Investigating RNA Polymerase Regulation in Antibiotic Producing *Streptomyces* Bacteria

By

Hayley Greenfield

Thesis submitted for the degree of Doctor of Philosophy

Biochemistry and Biomedicine
School of Life Sciences
University of Sussex
Falmer, UK

Mar 2022

Declaration

I hereby declare that this thesis has not been and will not be, submitted in whole or in part to another University for the award of any other degree.

Signature:……

University of Sussex

Hayley Greenfield

For the degree of Doctor of Philosophy

Investigating RNA Polymerase Regulation in Antibiotic Producing *Streptomyces* Bacteria

Summary

Organisms belonging to the *Streptomyces* genus produce a vast number of useful secondary metabolites, including over half of all clinically relevant antibiotics. Recent genome sequencing, however, has revealed large clusters of genes in *Streptomyces* that are not expressed under standard growth conditions and that hold potential for the discovery of novel antibiotics. Several studies have shown that certain RNA polymerase mutations can increase antibiotic production, as well as stimulate the synthesis of previously undetected secondary metabolites. However, the basis of this is not understood and little is known about how RNA polymerase levels are controlled in *Streptomyces*.

The analysis of RNA-seq data from *S. coelicolor* and *S. venezuelae* revealed that the *rpoBC* operon, encoding the large β and β' subunits of RNA polymerase, is subjected to a largely unstudied form of gene regulation known as reiterative transcription (RT) at the transcriptional start site. This is the first identified case of RT in *Streptomyces*, with this study focussing on further understanding this nucleotide based control of gene expression. Mutagenesis of this TSS was carried out revealing RT has a detrimental effect on the expression of *rpoBC*. Further use of reporter-based transcriptional fusions also revealed several regulatory elements are present within the 5' UTR region of *rpoB* also affecting expression of these large RNA polymerase subunits. This contribution to a greater understanding of how RNA polymerase levels are controlled in *Streptomyces*, might enable novel approaches towards the synthesis of novel inhibitory compounds and therefore contribute to overcoming the current antimicrobial resistance crisis.

Acknowledgments

Firstly, I would like to thank my supervisor, Dr Mark Paget for his support and guidance through this PhD project and for giving me the opportunity to develop my scientific skills.

This PhD is dedicated to my supervisor Dr Stephen Hare. He is the reason not only that this PhD was completed, but that I, myself made it to the end. Thank you Stephen for the endless early morning chats about science and life, and for always sitting with me through the tears (and being honest if my mascara was all the over the place!). You gave me so much of your time to discuss experiments with me, and I will never forget how you constantly reassured me of my scientific abilities and uplifted me when yet another *in vitro* transcription assay failed. Your passion and amazing brain, and your kindness and compassion for everyone who stepped through our lab doors forever amazed me, and I feel honoured I got to be supervised by you. I can't wait till we are reunited again for a trail run in the sky.

Thank you to everyone on corridor 2C for making these 4 years of my life the best they could be. Thank you to Frankie, for always making me laugh and the consistent supportive lab chats, they were sorely missed when you moved. Thank you to Ali (desk partner) and Daniel also for your support, I feel so lucky to have spent my time with such a lovely and grounding group of PhD students. Thank you to both Tracy Nissan and Simon Morley for always grounding me and providing supportive office chats (and excessive amounts of biscuits), and Dr Murat Eravci for assisting me with the running of Paget lab, and data analysis.

Thank you to all my friends for getting me through some of the toughest years of my life. To Ella (and moss) and Megan for always being down for a walk and coffee. To Alex and Jordan for the pudding dates, and to Charlotte for always being there for rant-runs and mcflurry dates, these times meant more than I could ever say. My further thanks are extended to my second family Chloe and Joanna, for always listening to my PhD rants. And thank you to my soulmate Em, you have been there through everything since we were little, and I will never be able to say how grateful I am to have had your support over these years.

Lastly, I would like to thank my parents and my brother Matthew for always being supportive, and always listening to my lab dramas, and putting up with a stressed me

at home (even if you did get baked goods out of it!). I will never be able to tell you how good you were for me during this time, and I know how lucky I am to have been surrounded by you all both during lockdown and over the course of my PhD.

Contents

Table of figures

List of tables

Abbreviations

aac(3)IV - apramycin resistance gene

aadA - Spectinomycin resistance gene

CFU – colony-forming unit

CTD – C- terminal domain

dsb – double stranded break

dsDNA – double stranded DNA

ECF – Extra-Cytoplasmic function

iNTP – initiating ribonucleotide

h – hour

(p) ppGpp – Guanosine tetraphosphate (pentaphophate)

nt - nucleotides

NTD – N-terminal domain

NTP- Nucleoside triphosphate

RNAP – RNA Polymerase

rRNA – ribosomal RNA

RPc – RNA Polymerase closed complex

RPo – RNA Polymerase open complex

RT- Reiterative transcription

ssDNA – single stranded DNA

StrepDB – The *Streptomyces* annotation server

YEME – Yeast Extract-Malt Extract medium

Chapter 1: Introduction

1.1. Overview

The gram-positive *Streptomyces* genus is well known for large genomes of a high GC content and the ability to produce a broad range of useful secondary metabolites. However, recent genome sequencing has revealed a far greater capacity for secondary metabolite production than previously imagined, although many potential biosynthetic genes appear not to be expressed under standard conditions. Several studies have shown that mutations in the key transcriptional enzyme RNA polymerase (RNAP) can activate these silent genes, allowing products such as antibiotics to be assayed (Hosaka *et al*., 2009; Ochi & Hosaka, 2013).

This study focusses on the control of the *rpoBC* operon, encoding the large β and β' subunits of the RNAP enzyme, to further understand control of this enzyme in *Streptomyces*, which might lead to new approaches to activate silent gene clusters.

1.1.1 The Actinobacteria phylum

Organisms belonging to the largest bacterial phylum, Actinobacteria, are characterised as Gram-positive bacteria containing genomic DNA of high GC-content, which ranges from ~51% to 71% for such organisms as *Corynebacterium* spp. and *Streptomyces spp.*, respectively (Ventura *et al.*, 2007). The phylum includes a range of morphologies including rod-coccoid, coccoid, or differentiated branched mycelial as seen with *Streptomyces* (Ul-Hassan and Wellington, 2009).

Actinobacteria are often isolated from the natural environment, and are well known for their importance for soil biodegradation, by production of extracellular enzymes into the environment, assisting with the decomposition of decaying plant, fungal and animal matter. Most therefore encode an extensive metabolism allowing them to adapt to environmental change and compete with other organisms (Stubbendieck, Vargas-Bautista and Straight, 2016; Behie *et al.*, 2017) through, for example, the production of chemically diverse antibiotics (Hoskisson and Fernández-Martínez, 2018).

