcriptional *rtcBA-lacZ* fusions, a potential RtcR binding region has been uncovered, which is centred at -144 relative to the transcription start site The region, termed *rtcO*, includes an IR sequence of ATATC-N_{12}-GATAT. Mutation of the IR indicates that it forms part of an RtcR UAS but that additional bases or UASs may also contribute to RtcR binding. In addition, a requirement for IHF in σ^{54}-dependent transcription of *rtcBA* has been established and found to be mediated in a DNA-dependent manner via interaction with at least one of the IHF binding sites in the *rtcBA* promoter. Investigation into the regulation of the bEBP RtcR has revealed the minimum promoter region necessary for *rtcR* transcription, which appears to be regulated by σ^{70}-associated RNAP. A possible autoregulatory feedback loop for *rtcR* transcription has also been identified by which RtcR is postulated to maintain constant and relatively low levels, possibly by binding to *rtcO*, which overlaps the likely *rtcR* promoter region. These findings have contributed to the characterisation of *rtcBA* regulation by the σ^{54}-dependent promoter and are consistent with the known features of transcription by σ^{54}-associated RNAP (Bush and Dixon, 2012; Shingler, 2010). However, in searching for conditions where *rtcBA* transcription was activated, a more complex system of regulation for *rtcBA* was revealed, as a number of RtcR- and σ^{54}-independent conditions were identified. Therefore, it appears that *rtcBA* transcription may be activated from both σ^{54}- and σ^{70}-dependent promoters. Characterising the potential σ^{70}-regulated promoter and identifying the different stimuli for activation of the σ^{54}- and σ^{70}-dependent promoters would further contribute to our understanding of *rtcBA* regulation in *E. coli* and may be applicable to other bacteria with an *rtcBA-rtcR* locus.

We have found no evidence to suggest that *rtcBA* is necessary for the growth of *E. coli* in standard laboratory conditions. This is in contrast to a recent publication that shows a growth phenotype associated with deletion of *rtcBA* and suggests that the *rtcBA* operon is involved in ribosome homeostasis (Engl et al., 2016). In this study, deletion of the *rtcBAR* genes had no effect on the growth rate of *E. coli*, which is supported by the findings of others (Genschik et al., 1998; Temmel et al., 2016) and qPCR analysis and *lacZ* fusion data suggest that levels of *rtcBA* transcription are very low in the absence of activating conditions and/or activated RtcR. Furthermore, rRNA stability was found to be unaffected by deletion of *rtcB*. In view of this,

rtcBA was considered to be induced in response to stress and we searched for cellular conditions that activated *rtcBA* transcription.

Genetic screening resulted in the identification of an activator of *rtcBA.* Expression of Rof, an inhibitor of Rho transcription termination, was found to induce transcription of *rtcBA* in an RtcR-dependent manner, suggesting that RtcB may have a role during dysregulated transcription termination. The function of RtcB in this context is unclear and the signal that induces the CARF-like regulatory domain of RtcR during expression of Rof is unknown (Makarova et al., 2014). Little is known about the conditions in which Rof inhibits Rho activity (Gutierrrez et al., 2007; Pichoff et al., 1998), thus investigation into the regulation of *rof* may enable further characterisation of this phenotype. The levels of *rtcBA* were also found to increase in response to a variety of challenges to the translation apparatus, suggesting that *rtcBA* may respond to the translational status of the cell. This was further investigated using CRAC to identify RNA substrates for RtcB *in vivo* during expression of the endoribonuclease MazF and treatment with the 23S rRNA inhibitor chloramphenicol. Although 12 candidate RNAs were identified in these conditions, it is not clear whether they represent biologically relevant targets of RtcB as their levels were unaffected by the absence of *rtcB.* The function of RtcB in these conditions and during Rof expression could be further analysed by optimising the CRAC protocol. For example, chromosomal *rtcB* could be FTH-tagged to allow expression to be regulated naturally and more robust controls could be used, including an RtcB variant with defective RNA binding (Englert et al., 2012; Maughan and Shuman, 2016). Alternatively, a different approach could be employed such as RNA-Seq, to analyse global changes in the transcriptome in relevant conditions in the presence and absence of RtcB.

In summary, here we have characterised the regulation of bacterial *rtcB* in the context of the σ^{54}-dependent *rtcBA* operon of *E. coli.* We have located a novel DNA element, *rtcO,* confirmed and identified essential regulatory roles for RtcR and IHF respectively and discovered a potential negative feedback mechanism for *rtcR* transcription. Furthermore, we have identified several conditions in which *rtcBA* transcription is activated *in vivo,* including expression of the Rho inhibitory protein Rof and exposure to several agents that target translation, for example MazF and chloramphenicol. Recent publications by Engl et al. (2016) and Temmel et al. (2016) regarding the function of RtcB in *E. coli* complement the findings presented in this study. In particular, Engl et al. (2016) also observed induction of *rtcBA* when translation was impaired by various agents and genetic lesions, while Temmel et al. (2016) characterised a role for RtcB in the repair of 16S rRNA during the MazF stress

response. Together with our data, this indicates that RtcB functions in repair processes in a variety of cellular stresses in *E. coli*. The current state of knowledge regarding the regulation and function of RtcB in *E. coli* is summarised in Figure 38.

As the only bacterial RtcB to be functionally characterised to date, *E. coli* RtcB provides the first indication into the role of this family of proteins in bacteria, as previously described functions in archaea and eukaryotes in tRNA maturation (Popow et al., 2011) and the UPR (Lu et al., 2014) are not relevant to bacteria. Owing to the conservation of the *rtcBA-rtcR* locus across four phyla, the regulatory and functional characteristics of *E. coli* RtcB may be applicable to other bacteria. However, the identification of RtcB paralogues in *M. xanthus* with distinct RNA ligase, DNA ligase and DNA capping efficiencies *in vitro* (Maughan and Shuman, 2015), combined with the coevolution of RtcB with archease, a cofactor known to enhance the *in vitro* activity of RtcB, in some bacteria (Desai et al., 2015) suggests that characterisation of RtcB from multiple bacterial species would be beneficial. It also appears that RtcB has at least one more function in eukaryotes, as a role for RtcB in inhibition of axon regeneration in *C. elegans* was found to be independent of tRNA maturation and the UPR (Kosmaczewski et al., 2015). Therefore, while considerable progress has been made in defining the *in vitro* and *in vivo* activity of RtcB in recent years, further investigation may contribute to a greater understanding of the spectrum of biological functions of RtcB and the regulatory mechanisms governing its activity in organisms from all domains of life.

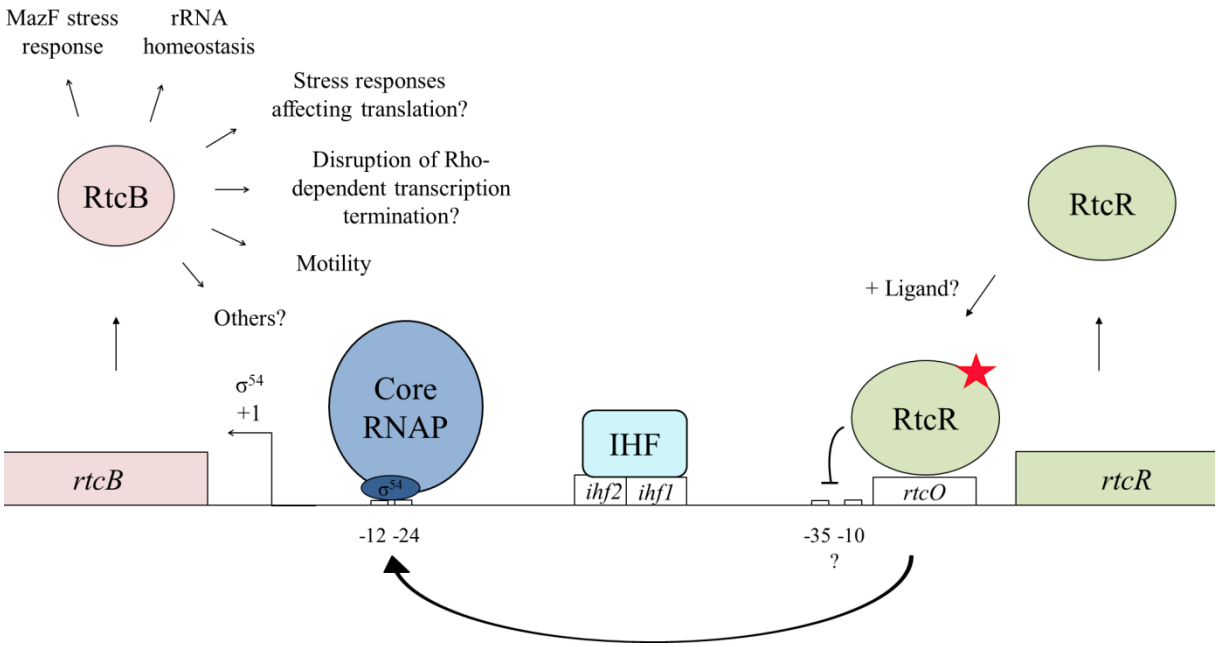
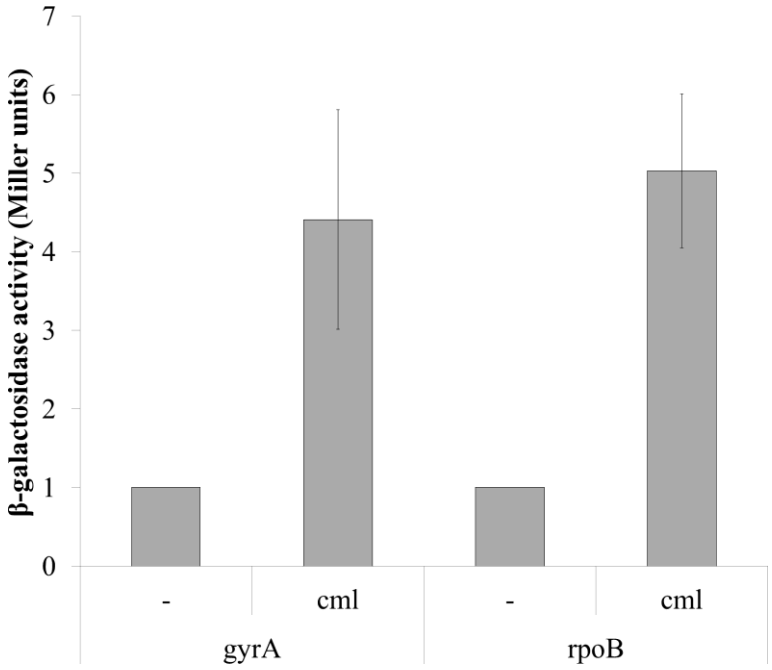