However, the Actinobacteria phylum also includes pathogenic and commensal organisms. *Mycobacteria* spp. and *Nocardia* spp., include organisms known to cause tuberculosis (*M. tuberculosis*) (Delogu, Sali and Fadda, 2013) and leprosy (*M. leprae*) (Sasaki *et al.*, 2001), and nocardiosis, respectively (Karam and Siadati, 2021; Yadav *et al.*, 2021). Bifidobacteria, also belonging to this phylum is a well-known intestinal

commensal that contributes several probiotic features (Lievin *et al.*, 2000; Barka *et al.*, 2016).

1.1.2 *Streptomyces* spp.

Within the Actinobacteria phylum, the *Streptomyces* genus is the largest, with members ubiquitous in nature, colonising both terrestrial and aquatic environments (Takizawa, Colwell and Hill, 1993; Lee *et al.*, 2014; Rashad *et al.*, 2015). The genus has provided a vast number of secondary metabolites including many clinically important antibiotics, and was the source of most key compounds during the "Golden Age" of antimicrobial discovery (Hodgson, 2000; Bentley *et al.*, 2002).

1.1.3 *Streptomyces coelicolor* A3 (2)

Streptomyces coelicolor A3 (2) was developed as a genetic model in the 1960s and now represents one of the most well studied organisms within the *Streptomyces* genus (Hopwood, 1999). Key to the decision to develop this strain as a genetic model is its production of two pigmented secondary metabolites, allowing the simple isolation of mutants (see below).

1.1.3.1 Genome of *S. coelicolor*

The *S. coelicolor* A3 (2) genome is a single linear chromosome, consisting of 8,667,507 bp of high GC content (72.12%) with 7,825 predicted coding sequences (Bentley *et al.*, 2002). The origin of replication (*oriC*) is in the centre of the chromosome, with replication initiating here bidirectionally. Essential genes, such as those encoding for transcription, translation, DNA replication, cell division and amino acid biosynthesis, tend to be located towards the centre of the genome, with genes associated with non-essential functions, including secondary metabolite production tending to be located at the distal arms of the chromosome (Bentley *et al.*, 2002).

The main model strain, M145, is a prototrophic derivative of the A3 (2) strain that has had its two plasmids SCP1 and SCP2 removed; these low-copy plasmids are in a linear and circular form, and consist of 365 kb and 31 kb, respectively (Kinashi and Shimaji-Murayama, 1991; Haug *et al.*, 2003).

More recently, the *Streptomyces venezuelae* organism has been proposed a newer model system, due to its ability to sporulate in liquid culture, not apparent for growth of *S. coelicolor*, allowing more representative analysis of the cellular differentiation within this genus (Bibb *et al.*, 2012; Bush *et al.*, 2013).

1.1.3.2. Life cycle of *S. coelicolor*

Most streptomycetes, including *S. coelicolor*, have a distinctive and complex life cycle when compared to most bacteria with growth more closely resembling that of filamentous fungi. They grow as a highly branched vegetative mycelium on solid medium, eventually leading to the formation of elevated aerial hyphae, upwards from the surface and differentiation into reproductive spores that can be dispersed in the local environment (Bentley *et al.*, 2002; Sigle *et al.*, 2015). A full cycle of growth is usually 3 to 10 days in *Streptomyces* (Fig 1.1.3.2.1) (Reviewed by Flärdh *et al.*, 2012).

Upon the detection of an appropriate source of nutrients, a previously dormant *S. coelicolor* spore is activated for growth by germination involving the formation of 1 or 2 germ tubes, that extend from the spore (Fig 1.1.3.2.1B) (Flärdh and Buttner, 2009). Further growth occurs by hyphal tip extension and branching from these germ tubes forming *Streptomyces* network of vegetative mycelium (Fig 1.1.3.2C). Upon the depletion of nutrients in the surrounding environment, aerial hyphae are produced from the vegetative mycelium, extending upwards into the air (Figs 1.1.3.2.1 D and E). A characteristic of this hyphae is its fuzzy appearance which represents a hydrophobic sheath caused by the production of the surfactant peptide SapB. Several other genes required for the formation of aerial hyphae are usually designated *bld* (bald) genes, due to their bald appearance that results when these genes are mutated, causing an inability to erect aerial hyphae (Flärdh and Buttner, 2009). Aerial hyphae grow to initially form a long non-septated sporagenic (multinucleoidal) cell containing >50 copies of the genome. Before septation takes place, growth is arrested before multiple rounds of cell division takes place producing uninucleoidal pre-spores that then further progress into mature grey pigmented spores (Schwedock *et al.*, 1997).

Streptomyces are often isolated from the natural environment, with the production of spores enabling them to survive harsh conditions where there is little nutrient and water availability, as well as acting as a mechanism of dispersal. The sporulation

process requires several proteins, with mutations in the corresponding *whi* (white) genes causing strains to lack both the ability to produce mature spores and their characteristic grey pigment, resulting in the production of white colonies. Several proteins are involved in growth-phase transitions in *Streptomyces*, and are discussed further in appendix Section 7.8.

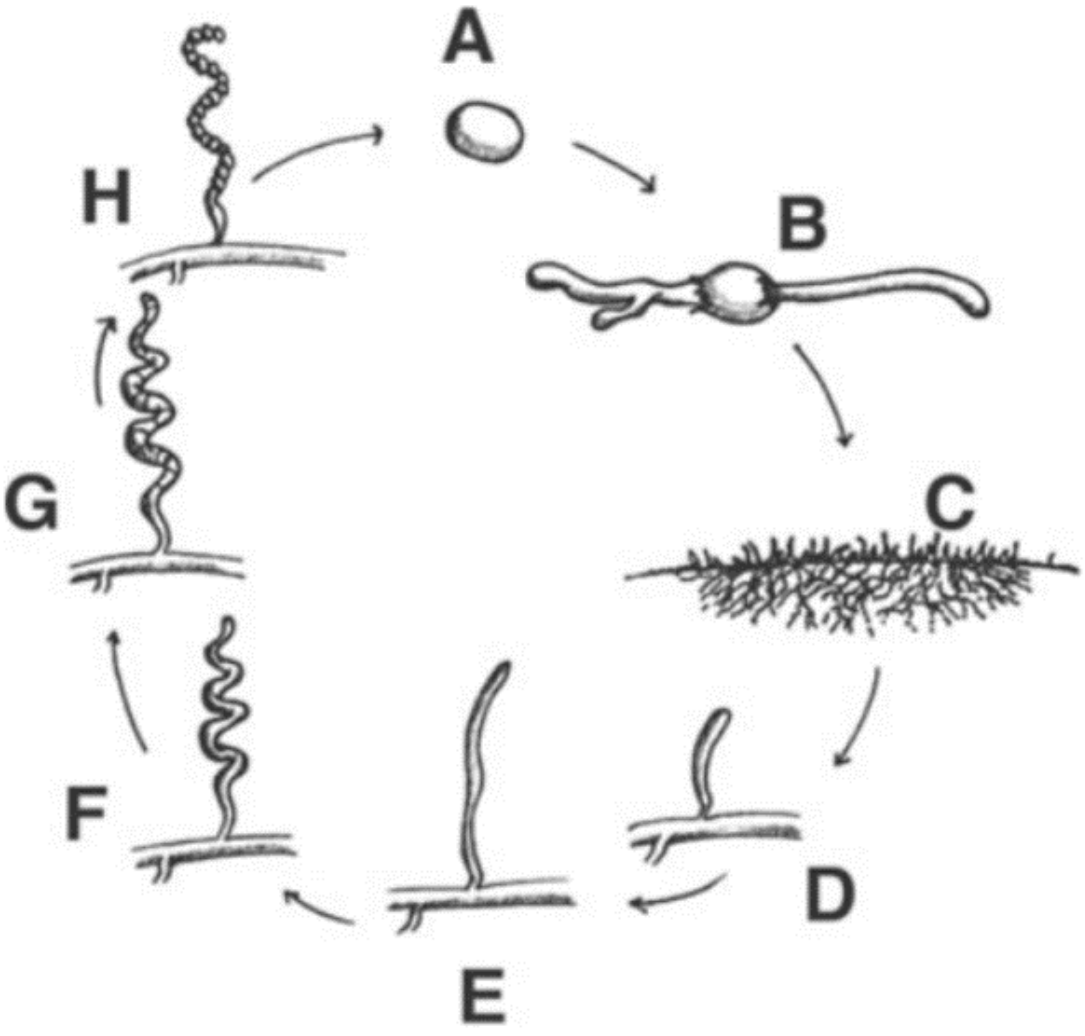


Figure 1.1.3.2.1: The life cycle of *Streptomyces coelicolor.* The initial spore (A) is germinated leading to germ-tube formation on solid medium (B). Substrate mycelium and aerial hyphae are then produced from these germ tubes (C). D and E present a close up of the aerial hyphae, and the multigenomic hyphae lacking septa, respectively, with F and G showing the development of the aerial hyphae into a multigenomic (sporagenic) cell and unigenomic pre-mature spores, respectively, before transition into mature spores seen in H. Copyright licence number: 5197831030574.

1.1.3.3 Antibiotic production and regulation in *S. coelicolor*

Streptomyces as a genus are well known to produce a large number of useful secondary metabolites which includes ~80% of the antibiotics used today, as well as antifungals, antihelminthics, antivirals and immunosuppressants (Watve *et al.*, 2001; de Lima Procópio *et al.*, 2012; Hoskisson and Fernández-Martínez, 2018).

S. coelicolor is itself known to encode 22 biosynthetic gene clusters associated with secondary metabolism that contain the genes necessary for the production of antibiotics, as well as other molecules such as siderophores, terpenoids and lipids (Bentley *et al.*, 2002). This includes 5 structurally distinct antibiotics: two pigmented antibiotics, the polyketide Actinorhodin (ACT) (Wright and Hopwood, 1976) and tripyrrole undecylprodigiosin (RED) (Rudd and Hopwood, 1980) that produce blue and red pigments, respectively, as well as the cryptic polyketide coelimycin P1 (CPK) (Pawlik, Kotowska and Kolesiński, 2010), the non-ribosomal lipopeptide calcium-dependent antibiotic (CDA) (Hopwood and Wright, 1983) and the cyclopentanone methylenomycin (MM) which is encoded on the linear SCP1 plasmid (Fig 1.1.3.3) (Wright and Hopwood, 1976).

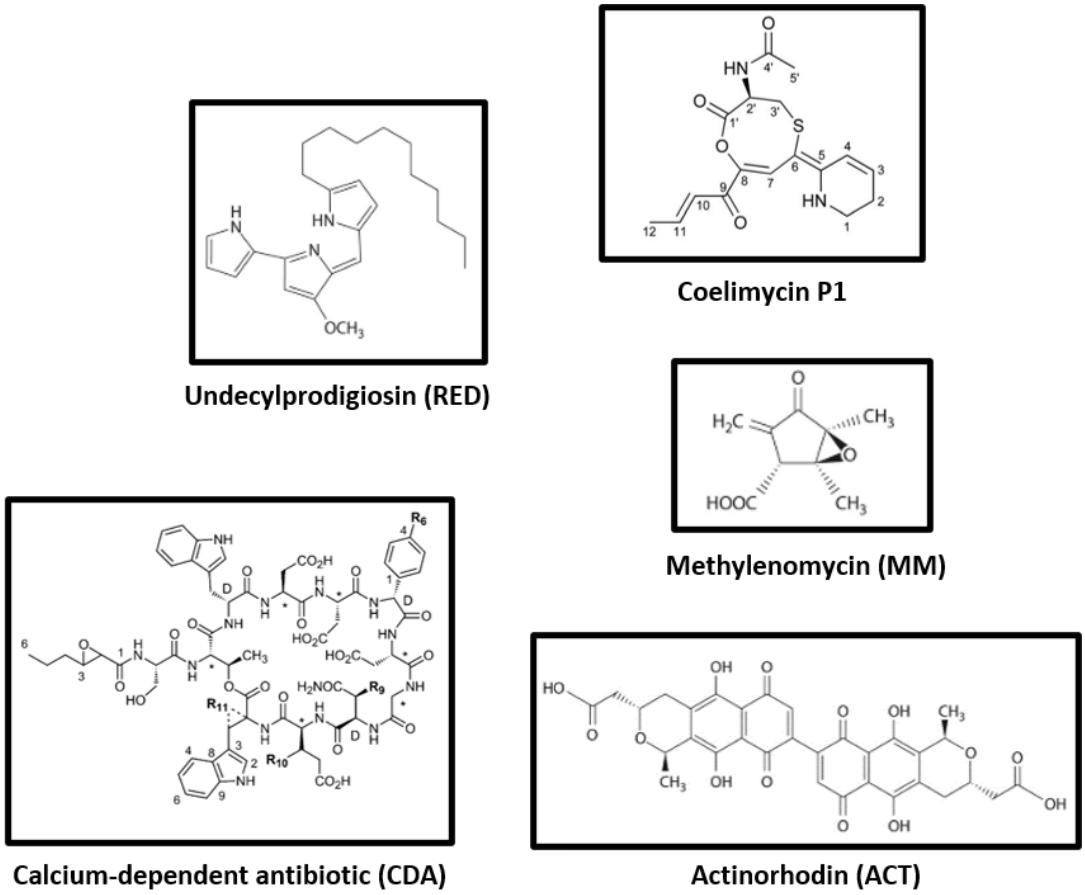


Figure 1.1.3.3: The structure of the five antibiotic compounds produced by *S. coelicolor*. Structures of undecylprodigiosin (RED), coelimycin P1, calcium-dependent antibiotic (CDA), methylenomycin (MM) and actinorhodin (ACT). (Figs extracted from Gomez-Escribano *et al.*, 2012; Hojati *et al.*, 2002; Liu *et al.*, 2013)**.**