Figure 38 Overview of the regulation and function of the *rtcBA* operon of *E. coli.*

Transcription of *rtcBA* by σ^{54}-associated RNA polymerase binding at the -12/-24 promoter elements is initiated via interaction with activated RtcR. RtcR is expressed from a σ^{70} family-dependent promoter and binds to upstream activating sequence *rtcO*, as indicated, which may mediate negative autoregulation of *rtcR* transcription. The CARF-like regulatory domain of RtcR is activated by as yet undetermined signals (red star), derepressing the central ATPase domain of RtcR. Interaction between σ^{54} and the ATPase domain is facilitated by IHF-induced DNA bending from the IHF binding region (*ihf1* and *ihf2*), as indicated (curved arrow). This results in transcription of *rtcB* from +1. RtcB then mediates a range of cellular processes, as indicated, via its RNA ligase and possibly DNA ligase or capping activities.

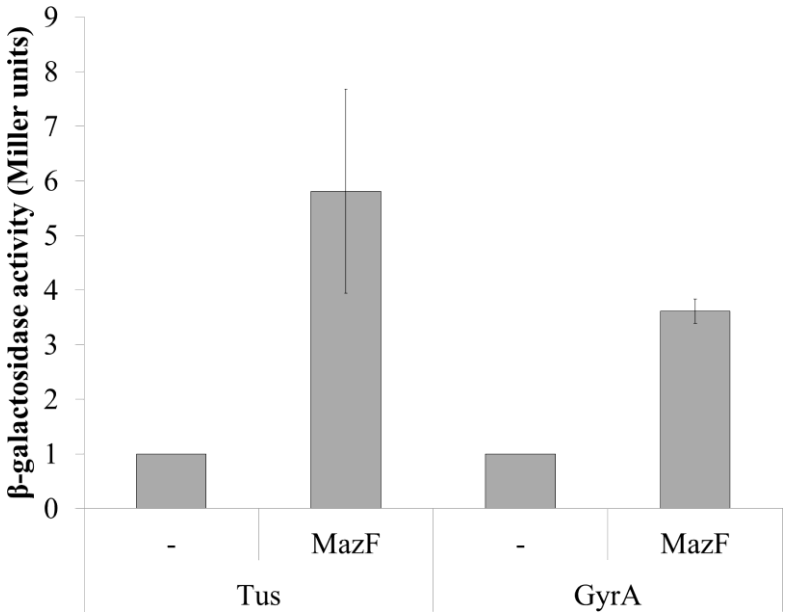
Appendix A



Levels of *rtcB* mRNA increase during chloramphenicol treatment (continued).

E. coli MG1655 was grown exponentially in LB medium then treated with 50 µg/ml chloramphenicol (cml) for 30 minutes. RNA was extracted from each sample, reverse transcribed and the fold change in transcript levels of *rtcB* relative to untreated MG1655 (-) was determined by qPCR. qPCR data were analysed using the Pfaffl method and normalised against the *gyrA* and *rpoB* genes as indicated. Data from three independent experiments are shown as the mean relative fold change ± standard deviation.

Appendix B



Levels of *rtcB* mRNA increase during MazF expression (continued).

E. coli MG1655 carrying carrying pSC3326 was grown exponentially in LB medium then expression of MazF was induced with 0.2% arabinose for 30 minutes. RNA was extracted from each sample, reverse transcribed and the fold change in transcript levels of *rtcB* relative to MG1655 carrying empty pBAD33 (-) was determined by qPCR. qPCR data were analysed using the Pfaffl method and normalised against the *tus* and *gyrA* genes as indicated. Data from three independent experiments are shown as the mean relative fold change ± standard deviation.

References

Abdou, L., Chou, H.T., Haas, D., and Lu, C.D. (2011). Promoter recognition and activation by the global response regulator CbrB in *Pseudomonas aeruginosa*. Journal of Bacteriology *193*, 2784-2792.

Amitsur, M., Levitz, R., and Kaufmann, G. (1987). Bacteriophage-T4 anticodon nuclease, polynucleotide kinase and RNA ligase reprocess the host lysine transfer-RNA. EMBO Journal *6*, 2499-2503.

Aravind, L., Anantharaman, V., Balaji, S., Babu, M.M., and Iyer, L.M. (2005). The many faces of the helix-turn-helix domain: Transcription regulation and beyond. FEMS Microbiology Reviews *29*, 231-262.

Arfin, S.M., Long, A.D., Ito, E.T., Tolleri, L., Riehle, M.M., Paegle, E.S., and Hatfield, G.W. (2000). Global gene expression profiling in *Escherichia coli* K12 - The effects of integration host factor. Journal of Biological Chemistry *275*, 29672-29684.

Arn, E.A., and Abelson, J. (1998). RNA ligases: Function, mechanism, and sequence conservation. RNA Structure and Function *35*, 695-726.

Arn, E.A., and Abelson, J.N. (1996). The 2'-5' RNA ligase of *Escherichia coli* - Purification, cloning, and genomic disruption. Journal of Biological Chemistry *271*, 31145-31153.

Austin, S., and Dixon, R. (1992). The prokaryotic enhancer binding-protein NTRC has an ATPase activity which is phosphorylation and DNA-dependent. EMBO Journal *11*, 2219-2228.

Auxilien, S., El Khadali, F., Rasmussen, A., Douthwaite, S., and Grosjean, H. (2007). Archease from *Pyrococcus abyssi* improves substrate specificity and solubility of a tRNA m(5)C Methyltransferase. Journal of Biological Chemistry *282*, 18711-18721.

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L., and Mori, H. (2006). Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. Molecular Systems Biology *2*, doi:10.1038/msb4100050.

Barrios, H., Valderrama, B., and Morett, E. (1999). Compilation and analysis of sigma 54-dependent promoter sequences. Nucleic Acids Research *27*, 4305-4313.

Berger, D.K., Narberhaus, F., and Kustu, S. (1994). The isolated catalytic domain of NifA, a bacterial enhancer binding protein, activates trnascription *in vitro:* Activation is inhibited by NifL. Proceedings of the National Academy of Sciences USA *91*, 103-107.

Bertoni, G., Fujita, N., Ishihama, A., and de Lorenzo, V. (1998). Active recruitment of sigma 54-RNA polymerase to the Pu promoter of *Pseudomonas putida*: role of IHF and alphaCTD. The EMBO Journal *17*, 5120-5128.

Blank, K., Hensel, M., and Gerlach, R.G. (2011). Rapid and highly efficient method for scarless mutagenesis within the *Salmonella enterica* chromosome. Plos One *6*, doi: 10.1371/journal.pone.0015763.

Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlack, M.C., Heyneker, H.L., Boyer, H.W., Crosa, J.H., and Falkow, S. (1977). Construction and characterisation of new cloning vehicles. II. A multipurpose cloning system. Gene *2*, 95-113.

Bordes, P., Wigneshweraraj, S.R., Chaney, M., Dago, A.E., Morett, E., and Buck, M. (2004). Communication between E sigma 54, promoter DNA and the conserved threonine residue in the GAFTGA motif of the PspF sigma 54-dependent activator during transcription activation. Molecular Microbiology *54*, 489-506.

Bordes, P., Wigneshweraraj, S.R., Schumacher, J., Zhang, X., Chaney, M., and Buck, M. (2003). The ATP hydrolyzing transcription activator phage shock protein F of *Escherichia coli*: Identifying a surface that binds σ54. Proceedings of the National Academy of Sciences *100*, 2278-2283.

Bose, D., Pape, T., Burrows, P.C., Rappas, M., Wigneshweraraj, S.R., Buck, M., and Zhang, X. (2008). Organization of an activator-bound RNA polymerase holoenzyme. Molecular Cell *32*, 337-346.

Bradford, M. (1976). A rapid and sensitive method for the quantification of microgram quantities of a protein utilize the principle of protein-dye binding. Analytical Biochemistry *72*, 248-254.

Brodersen, D.E., Clemons, W.M., Jr., Carter, A.P., Morgan-Warren, R.J., Wimberly, B.T., and Ramakrishnan, V. (2000). The structural basis for the action of the antibiotics tetracycline, pactamycin, and hygromycin B on the 30S ribosomal subunit. Cell *103*, 1143-1154.

Brown, D.R., Barton, G., Pan, Z., Buck, M., and Wigneshweraraj, S. (2014). Nitrogen stress response and stringent response are coupled in *Escherichia coli*. Nat Commun *5*, 4115.

Buck, M., and Cannon, W. (1992). Specific binding of the transcription factor sigma 54 to promoter DNA. Nature *358*, 422-424.

Buck, M., Gallegos, M.T., Studholme, D.J., Guo, J., and Gralla, J.D. (2000). The bacterial enhancer-dependent sigma 54 transcription factor. Journal of Biological Chemistry *182*, 4129-4136.

Buck, M., Miller, S., Drummond, M., and Dixon, R. (1986). Upstream activator sequences are present in the promoters of nitrogen fixation genes. Nature *320*, 374-378.

Burgess, R.R., and Travers, A.A. (1969). Factor stimulating transcription by RNA polymerase. Nature *221*, 43-46.

Burroughs, A.M., and Aravind, L. (2016). RNA damage in biological conflicts and the diversity of responding RNA repair systems. Nucleic Acids Research *44*, 8525-8555.

Burrows, P.C., Severinov, K., Ishihama, A., Buck, M., and Wigneshweraraj, S.R. (2003). Mapping sigma(54)-RNA polymerase interactions at the -24 consensus promoter element. Journal of Biological Chemistry *278*, 29728-29743.

Burrows, P.C., Wigneshweraraj, S.R., and Buck, M. (2008). Protein-DNA interactions that govern AAA+ activator-dependent bacterial transcription initiation. Journal of Molecular Biology *375*, 43-58.

Bush, M., and Dixon, R. (2012). The role of bacterial enhancer binding proteins as specialized activators of sigma 54-dependent transcription. Microbiology and Molecular Biology Reviews *76*, 497-529.

Bush, M., Ghosh, T., Sawicka, M., Moal, I.H., Bates, P.A., Dixon, R., and Zhang, X.D. (2015). The structural basis for enhancer-dependent assembly and activation of the AAA transcriptional activator NorR. Molecular Microbiology *95*, 17-30.

Canaves, J.M. (2004). Predicted role for the archease protein family based on structural and sequence analysis of TM1083 and MTH1598, two proteins structurally characterized through structural genomics efforts. Proteins-Structure Function and Bioinformatics *56*, 19-27.

Cannon, W., Claveriemartin, F., Austin, S., and Buck, M. (1994). Identification of a DNA contacting surface in the transcription factor sigma 54. Molecular Microbiology *11*, 227-236.

Cannon, W., Gallegos, M.T., Casaz, P., and Buck, M. (1999). Amino-terminal sequences of sigma N (sigma 54) inhibit RNA polymerase isomerization. Genes & Development *13*, 357-370.

Carmona, M., Claverie, M.F., and Magasanik, B. (1997). DNA bending and the initiation of transcription at sigma 54-dependent bacterial promoters. Proceedings of the National Academy of Sciences of the United States of America *94*, 9568-9572.

Chakravarty, A.K., and Shuman, S. (2011). RNA 3'-phosphate cyclase (RtcA) catalyzes ligase-like adenylylation of DNA and RNA 5'-monophosphate ends. Journal of Biological Chemistry *286*, 4117-4122.

Chakravarty, A.K., and Shuman, S. (2012). The sequential 2',3'-cyclic phosphodiesterase and 3'-phosphate/5'-OH ligation steps of the RtcB RNA splicing pathway are GTP-dependent. Nucleic Acids Research *40*, 8558-8567.

Chakravarty, A.K., Smith, P., and Shuman, S. (2011). Structures of RNA 3'-phosphate cyclase bound to ATP reveal the mechanism of nucleotidyl transfer and metal-assisted catalysis.

Proceedings of the National Academy of Sciences of the United States of America *108*, 21034-21039.

Chakravarty, A.K., Subbotin, R., Chait, B.T., and Shuman, S. (2012). RNA ligase RtcB splices 3'-phosphate and 5'-OH ends via covalent RtcB-(histidinyl)-GMP and polynucleotide-(3')pp(5')G intermediates. Proceedings of the National Academy of Sciences of the United States of America *109*, 6072-6077.

Chan, P.P., and Lowe, T.M. (2009). GtRNAdb: A database of transfer RNA genes detected in genomic sequence. Nucleic Acids Res *37*, 93-97.

Chaney, M., and Buck, M. (1999). The sigma 54 DNA-binding domain includes a determinant of enhancer responsiveness. Molecular Microbiology *33*, 1200-1209.

Chaney, M., Grande, R., Wigneshweraraj, S.R., Cannon, W., Casaz, P., Gallegos, M.T., Schumacher, J., Jones, S., Elderkin, S., Dago, A.E., *et al.* (2001). Binding of transcriptional activators to sigma 54 in the presence of the transition state analog ADP-aluminum fluoride: insights into activator mechanochemical action. Genes & Development *15*, 2282-2294.

Chauleau, M., Das, U., and Shuman, S. (2015). Effects of DNA(3')pp(5')G capping on 3' end repair reactions and of an embedded pyrophosphate-linked guanylate on ribonucleotide surveillance. Nucleic Acids Research *43*, 3197-3207.

Chen, B., Doucleff, M., Wemmer, D.E., De Carlo, S., Huang, H.H., Nogales, E., Hoover, T.R., Kondrashkina, E., Guo, L., and Nixon, B.T. (2007a). ATP ground- and transition states of bacterial enhancer binding AAA plus ATPases support complex formation with their target protein, sigma 54. Structure *15*, 429-440.

Chen, X., Taylor, D.W., Fowler, C.C., Galan, J.E., Wang, H.-W., and Wolin, S.L. (2013). An RNA degradation machine sculpted by Ro autoantigen and noncoding RNA. Cell *153*, 166-177.

Chen, X.G., Wurtmann, E.J., Van Batavia, J., Zybailov, B., Washburn, M.P., and Wolin, S.L. (2007b). An ortholog of the Ro autoantigen functions in 23S rRNA maturation in *D. radiodurans*. Genes & Development *21*, 1328-1339.

Cheng, H.R., and Jiang, N. (2006). Extremely rapid extraction of DNA from bacteria and yeasts. Biotechnol Lett *28*, 55-59.

Cherepanov, A.V., and de Vries, S. (2002). Kinetic mechanism of the Mg2+-dependent nucleotidyl transfer catalyzed by T4 DNA and RNA ligases. Journal of Biological Chemistry *277*, 1695-1704.

Cherepanov, P.P., and Wackernagel, W. (1995). Gene disruption in *Escherichia coli*: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. Gene *158*, 9-14.

Christensen, S.K., Pedersen, K., Hansen, F.G., and Gerdes, K. (2003). Toxin-antitoxin loci as stress-response-elements: ChpAK/MazF and ChpBK cleave translated RNAs and are counteracted by tmRNA. Journal of Molecular Biology *332*, 809-819.

Chung, C.T., and Miller, R.H. (1988). A rapid and convenient method for the preparation and storage of competent bacterial cells. Nucleic Acids Research *16*, 3580-3580.

Cronan, J.E. (2006). A family of arabinose-inducible *Escherichia coli* expression vectors having pBR322 copy control. Plasmid *55*, 152-157.

Culver, G.M., McCraith, S.M., Consaul, S.A., Stanford, D.R., and Phizicky, E.M. (1997). A 2'-phosphotransferase implicated in tRNA splicing is essential in *Saccharomyces cerevisiae*. Journal of Biological Chemistry *272*, 13203-13210.

D'Autreaux, B., Tucker, N.P., Dixon, R., and Spiro, S. (2005). A non-haem iron centre in the transcription factor NorR senses nitric oxide. Nature *437*, 769-772.

Dago, A.E., Wigneshweraraj, S.R., Buck, M., and Morett, E. (2007). A role for the conserved GAFTGA motif of AAA plus transcription activators in sensing promoter DNA conformation. Journal of Biological Chemistry *282*, 1087-1097.

Das, S., Noe, J.C., Paik, S., and Kitten, T. (2005). An improved arbitrary primed PCR method for rapid characterization of transposon insertion sites. Journal of microbiological methods *63*, 89-94.

Das, U., Chakravarty, A.K., Remus, B.S., and Shuman, S. (2013). Rewriting the rules for end joining via enzymatic splicing of DNA 3'-PO4 and 5'-OH ends. Proceedings of the National Academy of Sciences USA *110*, 20437-20442.

Das, U., Chauleau, M., Ordonez, H., and Shuman, S. (2014). Impact of DNA(3')pp(5')G capping on repair reactions at DNA 3' ends. Proceedings of the National Academy of Sciences of the USA *111*, 11317-11322.

Das, U., and Shuman, S. (2013a). 2'-Phosphate cyclase activity of RtcA: a potential rationale for the operon organization of RtcA with an RNA repair ligase RtcB in *Escherichia coli* and other bacterial taxa. Rna *19*, 1355-1362.