Genes required for the synthesis of secondary metabolites tend to be organised into biosynthetic gene clusters. For example, the production of the pigmented antibiotics, ACT and RED, is reliant on genes encoded by clusters SCO5071-5092 and SCO5877-5898 in the *S. coelicolor* genome, respectively (Bentley *et al.*, 2002). The production of most secondary metabolites involves many steps, involving multiple intermediates, and modifying enzymes that are required to produce a final compound conferring biological activity (Hodgson, 2000). For example, the highly studied *act* cluster contains 20 ORFs and at least 6 transcripts (Fernández-Moreno *et al.*, 1991; Aceti and Champness, 1998).

The expression of most biosynthetic gene clusters is under tight regulation to ensure that expression only occurs under appropriate conditions. Most clusters have at least one cluster-situated regulator (CSR) that controls, positively or negatively,

transcription of the biosynthetic genes. Examples include actII-ORF (Arias, Fernández-Moreno and Malpartida, 1999), *cdaR* (Hojati *et al.*, 2002), redD (Narva and Feitelson, 1990) and redZ (Guthrie *et al.*, 1998), which are cluster-linked and activate the production of ACT, CDA and RED antibiotics, respectively. The CSRs themselves are often targeted as a single point of regulation, allowing expression control of the biosynthetic cluster in accordance with changing intracellular conditions. For example, the transcription factor, actII-ORF4, is controlled by multiple regulatory proteins including the two repressors, DasR and AbsA2(Rigali *et al.*, 2008; Lewis *et al.*, 2019). AbsA2 is part of the *absA* operon which also encodes *absA1* and *absA2* which together form a two-component system. AbsA1, a sensor kinase, phosphorylates the AbsA2 response regulator causing it to repress directly actII-ORF4 transcription, with null mutations of either AbsA1 or AbsA2 increasing ACT antibiotic production in *S. coelicolor* (Sheeler, MacMillan and Nodwell, 2005). AbsA2 is also seen to exert a negative effect on the transcription of RED and CDA also, displaying that the *absA* regulon also plays a global role in the regulation of antibiotic production affecting multiple pathway-specific activators (Aceti and Champness, 1998); AbsA2 also binds to the promoter regions of *redZ* and within multiple sites within the CDA gene cluster, including the *cdaR* promoter. (McKenzie and Nodwell, 2007; Lewis *et al.*, 2019).

The DasR global regulator, a GntR-like repressor, also works by binding to promoter regions of all CSRs in *S. coelicolor* (reviewed by Urem and colleagues (2016)). The DNA binding activity of DasR appears to be modulated by N-acetylglucosamine (GlcNAc) metabolism. Upon entry into the cytoplasm, GlcNAc is both phosphorylated and deacetylated to produce an allosteric inhibitor of DasR, GlcN-6P, further preventing the repression of actII-ORF4 by DasR (Fillenberg *et al.*, 2016). This presents an example of nutritional control of antibiotic synthesis, in this case by GlcNAc, that acts as both a nitrogen and carbon source, and whose presence stimulates ACT synthesis.

Other nutrient signals also influence the extent of antibiotic production. (p)ppGpp, discussed below (Chapter 1.2.4.2) are alarmones that are produced when amino acids are limiting in the cell, causing a stringent response, a stress response that appears to be almost ubiquitous in bacteria (Irving, Choudhury and Corrigan, 2021). In *Streptomyces*, the accumulation of (p)ppGpp causes a significant decrease in the

expression of genes associated growth, such as rRNA synthesis, with a simultaneous increase in genes associated with stress responses and secondary metabolism (Hesketh *et al.*, 2007a). In particular, increased (p)ppGpp levels cause the over production of both ACT and RED in *S. coelicolor* (Hesketh *et al.*, 2001; Kang *et al.*, 1998). However, the role of (p)ppGpp in antibiotic production varies between clusters and strains – for example a Δ*relA* strain of *S. clavuligerus*, which is deficient in its ability to produce (p)ppGpp, overproduces both cephamycin C and clavulanic acid (Gomez-Escribano *et al.*, 2008). However, the exact mechanism of (p)ppGpp action on biosynthetic gene clusters, and wider transcription patterns in the Actinobacteria is not well understood (Liu *et al.*, 2013).

1.1.3.3.1 Pleiotropic activation of silent biosynthetic gene clusters

As mentioned above, most biosynthetic gene clusters appear to not be expressed under standard laboratory conditions (Hoskisson and Fernández-Martínez, 2018). This is likely due to the large amount of genetic control that is invested into regulating these secondary metabolic clusters, to ensure the organism only produces them when essentially necessary. However, an intriguing method of inducing such silent clusters has been to generate antibiotic resistant mutants in both ribosomal proteins and RNAP coding genes (Hosaka *et al.*, 2009; Ochi and Hosaka, 2013). Of particular interest to this project are mutations in the *rpoB* gene, encoding the β-subunit of RNAP, which increased actinorhodin production 1.5 –9 fold compared to the M145 parent strain, and also triggered the production of a previously uncharacterised antibacterial class, Piperidamycin, from another *Streptomyces* strain (Gomez-Escribano & Bibb, 2011). This antibiotic was produced from a *Streptomyces mauvecolor* strain containing various point mutations in *rpoB* that conferred rifampicin resistance (Hosaka *et al.*, 2009). It was proposed that this increased production of piperidamycin was due to an increased affinity of mutant RNAP for piperidamycin biosynthetic promoters (Hosaka *et al.*, 2009).

1.2 Transcription initiation in bacteria

1.2.1 DNA dependent RNA polymerase

RNA polymerase (RNAP) is the multi-subunit enzyme (~400 kDa) involved in the transcription of genes encoded in the template DNA, to RNA, making it one of the main points of regulation in the control of global gene expression. The core RNAP capable of transcription elongation, is comprised of five subunits: ββ'2αω (Fig 1.2.1.1) (Archambault and Friesen, 1993, Saecker *et al.*, 2010). The first X-ray crystallographic models of RNAP from *Thermus aquaticus* revealed a claw-like structure with an active site that allows for the binding to template DNA, and production of an RNA product complementary to the DNA sequence being transcribed (Darst, Kubalek and Kornberg, 1989; Zhang *et al.*, 1999; Murakami, 2013; Sutherland and Murakami, 2018).