Das, U., and Shuman, S. (2013b). Mechanism of RNA 2',3'-cyclic phosphate end healing by T4 polynucleotide kinase-phosphatase. Nucleic Acids Research *41*, 355-365.

Datsenko, K.A., and Wanner, B.L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proceedings of the National Academy of Sciences of the USA *97*, 6640-6645.

De Carlo, S., Chen, B.Y., Hoover, T.R., Kondrashkina, E., Nogales, E., and Nixon, B.T. (2006). The structural basis for regulated assembly and function of the transcriptional activator NtrC. Genes & Development *20*, 1485-1495.

de Groot, A., Roche, D., Fernandez, B., Ludanyi, M., Cruveiller, S., Pignol, D., Vallenet, D., Armengaud, J., and Blanchard, L. (2014). RNA sequencing and proteogenomics reveal the importance of leaderless mRNAs in the radiation-tolerant bacterium *Deinococcus deserti*. Genome Biology and Evolution *6*, 932-948.

Dennis, P.P. (1976). Effects of chloramphenicol on the transcriptional activities of ribosomal RNA and ribosomal protein genes in *Escherichia coli*. Journal of Molecular Biology *108*, 535-546.

Desai, K.K., Beltrame, A.L., and Raines, R.T. (2015). Coevolution of RtcB and Archease created a multiple-turnover RNA ligase. Rna *21*, 1866-1872.

Desai, K.K., Bingman, C.A., Cheng, C.L., Jr, G.N.P., and Raines, R.T. (2014a). Structure of RNA 3'-phosphate cyclase bound to substrate RNA. Rna *20*, 1560-1566.

Desai, K.K., Bingman, C.A., Phillips, G.N., Jr., and Raines, R.T. (2013). Structures of the noncanonical RNA ligase RtcB reveal the mechanism of histidine guanylylation. Biochemistry *52*, 2518-2525.

Desai, K.K., Cheng, C.L., Bingman, C.A., Phillips, G.N., and Raines, R.T. (2014b). A tRNA splicing operon: Archease endows RtcB with dual GTP/ATP cofactor specificity and accelerates RNA ligation. Nucleic Acids Research *42*, 3931-3942.

Desai, K.K., and Raines, R.T. (2012). tRNA ligase catalyzes the GTP-dependent ligation of RNA with 3'-phosphate and 5'-hydroxyl termini. Biochemistry *51*, 1333-1335.

Ding, N.Z., He, M., He, C.Q., Hu, J.S., Teng, J.L., and Chen, J.G. (2010). Expression and regulation of FAAP in the mouse epididymis. Endocrine *38*, 188-193.

Doucleff, M., Malak, L.T., Pelton, J.G., and Wemmer, D.E. (2005). The C-terminal RpoN domain of sigma 54 forms an unpredicted helix-turn-helix motif similar to domains of sigma 70. Journal of Biological Chemistry *280*, 41530-41536.

Doucleff, M., Pelton, J.G., Lee, P.S., Nixon, B.T., and Wemmer, D.E. (2007). Structural basis of DNA recognition by the alternative sigma-factor, sigma 54. Journal of Molecular Biology *369*, 1070-1078.

Drawz, S.M., and Bonomo, R.A. (2010). Three decades of beta-lactamase inhibitors. Clinical microbiology reviews *23*, 160-201.

Drummond, M.H., Contreras, A., and Mitchenall, L.A. (1990). The function of isolated domains and chimaeric proteins constructed from the transcriptional activators NifA and NtrC of *Klebsiella pneumoniae*. Molecular Microbiology *4*, 29-37.

El Omari, K., Ren, J., Bird, L.E., Bona, M.K., Klarmann, G., LeGrice, S.F., and Stammers, D.K. (2006). Molecular architecture and ligand recognition determinants for T4 RNA ligase. Journal of Biological Chemistry *281*, 1573-1579.

Elderkin, S., Bordes, P., Jones, S., Rappas, M., and Buck, M. (2005). Molecular determinants for PspA-mediated repression of the AAA transcriptional activator PspF. Journal of Bacteriology *187*, 3238-3248.

Engl, C., Schaefer, J., Kotta-Loizou, I., and Buck, M. (2016). Cellular and molecular phenotypes depending upon the RNA repair system RtcAB of *Escherichia coli*. Nucleic Acids Research, doi: 10.1093/nar/gkw1628.

Englert, M., Sheppard, K., Aslanian, A., Yates, J.R., and Soell, D. (2011). Archaeal 3'-phosphate RNA splicing ligase characterization identifies the missing component in tRNA maturation. Proceedings of the National Academy of Sciences of the USA *108*, 1290-1295.

Englert, M., Xia, S., Okada, C., Nakamura, A., Tanavde, V., Yao, M., Eom, S.H., Konigsberg, W.H., Soell, D., and Wang, J. (2012). Structural and mechanistic insights into guanylylation of RNA-splicing ligase RtcB joining RNA between 3'-terminal phosphate and 5'-OH. Proceedings of the National Academy of Sciences of the USA *109*, 15235-15240.

Feklistov, A., Sharon, B.D., Darst, S.A., and Gross, C.A. (2014). Bacterial sigma factors: A historical, structural, and genomic perspective. Annual Review of Microbiology *68*, 357-376.

Fernández, S., de Lorenzo, V., and Pérez-Martin, J. (1995). Activation of the transcriptional regulator XylR of *Pseudomonas putida* by release of repression between functional domains. Molecular Microbiology *16*, 205-213.

Filipowicz, W. (2014). Making ends meet: a role of RNA ligase RTCB in unfolded protein response. Embo Journal *33*, 2887-2889.

Filipowicz, W., Konarska, M., Gross, H.J., and Shatkin, A.J. (1983). RNA 3'-terminal phosphate cyclase activity and RNA ligation in Hela-cell extract. Nucleic Acids Research *11*, 1405-1418.

Filipowicz, W., and Shatkin, A.J. (1983). Origin of splice junction phosphate in tRNAs processed by HeLa cell extract. Cell *32*, 547-557.

Filipowicz, W., Strugala, K., Konarska, M., and Shatkin, A.J. (1985). Cyclization of RNA 3'-terminal phosphate by cyclase from Hela cells proceeds via formation of N(3')pp(5')A activated intermediate. Proceedings of the National Academy of Sciences of the USA *82*, 1316-1320.

François, B., Russell, R.J.M., Murray, J.B., Aboul-ela, F., Masquida, B., Vicens, Q., and Westhof, E. (2005). Crystal structures of complexes between aminoglycosides and decoding A site oligonucleotides: role of the number of rings and positive charges in the specific binding leading to miscoding. Nucleic Acids Research *33*, 5677-5690.

Friedman, D.I. (1988). Integration host factor - a protein for all reasons. Cell *55*, 545-554.

Gallegos, M.T., and Buck, M. (1999). Sequences in σN determining holoenzyme formation and properties. Journal of Molecular Biology *288*, 539-553.

Gardner, A.M., Gessner, C.R., and Gardner, P.R. (2003). Regulation of the nitric oxide reduction operon (*norRVW*) in *Escherichia coli* - Role of NorR and sigma 54 in the nitric oxide stress response. Journal of Biological Chemistry *278*, 10081-10086.

Genschik, P., Billy, E., Swianiewicz, M., and Filipowicz, W. (1997). The human RNA 3'-terminal phosphate cyclase is a member of a new family of proteins conserved in Eucarya, Bacteria and Archaea. EMBO Journal *16*, 2955-2967.

Genschik, P., Drabikowski, K., and Filipowicz, W. (1998). Characterisation of the *Escherichia coli* RNA 3'-terminal phosphate cyclase and its sigma 54-regulated operon. Journal of Biological Chemistry *273*, 25516-25526.

Ghosh, T., Bose, D., and Zhang, X. (2010). Mechanisms for activating bacterial RNA polymerase. FEMS Microbiology Reviews *34*, 611-627.

Gomes, I., and Gupta, R. (1997). RNA splicing ligase activity in the Archaeon *Haloferax volcanii*. Biochemical and Biophysical Research Communications *237*, 588-594.

Goodrich, J.A., Schwartz, M.L., and McClure, W.R. (1990). Searching for and predicting the activity of sites for DNA-binding proteins - compilation and analysis of the binding-sites for *Escherichia coli* integration host factor (IHF). Nucleic Acids Research *18*, 4993-5000.

Gopel, Y., Luttmann, D., Heroven, A.K., Reichenbach, B., Dersch, P., and Gorke, B. (2011). Common and divergent features in transcriptional control of the homologous small RNAs GlmY and GlmZ in Enterobacteriaceae. Nucleic Acids Research *39*, 1294-1309.

Granneman, S., Kudla, G., Petfalski, E., and Tollervey, D. (2009). Identification of protein binding sites on U3 snoRNA and pre-rRNA by UV cross-linking and high-throughput analysis of cDNAs. Proceedings of the National Academy of Sciences USA *106*, 9613-9618.

Greer, C.L., Javor, B., and Abelson, J. (1983a). RNA ligase in bacteria - formation of a 2',5' linkage by an *Escherichia coli* extract. Cell *33*, 899-906.

Greer, C.L., Peebles, C.L., Gegenheimer, P., and Abelson, J. (1983b). Mechanism of action of a yeast RNA ligase in tRNA splicing. Cell *32*, 537-546.

Grylak-Mielnicka, A., Bidnenko, V., Bardowski, J., and Bidnenko, E. (2016). Transcription termination factor Rho: a hub linking diverse physiological processes in bacteria. Microbiology *162*, 433-447.

Guo, Y.L., and Gralla, J.D. (1997). DNA-binding determinants of sigma 54 as deduced from libraries of mutations. Journal of Bacteriology *179*, 1239-1245.

Gutierrrez, P., Kozlov, G., Gabrielli, L., Elias, D., Osborne, M.J., Gallouzi, I.E., and Gehring, K. (2007). Solution structure of YaeO, a Rho-specific inhibitor of transcription termination. Journal of Biological Chemistry *282*, 23348-23353.

Guzman, L.M., Belin, D., Carson, M.J., and Beckwith, J. (1995). Tight regulation, modulation, and high-level expression by vectors containing the arabinose P-BAD promoter. Journal of Bacteriology *177*, 4121-4130.

Hales, L.M., Gumport, R.I., and Gardner, J.F. (1994). Determining the DNA sequence elements required for binding integration host factor to two different target sites. Journal of Bacteriology *176*, 2999-3006.

Hamamichi, S., Rivas, R.N., Knight, A.L., Cao, S., Caldwell, K.A., and Caldwell, G.A. (2008). Hypothesis-based RNAi screening identifies neuroprotective genes in a Parkinson's disease model. Proceedings of the National Academy of Sciences USA *105*, 728-733.

Hausner, G., Hafez, M., and Edgell, D.R. (2014). Bacterial group I introns: mobile RNA catalysts. Mobile DNA *5*, http://www.mobilednajournal.com/content/5/1/8.

Hazan, R., Sat, B., and Engelberg-Kulka, H. (2004). *Escherichia coli* MazEF-mediated cell death is triggered by various stressful conditions. Journal of Bacteriology *186*, 3663-3669.

Hazari, Y.M., Bashir, A., Haq, E.u., and Fazili, K.M. (2016). Emerging tale of UPR and cancer: an essentiality for malignancy. Tumor Biology *37*, 14381-14390.

Ho, C.K., Wang, L.K., Lima, C.D., and Shuman, S. (2004). Structure and mechanism of RNA ligase. Structure *12*, 327-339.

Hollien, J. (2013). Evolution of the unfolded protein response. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research *1833*, 2458-2463.

Hu, J.S., Teng, J.L., Ding, N.Z., He, M., Sun, Y.H., Yu, A.C.H., and Chen, J.G. (2008). FAAP, a novel murine protein, is involved in cell adhesion through regulating vinculin-paxillin association. Frontiers in Bioscience-Landmark *13*, 7123-7131.

Huo, Y.X., Tian, Z.X., Rappas, M., Wen, J., Chen, Y.C., You, C.H., Zhang, X., Buck, M., Wang, Y.P., and Kolb, A. (2006). Protein-induced DNA bending clarifies the architectural organization of the sigma 54-dependent *glnAp2* promoter. Mol Microbiol *59*, 168-180.

Huo, Y.X., Zhang, Y.T., Xiao, Y., Zhang, X.D., Buck, M., Kolb, A., and Wang, Y.P. (2009). IHF-binding sites inhibit DNA loop formation and transcription initiation. Nucleic Acids Research *37*, 3878-3886.

Ishihama, A. (2000). Functional modulation of *Escherichia coli* RNA polymerase. Annual Review of Microbiology *54*, 499-518.

Jiang, Y., Pogliano, J., Helinski, D.R., and Konieczny, I. (2002). ParE toxin encoded by the broad-host-range plasmid RK2 is an inhibitor of *Escherichia coli* gyrase. Molecular Microbiology *44*, 971-979.

Jovanovic, G., Dworkin, J., and Model, P. (1997). Autogenous control of PspF, a constitutively active enhancer-binding protein of *Escherichia coli*. Journal of Bacteriology *179*, 5232-5237.

Jovanovic, G., and Model, P. (1997). PspF and IHF bind co-operatively in the psp promoter-regulatory region of *Escherichia coli*. Molecular Microbiology *25*, 473-481.

Jovanovic, G., Rakonjac, J., and Model, P. (1999). *In vivo* and *in vitro* activities of the *Escherichia coli* σ54 transcription activator, PspF, and its DNA-binding mutant, PspFΔHTH1. Journal of Molecular Biology *285*, 469-483.

Jurkin, J., Henkel, T., Nielsen, A.F., Minnich, M., Popow, J., Kaufmann, T., Heindl, K., Hoffmann, T., Busslinger, M., and Martinez, J. (2014). The mammalian tRNA ligase complex mediates splicing of XBP1 mRNA and controls antibody secretion in plasma cells. EMBO Journal *33*, 2922-2936.

Justino, M.C., Goncalves, V.M.M., and Saraiva, L.M. (2005). Binding of NorR to three DNA sites is essential for promoter activation of the flavorubredoxin gene, the nitric oxide reductase of *Escherichia coli*. Biochemical and Biophysical Research Communications *328*, 540-544.

Kanai, Y., Dohmae, N., and Hirokawa, N. (2004). Kinesin transports RNA: Isolation and characterization of an RNA-transporting granule. Neuron *43*, 513-525.

Kato, M., Shirouzu, M., Terada, T., Yamaguchi, H., Murayama, K., Sakai, H., Kuramitsu, S., and Yokoyama, S. (2003). Crystal structure of the 2′-5′ RNA ligase from *Thermus thermophilus* HB8. Journal of Molecular Biology *329*, 903-911.

Keppetipola, N., Nandakumar, J., and Shuman, S. (2007). Reprogramming the tRNA-splicing activity of a bacterial RNA repair enzyme. Nucleic Acids Research *35*, 3624-3630.

Keseler, I.M., Collado-Vides, J., Santos-Zavaleta, A., Peralta-Gil, M., Gama-Castro, S., Muniz-Rascado, L., Bonavides-Martinez, C., Paley, S., Krummenacker, M., Altman, T., *et al.* (2011). EcoCyc: a comprehensive database of Escherichia coli biology. Nucleic Acids Research *39*, D583-D590.

Kiebler, M.A., and Bassell, G.J. (2006). Neuronal RNA granules: Movers and makers. Neuron *51*, 685-690.

Kitagawa, M., Ara, T., Arifuzzaman, M., Ioka-Nakamichi, T., Inamoto, E., Toyonaga, H., and Mori, H. (2005). Complete set of ORF clones of *Escherichia coli* ASKA library (a complete set of *E. coli* K-12 ORF archive): Unique resources for biological research. DNA Research *12*, 291-299.

Kjems, J., and Garrett, R.A. (1988). Novel splicing mechanisms for the ribosomal RNA intron in the archaebacterium *Desulfurococcus mobilis*. Cell *54*, 693-703.

Konarska, M., Filipowicz, W., Domdey, H., and Gross, H.J. (1981). Formation of a 2'-phosphomonoester, 3',5'-phosphodiester linkage by a novel RNA ligase in wheat germ. Nature *293*, 112-116.

Kosmaczewski, S.G., Edwards, T.J., Han, S.M., Eckwahl, M.J., Meyer, B.I., Peach, S., Hesselberth, J.R., Wolin, S.L., and Hammarlund, M. (2014). The RtcB RNA ligase is an essential component of the metazoan unfolded protein response. EMBO Reports *15*, 1278-1285.

Kosmaczewski, S.G., Han, S.M., Han, B., Meyer, B.I., Baig, H.S., Athar, W., Lin-Moore, A.T., Koelle, M.R., and Hammarlund, M. (2015). RNA ligation in neurons by RtcB inhibits axon regeneration. Proceedings of the National Academy of Sciences USA *112*, 8451-8456.

Kustu, S., Santero, E., Keener, J., Popham, D., and Weiss, D. (1989). Expression of sigma 54 (NtrA)-dependent genes is probably united by a common mechanism. Microbiological Reviews *53*, 367-376.

Leonhartsberger, S., Huber, A., Lottspeich, F., and Bock, A. (2001). The *hydH/G* genes from *Escherichia coli* code for a zinc and lead responsive two-component regulatory system. Journal of Molecular Biology *307*, 93-105.

Liang, Z.-Y., Lai, H.-Y., Yang, H., Zhang, C.-J., Yang, H., Wei, H.-H., Chen, X.-X., Zhao, Y.-W., Su, Z.-D., Li, W.-C., *et al.* (2017). Pro54DB: a database for experimentally verified sigma 54 promoters. Bioinformatics *33*, 467-469.

Lintner, N.G., Frankel, K.A., Tsutakawa, S.E., Alsbury, D.L., Copie, V., Young, M.J., Tainer, J.A., and Lawrence, C.M. (2011). The structure of the CRISPR-associated protein Csa3 provides insight into the regulation of the CRISPR/Cas system. Journal of Molecular Biology *405*, 939-955.

Little, R., and Dixon, R. (2003). The amino-terminal GAF domain of *Azotobacter vinelandii* NifA binds 2-oxoglutarate to resist inhibition by NifL under nitrogen-limiting conditions. Journal of Biological Chemistry *278*, 28711-28718.

Lu, Y., Liang, F.-X., and Wang, X. (2014). A synthetic biology approach identifies the mammalian UPR RNA ligase RtcB. Molecular Cell *55*, 758-770.