In all bacteria the large β and β' subunits are encoded by the *rpoBC* operon and comprise the majority of each arm of the claw-like structure, with a central 27 Å wide channel, and have extensive interactions with each other (Zhang *et al.*, 1999). For the catalytic activity of this enzyme, a Mg^{2+} ion is required to bind to the back wall of the channel internal to the large subunits (Zhang *et al.*, 1999). The α-subunits, encoded by the *rpoA* gene and consisting of an N-Terminal Domain (α-NTD) and C-Terminal Domain (α-CTD) connected by a flexible linker, are important for initial assembly of the core RNAP and play a key role in regulation, but are not involved in enzyme catalysis. Dimerization of the α subunits promotes further assembly with the β and β' subunits through interaction with the α-NTD (Zhang *et al.*, 1999; Gourse, Ross and Gaal, 2000). The α-CTD is important for interaction with upstream DNA elements, as well as being a target for several DNA-binding transcription factors that act upstream of the core promoter elements (Ross *et al.*, 1993; Dove, Joung and Hochschild, 1997; Hochschild and Dove, 1998). It believed that the limiting factor in the formation of the core enzyme is the concentration of the β and β' subunits, as the α-subunit was observed to be in >2-fold excess when compared to the concentration of the large subunits (Dykxhoorn, St. Pierre and Linn, 1996).

The role of the smallest RNAP subunit, ω, encoded by *rpoZ*, is the least understood. This subunit is seen to be located at the surface of the holoenzyme (Fig 1.2.1.1), in

contact with the C-terminal tail of the β' subunit only. It is thought to play a role as a chaperone in later stages of RNAP assembly and is likely accessible by transcriptional factors (Burgess, 1969; Dove and Hochschild, 1998). It also plays a direct role in (p)ppGpp binding and the stringent response in *E. coli* and other proteobacteria (Section 1.2.4.2) (Mathew and Chatterji, 2006; Ross *et al.*, 2013). Unlike the other core subunits, ω is non-essential for growth in many bacteria including *E. coli* and *Streptomyces* (Ghosh, Ishihama and Chatterji, 2001; Santos-Beneit *et al.*, 2011; Kurkela *et al.*, 2021). However, in *Mycobacterium tuberculosis*, ω is crucial for viability, probably through its role in the assembly of the RNAP (Mao *et al.*, 2018).

For successful transcription initiation upstream of specific genes, a sixth dissociable sigma (σ) factor is required to bind to core RNAP, forming the RNAP holoenzyme, directing it to promoter regions that lie upstream of genes.

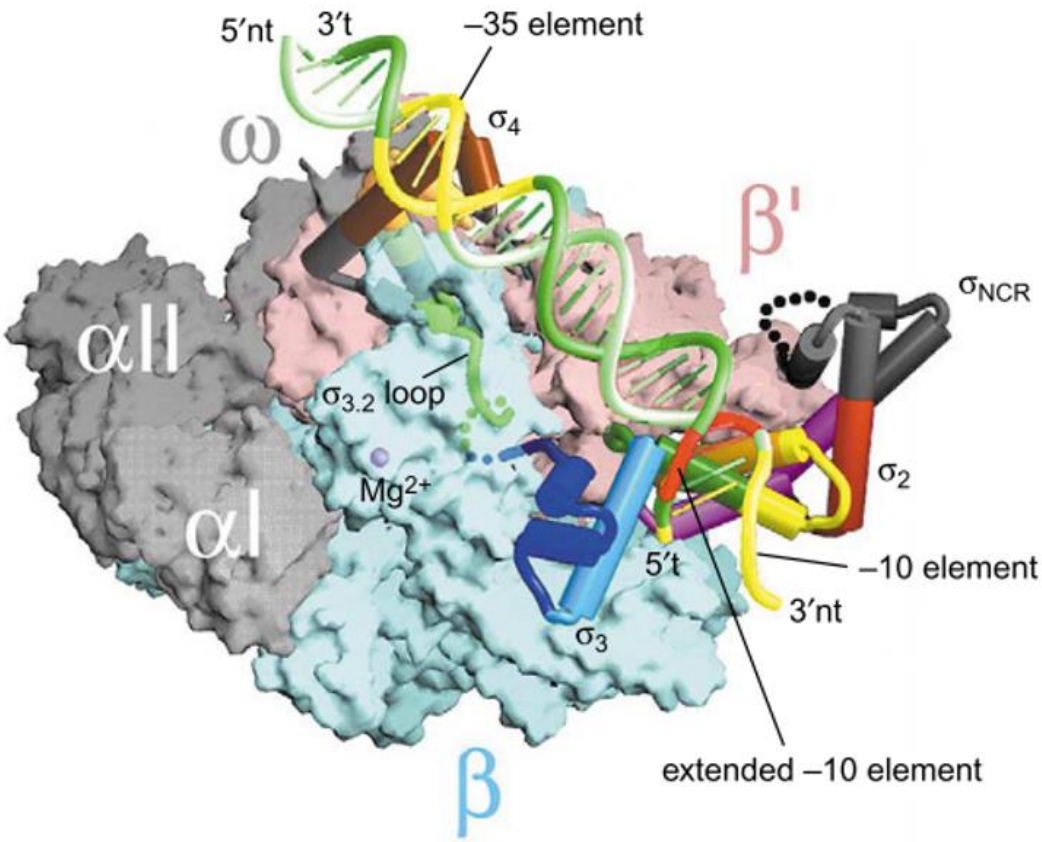


Figure 1.2.1.1: The structure of the RNA polymerase holoenzyme bound to promoter DNA. The molecular surface of the RNAP enzyme is shown surrounding the core promoter elements -35 and -10, which are indicated in yellow, on the DNA helix in the figure. The template and non-template strand of the DNA helix can be identified as dark green and light green, respectively with the extended -10 shown in red. The individual RNAP subunits are also indicated; β and β' are in blue and pink, respectively, with the two α subunits (αI and αII) and the ω subunit indicated in grey. The bound σ factor can be identified in multiple segments; the σ helices are shown as cylinders with differing conserved regions identified in different colours (σ_2 (red) σ_3 (light blue) σ_4 (brown)). The $\sigma_{3.2}$ loop is identified in pale green and the non-conserved (NCR) σ-region is identified as black cylinders (see Section 1.2.2 for σ-factor structure). Fig adapted from Murakami & Darst, 2003, copyright license number: 5193821173374.