Ma, H., Qi, M.Y., Zhang, X., Zhang, Y.L., Wang, L., Li, Z.Q., Fu, B., Wang, W.T., and Liu, D. (2014). HSPC117 is regulated by epigenetic modification and is involved in the migration of JEG-3 cells. International Journal of Molecular Sciences *15*, 10936-10949.

Maeda, H., Fujita, N., and Ishihama, A. (2000). Competition among seven *Escherichia coli* σ subunits: relative binding affinities to the core RNA polymerase. Nucleic Acids Research *28*, 3497-3503.

Magasanik, B. (1988). Reversible phosphorylation of an enhancer binding protein regulates the transcription of bacterial nitrogen utilisation genes. Trends in Biochemical Sciences *13*, 475-479.

Makarova, K.S., Anantharaman, V., Grishin, N.V., Koonin, E.V., and Aravind, L. (2014). CARE and WYL domains: ligand-binding regulators of prokaryotic defense systems. Frontiers in Genetics *5*, doi: 10.3389/fgene.2014.00102.

Makarova, K.S., Wolf, Y.I., and Koonin, E.V. (2013). Comparative genomics of defense systems in archaea and bacteria. Nucleic Acids Research *41*, 4360-4377.

Maughan, W.P., and Shuman, S. (2015). Characterization of 3 '-phosphate RNA ligase paralogs RtcB1, RtcB2, and RtcB3 from *Myxococcus xanthus* highlights DNA and RNA 5 '-phosphate capping activity of RtcB3. Journal of Bacteriology *197*, 3616-3624.

Maughan, W.P., and Shuman, S. (2016). Distinct contributions of enzymic functional groups to the 2',3'-cyclic phosphodiesterase, 3'-phosphate guanylylation, and 3'-ppG/5'-OH ligation steps of the *Escherichia coli* RtcB nucleic acid splicing pathway. Journal of Bacteriology *198*, 1294-1304.

McManus, M.T., Shimamura, M., Grams, J., and Hajduk, S.L. (2001). Identification of candidate mitochondrial RNA editing ligases from *Trypanosoma brucei*. Rna *7*, 167-175.

Miller, J.H. (1972). Experiments in Molecular Genetics (Cold Spring Harbour, New York: Cold Spring Harbour Laboratory Press).

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

Molina-Serrano, D., Marques, J., Nohales, M.-A., Flores, R., and Daros, J.-A. (2012). A chloroplastic RNA ligase activity analogous to the bacterial and archaeal 2'-5' RNA ligase. RNA Biology *9*, 326-333.

Mooney, R.A., Darst, S.A., and Landick, R. (2005). Sigma and RNA polymerase: An on-again, off-again relationship? Molecular Cell *20*, 335-345.

Morris, L., Cannon, W., Claveriemartin, F., Austin, S., and Buck, M. (1994). DNA distortion and nucleation of local DNA unwinding within sigma 54 (sigma N) holoenzyme closed promoter complexes. Journal of Biological Chemistry *269*, 11563-11571.

Nandakumar, J., and Shuman, S. (2004). How an RNA ligase discriminates RNA versus DNA damage. Molecular Cell *16*, 211-221.

Nash, H.A., Robertson, C.A., Flamm, E., Weisberg, R.A., and Miller, H.I. (1987). Overproduction of *Escherichia coli* integration host factor, a protein with nonidentical subunits. Journal of Bacteriology *169*, 4124-4127.

Nix, P., Hammarlund, M., Hauth, L., Lachnit, M., Jorgensen, E.M., and Bastiani, M. (2014). Axon regeneration gene identified by RNAi screening in *C. elegans*. Journal of Neuroscience *34*, 629-645.

North, A.K., and Kustu, S. (1997). Mutant forms of the enhancer-binding protein NtrC can activate transcription from solution. Journal of Molecular Biology *267*, 17-36.

O'Neill, E., Ng, L.C., Sze, C.C., and Shingler, V. (1998). Aromatic ligand binding and intramolecular signalling of the phenol-responsive sigma 54-dependent regulator DmpR. Molecular Microbiology *28*, 131-141.

Okada, C., Maegawa, Y., Yao, M., and Tanaka, I. (2006). Crystal structure of an RtcB homolog protein (PH1602-extein protein) from *Pyrococcus horikoshii* reveals a novel fold. Proteins - Structure Function and Bioinformatics *63*, 1119-1122.

Osterberg, S., del Peso-Santos, T., and Shingler, V. (2011). Regulation of alternative sigma factor use. Annual Review of Microbiology *65*, 37-55.

Pascal, J.M. (2008). DNA and RNA ligases: structural variations and shared mechanisms. Current Opinion in Structural Biology *18*, 96-105.

Pedersen, H., and Valentin-Hansen, P. (1997). Protein-induced fit: the CRP activator protein changes sequence-specific DNA recognition by the CytR repressor, a highly flexible LacI member. EMBO Journal *16*, 2108-2118.

Pedersen, K., Christensen, S.K., and Gerdes, K. (2002). Rapid induction and reversal of a bacteriostatic condition by controlled expression of toxins and antitoxins. Molecular Microbiology *45*, 501-510.

Pedersen, K., Zavialov, A.V., Pavlov, M.Y., Elf, J., Gerdes, K., and Ehrenberg, M. (2003). The bacterial toxin RelE displays codon-specific cleavage of mRNAs in the ribosomal A site. Cell *112*, 131-140.

Peebles, C.L., Ogden, R.C., and Knapp, G. (1979). Splicing of yeast tRNA precursors: a two-stage reaction. Cell *18*, 27-35.

Perez-Gonzalez, A., Pazo, A., Navajas, R., Ciordia, S., Rodriguez-Frandsen, A., and Nieto, A. (2014). hCLE/C14orf166 associates with DDX1-HSPC117-FAM98B in a novel transcription-dependent shuttling RNA-transporting complex. PLOS One *9*, doi:10.1371/journal.pone.0090957.

Perez-Martin, J., and De Lorenzo, V. (1995). Integration host factor supresses promiscuous activation of the sigma 54-dependent promoter *Pu* of *Pseudomonas putida*. Proceedings of the National Academy of Sciences USA *92*, 7277-7281.

Peters, J.M., Vangeloff, A.D., and Landick, R. (2011). Bacterial transcription terminators: The RNA 3'-end chronicles. Journal of Molecular Biology *412*, 793-813.

Pfaffl, M.W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Research *29*, 2002-2007.

Pichoff, S., Alibaud, L., Guedant, A., Castanie, M.P., and Bouche, J.P. (1998). An *Escherichia coli* gene (*yaeO*) suppresses temperature-sensitive mutations in essential genes by modulating Rho-dependent transcription termination. Molecular Microbiology *29*, 859-869.

Pope, W.H., Anders, K.R., Baird, M., Bowman, C.A., Boyle, M.M., Broussard, G.W., Chow, T., Clase, K.L., Cooper, S., Cornely, K.A., *et al.* (2014). Cluster M mycobacteriophages Bongo, PegLeg, and Rey with unusually large repertoires of tRNA isotypes. Journal of virology *88*, 2461-2480.

Popow, J., Englert, M., Weitzer, S., Schleiffer, A., Mierzwa, B., Mechtler, K., Trowitzsch, S., Will, C.L., Luehrmann, R., Soell, D., *et al.* (2011). HSPC117 is the essential subunit of a human tRNA splicing ligase complex. Science *331*, 760-764.

Popow, J., Jurkin, J., Schleiffer, A., and Martinez, J. (2014). Analysis of orthologous groups reveals archease and DDX1 as tRNA splicing factors. Nature *511*, 104-107.

Popow, J., Schleiffer, A., and Martinez, J. (2012). Diversity and roles of (t)RNA ligases. Cellular and Molecular Life Sciences *69*, 2657-2670.

Porter, S.C., North, A.K., Wedel, A.B., and Kustu, S. (1993). Oligomerization of NTRC at the *glnA* enhancer is required for transcriptional activation. Genes & Development *7*, 2258-2273.

Prysak, M.H., Mozdzierz, C.J., Cook, A.M., Zhu, L., Zhang, Y., Inouye, M., and Woychik, N.A. (2009). Bacterial toxin YafQ is an endoribonuclease that associates with the ribosome and blocks translation elongation through sequence-specific and frame-dependent mRNA cleavage. Molecular Microbiology *71*, 1071-1087.

Rappas, M., Bose, D., and Zhang, X.D. (2007). Bacterial enhancer-binding proteins: unlocking sigma 54-dependent gene transcription. Current Opinion in Structural Biology *17*, 110-116.

Ray, A., Zhang, S., Rentas, C., Caldwell, K.A., and Caldwell, G.A. (2014). RTCB-1 mediates neuroprotection via *XBP-1* mRNA splicing in the unfolded protein response pathway. Journal of Neuroscience *34*, 16076-16085.

Rehse, P.H., and Tahirov, H.T. (2005). Structure of a putative 2 '-5 ' RNA ligase from *Pyrococcus horikoshii*. Acta Crystallographica Section D-Biological Crystallography *61*, 1207-1212.

Reinberg, D., Arenas, J., and Hurwitz, J. (1985). The enzymatic conversion of 3'-phosphate terminated RNA chains to 2',3'-cyclic phosphate derivatives. Journal of Biological Chemistry *260*, 6088-6097.

Reinholdhurek, B., and Shub, D.A. (1992). Self-splicing introns in transfer-RNA genes of widely divergent bacteria. Nature *357*, 173-176.

Reitzer, L. (2003). Nitrogen assimilation and global regulation in *Escherichia coli*. Annual Review of Microbiology *57*, 155-176.

Reitzer, L., and Schneider, B.L. (2001). Metabolic context and possible physiological themes of sigma 54-dependent genes in *Escherichia coli*. Microbiology and Molecular Biology Reviews *65*, 422-+.

Reitzer, L.J., Movsas, B., and Magasanik, B. (1989). Activation of *glnA* transcription by nitrogen regulator-I (NRI)-phosphate in *Escherichia coli* - evidence for a long-range physical interaction between NRI-phosphate and RNA polymerase. Journal of Bacteriology *171*, 5512-5522.

Remus, B.S., Jacewicz, A., and Shuman, S. (2014). Structure and mechanism of *E. coli* RNA 2',3'-cyclic phosphodiesterase. Rna *20*, 1697-1705.

Remus, B.S., and Shuman, S. (2014). Distinctive kinetics and substrate specificities of plant and fungal tRNA ligases. Rna *20*, 462-473.

Rice, P., Longden, I., and Bleasby, A. (2000). EMBOSS: The European molecular biology open softwear suite. Trends in Genetics *16*, 276-277.

Ringquist, S., Shinedling, S., Barrick, D., Green, L., Binkley, J., Stormo, G.D., and Gold, L. (1992). Translation initiation in *Escherichia coli* - sequences within the ribosome binding site. Molecular Microbiology *6*, 1219-1229.

Rossignol, M., Basset, A., Espeli, O., and Boccard, F. (2001). NKBOR, a mini-Tn10-based transposon for random insertion in the chromosome of Gram-negative bacteria and the rapid recovery of sequences flanking the insertion sites in *Escherichia coli*. Research in microbiology *152*, 481-485.

Rubin, E.J., Akerley, B.J., Novik, V.N., Lampe, D.J., Husson, R.N., and Mekalanos, J.J. (1999). *In vivo* transposition of mariner-based elements in enteric bacteria and mycobacteria. Proceedings of the National Academy of Sciences USA *96*, 1645-1650.

Saecker, R.M., Record, M.T., and Dehaseth, P.L. (2011). Mechanism of bacterial transcription initiation: RNA polymerase - promoter binding, isomerization to initiation-competent open complexes, and initiation of RNA synthesis. Journal of Molecular Biology *412*, 754-771.

Sallai, L., and Tucker, P.A. (2005). Crystal structure of the central and C-terminal domain of the sigma 54-activator ZraR. J. Struct. Biol. *151*, 160-170.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular cloning: A laboratory manual (Cold Spring Harbour, New York: Cold Spring Harbour Laboratory Press).

Sasse-Dwight, S., and Gralla, J.D. (1988). Probing the *Escherichia coli glnALG* upstream activation mechanism *in vivo*. Proceedings of the National Academy of Sciences USA *85*, 8934-8938.

Sat, B., Hazan, R., Fisher, T., Khaner, H., Glaser, G., and Engelberg-Kulka, H. (2001). Programmed cell death in *Escherichia coli*: Some antibiotics can trigger *mazEF* lethality. Journal of Bacteriology *183*, 2041-2045.

Sauert, M., Wolfinger, M.T., Vesper, O., Muller, C., Byrgazov, K., and Moll, I. (2016). The MazF-regulon: a toolbox for the post-transcriptional stress response in *Escherichia coli*. Nucleic Acids Research *44*, 6660-6675.

Schaefer, J., Engl, C., Zhang, N., Lawton, E., and Buck, M. (2015). Genome wide interactions of wild-type and activator bypass forms of sigma 54. Nucleic Acids Research *43*, 7280-7291.

Schlunzen, F., Zarivach, R., Harms, J., Bashan, A., Tocilj, A., Albrecht, R., Yonath, A., and Franceschi, F. (2001). Structural basis for the interaction of antibiotics with the peptidyl transferase centre in eubacteria. Nature *413*, 814-821.

Sekine, S.-I., Tagami, S., and Yokoyama, S. (2012). Structural basis of transcription by bacterial and eukaryotic RNA polymerases. Current Opinion in Structural Biology *22*, 110-118.

Selvaraj, S., Sambandam, V., Sardar, D., and Anishetty, S. (2012). In silico analysis of DosR regulon proteins of *Mycobacterium tuberculosis*. Gene *506*, 233-241.

Senčilo, A., Jacobs-Sera, D., Russell, D.A., Ko, C.-C., Bowman, C.A., Atanasova, N.S., Österlund, E., Oksanen, H.M., Bamford, D.H., Hatfull, G.F., *et al.* (2013). Snapshot of haloarchaeal tailed virus genomes. RNA Biology *10*, 803-816.

Sharma, A., Leach, R.N., Gell, C., Zhang, N., Burrows, P.C., Shepherd, D.A., Wigneshweraraj, S., Smith, D.A., Zhang, X.D., Buck, M., *et al.* (2014). Domain movements of the enhancer-dependent sigma factor drive DNA delivery into the RNA polymerase active site: Insights from single molecule studies. Nucleic Acids Research *42*, 5177-5190.

Shen, V., and Bremer, H. (1977). Chloramphenicol-induced changes in the synthesis of ribosomal, transfer and messenger ribonucleic acids in *Escherichia coli*. Journal of Bacteriology *130*, 1096-1108.

Shimamoto, N., Kamigochi, T., and Utiyama, H. (1986). Release of the sigma subunit of *Escherichia coli:* DNA-dependent RNA polymerase depends mainly on time elapsed after the start of initiation, not on length of product RNA. Journal of Biological Chemistry *261*, 1859-1865.

Shingler, V. (2010). Signal sensory systems that impact sigma 54-dependent transcription. FEMS Microbiology Reviews *35*, 425-440.

Shingler, V., and Moore, T. (1994). Sensing of aromatic compounds by the DmpR transcriptional activator of phenol-catabolising *Psuedomonas* sp. Strains CF600. Journal of Bacteriology *176*, 1555-1560.

Shultzaberger, R.K., Chen, Z., Lewis, K.A., and Schneider, T.D. (2006). Anatomy of *Escherichia coli* sigma 70 promoters. Nucleic Acids Research *35*, 771-788.

Sidrauski, C., Cox, J.S., and Walter, P. (1996). tRNA ligase is required for regulated mRNA splicing in the unfolded protein response. Cell *87*, 405-413.

Siegel, A.R., and Wemmer, D.E. (2016). Role of the sigma 54 activator interacting domain in bacterial transcription initiation. Journal of Molecular Biology *428*, 4669-4685.

Silber, R., Malathi, V.G., and Hurwitz, J. (1972). Purification and properties of bacteriophage T4-induced RNA ligase. Proceedings of the National Academy of Sciences USA *69*, 3009-3013.

Sim, S., and Wolin, S.L. (2011). Emerging roles for the Ro 60-kDa autoantigen in noncoding RNA metabolism. Wiley Interdisciplinary Reviews - RNA *2*, 686-699.

Singh, S.M., and Panda, A.K. (2005). Solubilization and refolding of bacterial inclusion body proteins. Journal of Bioscience and Bioengineering *99*, 303-310.

Snider, J., Thibault, G., and Houry, W.A. (2008). The AAA+ superfamily of functionally diverse proteins. Genome Biol. *9*, 216-216.

Studholme, D.J., and Dixon, R. (2004). In silico analysis of the σ54-dependent enhancer-binding proteins in *Pirellula* species strain 1. FEMS Microbiology Letters *230*, 215-225.

Sulthana, S., Basturea, G.N., and Deutscher, M.P. (2016). Elucidation of pathways of ribosomal RNA degradation: an essential role for RNase E. Rna *22*, 1163-1171.

Svenningsen, S.L., Kongstad, M., Stenum, T.S., Munoz-Gomez, A.J., and Sorensen, M.A. (2017). Transfer RNA is highly unstable during early amino acid starvation in *Escherichia coli*. Nucleic Acids Research *45*, 793-804.

Sysoeva, T.A., Chowdhury, S., Guo, L., and Nixon, B.T. (2013). Nucleotide-induced asymmetry within ATPase activator ring drives σ54–RNAP interaction and ATP hydrolysis. Genes & Development *27*, 2500-2511.

Sze, C.C., Laurie, A.D., and Shingler, V. (2001). *In vivo* and *in vitro* effects of integration host factor at the DmpR-regulated sigma 54-dependent Po promoter. Journal of Bacteriology *183*, 2842-2851.

Tanaka, N., Chakravarty, A.K., Maughan, B., and Shuman, S. (2011a). Novel mechanism of RNA repair by RtcB via sequential 2',3'-cyclic phosphodiesterase and 3'-phosphate/5'-hydroxyl ligation reactions. Journal of Biological Chemistry *286*, 43134-43143.

Tanaka, N., Meineke, B., and Shuman, S. (2011b). RtcB, a novel RNA ligase, can catalyze tRNA splicing and *HAC1* mRNA splicing *in vivo*. Journal of Biological Chemistry *286*, 30253-30257.

Tanaka, N., and Shuman, S. (2011). RtcB is the RNA ligase component of an *Escherichia coli* RNA repair operon. Journal of Biological Chemistry *286*, 7727-7731.

Tanaka, N., Smith, P., and Shuman, S. (2010). Structure of the RNA 3'-phosphate cyclase-adenylate intermediate illuminates nucleotide specificity and covalent nucleotidyl transfer. Structure *18*, 449-457.