1.2.2 Structure and function of sigma factors

The sixth RNAP subunit, known as a sigma factor, is essential for transcription initiation through its interaction with specific conserved promoter regions (Burgess *et al.*, 1969). There are two major classes of σ factor - the σ^{70} family, named after the *E. coli* principal sigma factor σ^{70}, and the σ^{54} family, named after *E. coli* σ^{54} which is involved in nitrogen metabolism and stress responses (Hirschman *et al.*, 1985; Buck and Cannon, 1992; Lonetto, Gribskov and Gross, 1992; Zhang and Buck, 2015) . However, since no examples of the σ^{54} class have been discovered in Actinobacteria, this class will not be discussed further.

Members of the σ^{70} family are responsible for interacting directly with the -35 and -10 promoter elements, as well as the extended -10 promoter elements (Fig 1.2.2.1), however they play several additional key roles in transcription initiation. As well as recruiting RNAP to specific promoter regions, σ factors facilitate DNA melting and the formation of the "transcription bubble" near the TSS producing an initial open complex (RPo) as well as contributing to de novo RNA synthesis at the active site. All bacteria contain an essential primary σ factor responsible for initiating transcription at most promoters (σ^{70} in *E. coli*), but also additional "alternative" σ factors that redirect RNAP to other sets of promoters: for example *E. coli* and *S. coelicolor* have 6 and 64 alternative sigma factors, respectively (Bentley *et al.*, 2002; Gruber and Gross, 2003; Tripathi, Zhang and Lin, 2014).

Sigma factors belonging to the σ^{70} family can further be phylogenetically divided into four groups (I-IV), which can reflect differences in function, with variable presence/absence of four domains within the whole sigma factor structure; these include domains $\sigma_{1.1}$, σ_2, σ_3 and σ_4 which can constitute the sigma factor and within them contain specific conserved regions identified as σR1.1, σR1.2-2.4, σR3.0-3.2 and σR4.1-4.2, respectively (see Fig 1.2.2.1). The principal and essential sigma factors comprise Group I, which contain all four domains and are structurally similar to σ^{70}. Group II are similar in structure to principal sigma factors, however, lack the $\sigma_{1.1}$ domain and are not essential for growth (Paget, 2015). Group III again usually lack the $\sigma_{1.1}$ region, however, also lack the highlighted non-conserved region and are more diverse in their role, with these sigma factors responsible for changes in gene expression during stress response, such as heat shock, as well as having involvement

in sporulation and flagella synthesis (Paget and Helmann, 2003). The conserved promoter elements required for interactions with the group III sigma factors also differ from those in groups I and II (Paget and Helmann, 2003). Group IV, also designated Extra-Cytoplasmic Function (ECF) sigma factors, contains the most diverse and largest number of sigma factors (Lonetto *et al.*, 1994; Helmann, 2002). These include those involved in stress responses, usually in response to environmental/ extracellular signals, and contain only the σ_2 and σ_4 domains as well as the equivalent of the 3.2 linker (Fig 1.2.2.1).

The presence of the domains determines interactions with the promoter DNA, thus explaining the differing activity at promoters of the sigma factor groups. The -10 and -35 promoter elements for σ^{70} which are centred roughly -10 and -35 nucleotides away from TSS, have consensus sequence for binding of TATAAT and TTGACA, respectively. In *S. coelicolor,* the conserved regions differed only slightly as TANNNT and NTGACC for the -10 and -35 elements, respectively (N can be replaced by any nucleotide) (Jeong *et al.*, 2016a). Some promoters regions also contain an extended -10 element (TGNTATAAT) , that can allow sufficient transcription initiation without the presence of a -35 element (Barne, Bown and Minchin, 1997).

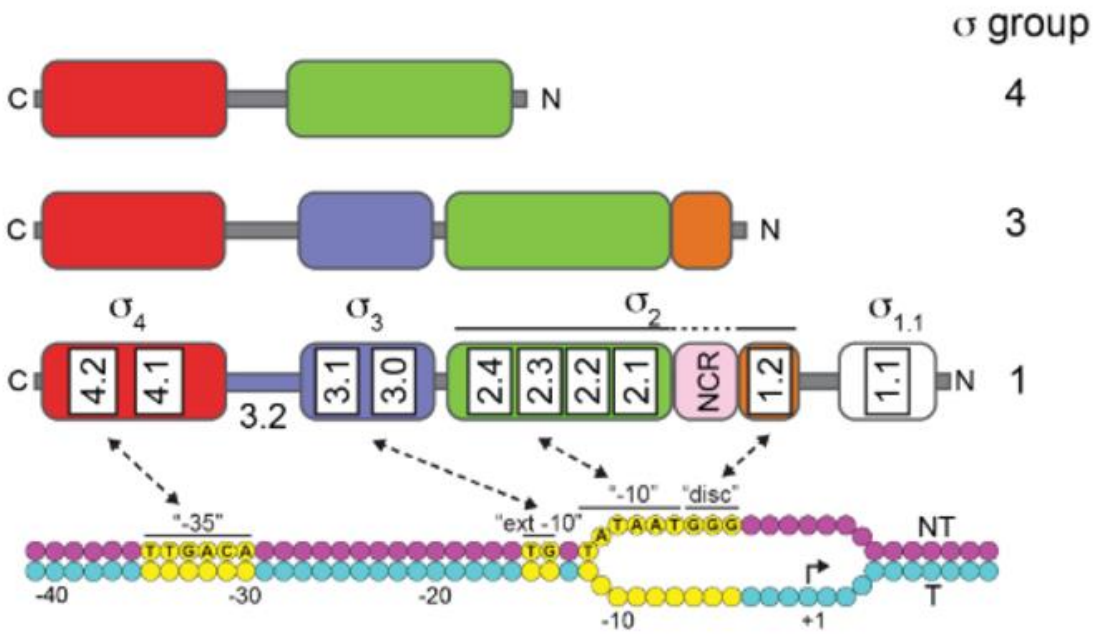


Figure 1.2.2.1: The structure, organisation and conserved regions of the differing σ^{70} family. The conserved regions are noted within the 4 domains observed for σ group 1; $\sigma_{1.1}$, σ_2, σ_3 and σ_4 are shown in white, green, blue and red, respectively. The interactions of the differing σ domains with the non-template (NT) promoter DNA are shown. T represents the template strand and NCR is the non-conserved region located within the 1.2 and 2.1 domains. Note that the template strand contains the consensus sequence for *E. coli*. Figure edited and extracted from M. S. Paget, 2015. The Generic License attributed to Mark Paget and can be accessed in Biomolecules, 2015, 5(3), 1245-1265). This figure is licensed under a Creative Commons Attribution (CC BY).

1.2.2.1 The domains of the principal sigma factors

The presence of the differing sigma factor domains, determines the function and further the classified group these sigma factors belong to. These domains further determine interactions with the DNA template, thus providing specificity during transcription initiation.

The $\sigma_{1.1}$ region

The $\sigma_{1.1}$ region, present at the NTD and only in group I principal sigma factors, consists of three helices that make up the $\sigma_{1.1}$ region, which is connected to $\sigma_{1.2}$ by a flexible linker (Fig 1.2.2.1). This region functions to prevent the lone binding of these sigma factors to promoter elements, ensuring contacts are only made when the sigma factor is bound to the core enzyme (Schwartz *et al.*, 2005). This inhibition is due to the negative charge of the $\sigma_{1.1}$ region, which affects binding of the σ_2 and σ_4 to

promoter DNA, when the sigma factor is not bound to the RNAP core (Dombroski, Walter and Gross, 1993); $\sigma_{1.1}$ has been shown in *E. coli* structures to interact with $\sigma_{4.2,}$ in free state (not RNAP bound), preventing lone binding of promoter DNA (Dombroski, Walter and Gross, 1993; Zachrdla *et al.*, 2017). $\sigma_{1.1}$ also plays an additional role in the transcription initiation by assisting in the isomerisation of RNAP closed complex (RPc) to open (RPo) (Vuthoori *et al.*, 2001; Bae *et al.*, 2013).

The σ_2 domain

The σ_2 domain is the most conserved and makes extensive contacts with the core RNAP through interactions with the β' subunit (Young *et al.*, 2001; Zachrdla *et al.*, 2017). The domain is crucial for the recognition and binding of the -10 consensus sequence ($_{-12}$TATAAT$_{-7}$), recognising these bases on single-stranded non-template DNA. During DNA melting, conserved bases at -7 (T) and -11 (A) flip out of the double helix and into hydrophilic and hydrophobic pockets in $\sigma_{2,}$ respectively, which stabilizes the open complex (Feklistov and Darst, 2011; Zhang *et al.*, 2012). This DNA melting further facilitates the growth of the transcription bubble, and causes a 90° bend in the -10 promoter DNA which causes downstream DNA to be inserted into the cleft that represents the active site of the RNAP, allowing transcription to begin (Bae *et al.*, 2013; Feklistov, 2013).

The σ_3 domain

The σ_3 domain, which spans regions 3.0-3.2, is composed of three helices and is important for the stabilisation of RPo by interacting with DNA upstream of the -10 element. In particular the $\sigma_{3.0}$ region is able to bind in major groove of extended -10 elements ($T_{-15}G_{-14}$ in *E. coli*, discussed above), which removes the need for a -35 element for binding and initiation of transcription (Barne, Bown and Minchin, 1997; Mitchell *et al.*, 2003).

The $\sigma_{3.2}$ region is the flexible linker region between domains 2 and 4, and includes the σ finger, which penetrates the active site interacting directly with template DNA upstream, with a possible role in correctly positioning the template DNA in the active site ready for RNA synthesis (Oguienko *et al.*, 2021). Interaction between σ finger and nascent RNA is thought to also be crucial for the release of the sigma factor from RNAP, allowing promoter escape and transition into transcription elongation; synthesis of RNA extends towards and clashes with the σ finger, causing either

abortive transcription or pausing and release of sigma factor (Pupov *et al.*, 2014; Petushkov *et al.*, 2017; Li *et al.*, 2020). The σ finger is particularly important for transcription activation with alternative ECF sigma factors, with mutations in this region causing decreased transcription due to impaired promoter escape (Oguienko *et al.*, 2021).

The σ_4 domain

The σ_4 domain includes a classical helix-turn-helix DNA motif that binds to the -35 element, causing a bend (30°) in the DNA and placing the upstream DNA in closer proximity to the holoenzyme, enhancing the interactions between the two (Murakami and Darst, 2003; Murakami, 2013). This region is also a key contact point for many transcriptional activators that bind just upstream or overlapping the -35 element (Dove, Darst and Hochschild, 2003).

1.2.2.2 Sigma factors in *S. coelicolor*

As discussed above, *S. coelicolor* is known to encode 65 sigma factors, all belonging to the σ^{70} family (Bentley *et al.*, 2002). The principal sigma factor is σ^{HrdB} (Buttner, Chater and Bibb, 1990), and although there are three related sigma factors (Class 2 σ^{HrdA}, σ^{HrdC} and σ^{HrdD}) only $\sigma^{HrdB\ is}$ essential for growth (Buttner, Chater and Bibb, 1990).

Aside from these, 10 sigma factors are related to the σ^B stress response factor found in *B. subtilis*, which includes σ^F , σ^B and σ^{WhiG} that are responsible for spore maturation, morphological development and sporulation, respectively (Kelemen *et al.*, 1996; Tan *et al.*, 1998; Cho *et al.*, 2001; Paget and Helmann, 2003).

The highest number of sigma factors belong to the ECF sigma factor family, with 45 identifying within this group. This includes σ^R, that redirects transcription during oxidative stress, increasing the expression of a disulphide reductase system to overcome and resist the stress (Paget, 1998). For example, σ^R specifically activates promoters of the *trxBA* operon that encodes thioredoxin reductase and thioredoxin. The defined regulon for promoters under control of σ^R was determined, which consisted of 108 genes, including the HspR regulon, ribosomal associated genes, *hrdB* and *relA* (Section 1.2.5.2) (Kallifidas *et al.*, 2010; Kim *et al.*, 2012). The activity of σ^R itself is controlled by its anti-sigma factor RsrA, which is encoded downstream and

co-transcribed with *sigR* (Kang *et al.*, 1999). RsrA in a reduced state, binds and inactivates σ^R. However, upon the exposure of an oxidising agent, such as diamide, RsrA forms an intramolecular disulphide bond between its redox-sensitive cysteine residues, resulting in the release of σ^R (Paget, 1998; Paget *et al.*, 2001).

Interestingly, σ^{shbA}, another ECF sigma factor that is highly conserved within *Streptomyces*, was isolated from *S. griseus,* and determined to be, like σ^R, also in control of transcription of *hrdB* in *S. griseus* (Otani *et al.*, 2013). Upon deletion of σ^{shbA} a large amount of housekeeping genes were downregulated, which resulted in severe growth defects, likely due to the decreased transcription of σ^{HrdB}. The consensus sequences for the -35 and -10 elements for activation by σ^{ShbA} were determined as CGTAAC and CGATGA, respectively (Otani *et al.*, 2013).

1.2.3 Transcription in bacteria.

1.2.3.1 Initiation

After the correct assembly of the holoenzyme, the enzyme can bind and initiate transcription at specific promoters in accordance with the sigma factor that is bound to the core RNAP. This is the first stage of the transcription cycle and consists of several stages from promoter recognition and binding, through to escape of the RNAP from the promoter, and transition into the elongation phase of transcription. Regulation of transcription largely occurs at this stage, due to the impact of not only sigma factors that can recognise different promoters, but also additional transcription factors that can bind to DNA and either activate or repress transcription from specific promoters (Browning and Busby, 2004; Paget, 2015; Jensen and Galburt, 2021).