Taylor, M., Butler, R., Chambers, S., Casimiro, M., Badii, F., and Merrick, M. (1996). The RpoN-box motif of the RNA polymerase sigma factor sigma N plays a role in promoter recognition. Molecular Microbiology *22*, 1045-1054.

Temmel, H., Muller, C., Sauert, M., Vesper, O., Reiss, A., Popow, J., Martinez, J., and Moll, I. (2016). The RNA ligase RtcB reverses MazF-induced ribosome heterogeneity in *Escherichia coli*. Nucleic Acids Research, doi: 10.1093/nar/gkw1018.

Tintut, Y., and Gralla, J.D. (1995). PCR mutagenesis identifies a polymerase-binding sequence of sigma 54 that includes a sigma 70 homology region. Journal of Bacteriology *177*, 5818-5825.

Toro, N., Jimenez-Zurdo, J.I., and Garcia-Rodriguez, F.M. (2007). Bacterial group II introns: not just splicing. FEMS Microbiology Reviews *31*, 342-358.

Travers, A.A., and Burgess, R.R. (1969). Cyclic re-use of the RNA polymerase sigma factor. Nature *222*, 537-540.

Tree, J.J., Granneman, S., McAteer, S.P., Tollervey, D., and Gally, D.L. (2014). Identification of bacteriophage-encoded anti-sRNAs in pathogenic *Escherichia coli*. Molecular Cell *55*, 199-213.

Trujillo, M.A., Roux, D., Fueri, J.P., Samuel, D., Cailla, H.L., and Rickenberg, H.V. (1987). The occurrence of 2'-5' oligoadenylates in *Escherichia coli*. European Journal of Biochemistry *169*, 167-173.

Tucker, N., D'Autreaux, B., Spiro, S., and Dixon, R. (2005). DNA binding properties of the *Escherichia coli* nitric oxide sensor NorR: towards an understanding of the regulation of flavorubredoxin expression. Biochemical Society Transactions *33*, 181-183.

Untergasser, A. (2006). DNA miniprep - rapid boiling method. In http://www.untergasser.de/lab/protocols/miniprep_rapid_boiling_v1_0.htm.

Vesper, O., Amitai, S., Belitsky, M., Byrgazov, K., Kaberdina, A.C., Engelberg-Kulka, H., and Moll, I. (2011). Selective translation of leaderless mRNAs by specialized ribosomes generated by MazF in *Escherichia coli*. Cell *147*, 147-157.

Vidangos, N., Maris, A.E., Young, A., Hong, E.M., Pelton, J.G., Batchelor, J.D., and Wemmer, D.E. (2013). Structure, function, and tethering of DNA-binding domains in sigma 54 transcriptional activators. Biopolymers *99*, 1082-1096.

Walker, J.E., Saraste, M., Runswick, M.J., and Gay, N.J. (1982). Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. The EMBO Journal *1*, 945-951.

Walter, P., and Ron, D. (2011). The unfolded protein response: From stress pathway to homeostatic regulation. Science *334*, 1081.

Wang, J.T., and Gralla, J.D. (1996). The transcription initiation pathway of sigma 54 mutants that bypass the enhancer protein requirement - Implications for the mechanism of activation. Journal of Biological Chemistry *271*, 32707-32713.

Wang, L.K., and Shuman, S. (2005). Structure–function analysis of yeast tRNA ligase. Rna *11*, 966-975.

Wang, Y.Y., Hai, T., Liu, Z.C., Zhou, S.Y., Lv, Z., Ding, C.H., Liu, L., Niu, Y.Y., Zhao, X.Y., Tong, M., *et al.* (2010). HSPC117 deficiency in cloned embryos causes placental abnormality and fetal death. Biochemical and Biophysical Research Communications *397*, 407-412.

Wassem, R., de Souza, E.M., Yates, M.G., Pedrosa, F.D., and Buck, M. (2000). Two roles for integration host factor at an enhancer-dependent *nifA* promoter. Molecular Microbiology *35*, 756-764.

Weaver, F. (2012). Molecular biology, 5 edn (New York: McGraw-Hill).

Webb, S., Hector, R.D., Kudla, G., and Granneman, S. (2014). PAR-CLIP data indicate that Nrd1-Nab3-dependent transcription termination regulates expression of hundreds of protein coding genes in yeast. Genome Biol. *15*, doi: 10.1186/gb-2014-1115-1181-r1188.

Wigneshweraraj, S., Bose, D., Burrows, P.C., Joly, N., Schumacher, J., Rappas, M., Pape, T., Zhang, X., Stockley, P., Severinov, K., *et al.* (2008). Modus operandi of the bacterial RNA polymerase containing the sigma 54 promoter-specificity factor. Molecular Microbiology *68*, 538-546.

Winther, K.S., and Gerdes, K. (2009). Ectopic production of VapCs from Enterobacteria inhibits translation and trans-activates YoeB mRNA interferase. Molecular Microbiology *72*, 918-930.

Winther, K.S., and Gerdes, K. (2011). Enteric virulence associated protein VapC inhibits translation by cleavage of initiator tRNA. Proceedings of the National Academy of Sciences USA *108*, 7403-7407.

Winther, K.S., Tree, J.J., Tollervey, D., and Gerdes, K. (2016). VapCs of *Mycobacterium tuberculosis* cleave RNAs essential for translation. Nucleic Acids Research *44*, 9860-9871.

Wolin, S.L., Belair, C., Boccitto, M., Chen, X., Sim, S., Taylor, D.W., and Wang, H.-W. (2013). Non-coding Y RNAs as tethers and gates: Insights from bacteria. RNA Biology *10*, 1602-1608.

Wong, C.N., Tintut, Y., and Gralla, J.D. (1994). The domain-structure of sigma 54 as determined by analysis of a set of deletion mutants. Journal of Molecular Biology *236*, 81-90.

Wosten, M.M.S.M. (1998). Eubacterial sigma-factors. FEMS Microbiology Reviews *22*, 127-150.

Xiang, C., Wang, Y., Zhang, H., and Han, F. (2017). The role of endoplasmic reticulum stress in neurodegenerative disease. Apoptosis *22*, 1-26.

Xu, H., and Hoover, T.R. (2001). Transcriptional regulation at a distance in bacteria. Current Opinion in Microbiology *4*, 138-144.

Yamamoto, K., Hirao, K., Oshima, T., Aiba, H., Utsumi, R., and Ishihama, A. (2005). Functional characterization in vitro of all two-component signal transduction systems from *Escherichia coli*. Journal of Biological Chemistry *280*, 1448-1456.

Yang, X., and Lewis, P.J. (2010). The interaction between bacterial transcription factors and RNA polymerase during the transition from initiation to elongation. Transcription *1*, 66-69.

Yang, Y., Darbari, V.C., Zhang, N., Lu, D., Glyde, R., Wang, Y.P., Winkelman, J.T., Gourse, R.L., Murakami, K.S., Buck, M., *et al.* (2015). Structures of the RNA polymerase-sigma 54 reveal new and conserved regulatory strategies. Science *349*, 882-885.

Yoshihisa, T. (2014). Handling tRNA introns, archaeal way and eukaryotic way. Frontiers in Genetics *5*, 213.

Zhang, N., Darbari, V.C., Glyde, R., Zhang, X.D., and Buck, M. (2016). The bacterial enhancer-dependent RNA polymerase. Biochemical Journal *473*, 3741-3753.

Zhang, N., Joly, N., Burrows, P.C., Jovanovic, M., Wigneshweraraj, S.R., and Buck, M. (2009). The role of the conserved phenylalanine in the sigma 54-interacting GAFTGA motif of bacterial enhancer binding proteins. Nucleic Acids Research *37*, 5981-5992.

Zhang, Y., Zhang, J., Hoeflich, K.P., Ikura, M., Qing, G., and Inouye, M. (2003). MazF cleaves cellular mRNAs specifically at ACA to block protein synthesis in *Escherichia coli*. Molecular Cell *12*, 913-923.

Zhang, Y.L., Zhang, J.J., Hara, H., Kato, I., and Inouye, M. (2005). Insights into the mRNA cleavage mechanism by MazF, an mRNA interferase. Journal of Biological Chemistry *280*, 3143-3150.

Zhao, K., Liu, M., and Burgess, R.R. (2010). Promoter and regulon analysis of nitrogen assimilation factor, sigma 54, reveal alternative strategy for *E. coli* MG1655 flagellar biosynthesis. Nucleic Acids Research *38*, 1273-1283.

Zillmann, M., Gorovsky, M.A., and Phizicky, E.M. (1991). Conserved mechanism of tRNA splicing in eukaryotes. Molecular and Cellular Biology *11*, 5410-5416.

Zofallova, L., Guo, Y.H., and Gupta, R. (2000). Junction phosphate is derived from the precursor in the tRNA spliced by the archaeon *Haloferax volcanii* cell extract. Rna *6*, 1019-1030.

Zundel, M.A., Basturea, G.N., and Deutscher, M.P. (2009). Initiation of ribosome degradation during starvation in *Escherichia coli*. Rna *15*, 977-983.

Zuo, Y.H., and Steitz, T.A. (2015). Crystal structures of the *E. coli* transcription initiation complexes with a complete bubble. Molecular Cell *58*, 534-540.