Using footprinting and cryo-EM methods, the first interactions were shown to involve the αCTD of RNAP and DNA upstream from the -35 element, with further interactions then made between sigma regions 4.2 and 2.3 and the specific -35 and -10 promoter elements, respectively (Feklistov and Darst, 2011; Saecker, and DeHaseth, 2011; Zhang *et al.*, 2012; Murakami, 2013, Narayanan *et al.*, 2018). Upon the binding of the holoenzyme to its specific promoter DNA, an initial closed complex is formed (RPc).

At this point the DNA lies on the outer face of one side of the RNAP, outside of the active site channel (Murakami and Darst, 2003).

After this has occurred, several conformational changes, involving several unstable intermediates occurs both in the DNA and enzyme, resulting in the DNA around the TSS opening, forming the open complex (RPo). The first unstable intermediate (I_1, intermediate 1) consists again of a closed complex (RPc) however, a 90° bend in the DNA at the upstream end of the -10 element results in downstream dsDNA (~+20-25) being inserted into the active site cleft (Saecker *et al.*, 2002; Plaskon *et al.*, 2021). The I_1 intermediate itself is further stabilised by RNAP interactions with DNA upstream of -35 hexamer, made initially via the contacts made by the α-CTD domain (see above) (Sclavi *et al.*, 2005; Saecker, Record Jr and DeHaseth, 2011). Duplex DNA (from -11 - +20) is thought to occupy the active-site cleft in RPo, with the opening of the DNA from -11 - +2 occurring in one kinetic step to form the second intermediate (I_2), placing the +1 nucleotide at the RNAP active site (Davis *et al.*, 2007; Saecker and DeHaseth, 2011). This transition of I_1 to I_2 is highly dependent on temperature (Saecker *et al.*, 2002).

A final intermediate I_3 is made before the final transition to the RPo complex, which results from the folding and assembly of β and β' mobile regions into a clamp structure, bound to downstream dsDNA, covering residues ~ +10 to +20 (Hudson *et al.*, 2009; Opalka *et al.*, 2010; Plaskon *et al.*, 2021). The final RPo conformation results from the tightening of this jaw and a stabilisation of the open DNA state. This places the +1 template strand within the RNAP active site, allowing the binding of NTPs and initiation of transcription (Saecker *et al.*, 2002; Kontur *et al.*, 2008). A simplified version of these initiation events can be seen in Figure 1.2.3.1. The synthesis of RNA of up to 12 nt then causes the extension of this RNA towards to the exit channel, causing the displacement of the σ finger ($\sigma_{3.2}$)(see above), which destabilises the interactions with -35 and -10 elements by regions $\sigma_{4.2}$ and $\sigma_{2.3,}$ respectively, causing the release of the sigma factor and entry into the elongation cycle of transcription (Murakami and Darst, 2003). Prior to this entry into elongation however, RNAP undergoes multiple rounds of abortive transcription, which results in the synthesis of several short RNA transcripts, up to ~10 nt in size. This abortive transcription causes scrunching of the DNA, due to the contacts between RNAP and promoter DNA remaining in place, which is resolved either by the release of this small

abortive transcript (~10 nt), or the further extension to 12 nt which displaces the sigma factor and allows transition into elongation (Murakami and Darst, 2003; Zuo and Steitz, 2015). The extent of this abortive initiation is heavily reliant on the sequences surrounding the TSS (Han and Turnbough, 2014; Heyduk and Heyduk, 2018).

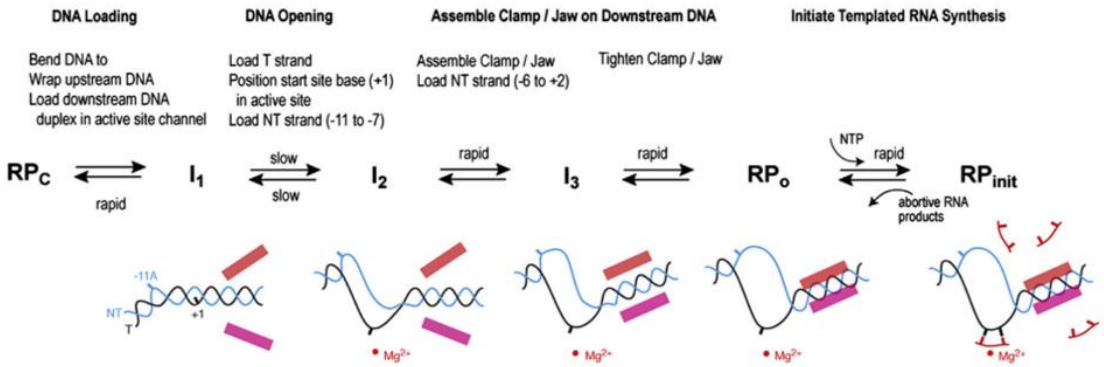


Figure 1.2.3.1: Stepwise images of RNA polymerase during different stages of transcription initiation. A summary of the transition from the RPc (RNAP closed complex) to RPo (RNAP open complex) . I_1, I_2 and I_3 represent the intermediate stages of this transition, resulting from large conformational changes in both RNAP and the DNA. The red and pink rectangles represent each arm of the claw produced by the β and β' subunits. Extracted from the review by Saecker *et al.*, 2011, with permission. License number: 5194790641590.

1.2.3.2 Elongation

Elongation, occuring after transition from initation and promoter escape, involves the further extension of the ~12 nt long nascent RNA made during initation. This transition is characterised by the conformational change in RNAP from initiation, resultant from the loss of the sigma factor (Vassylyev *et al.*, 2007). The extension of the RNA occurs repetitively, with ~50-100 nucleotides added per second, in a processive fashion (Roberts, Shankar and F, 2008). However 'backtracking' of the enzyme can also occur by movement backwards on DNA, often a result of a mispairing in the RNA:DNA hybrid or a mis-incorporation into the nascent RNA (Roberts, *et al.*,

2008). To resolve this backtracking and resolve arrest of transcription, elongation factors GreA and GreB are required. These proteins bind using their CTD to the edge of the secondary channel within RNAP, the channel NTPs enter for access to the enzyme active site (Roberts, Shankar and F, 2008). Here the Gre factors catalyse the removal of the 3' of the RNA allowing for synthesis to restart at the 3' OH primer end created from this hydrolysis (Borukhov *et al.*, 1992). These Gre factors are also important for promotion of elongation from the initation step of transcription, by supressing abortive transcription. Other factors that work to increase the rate of elongation are NusG and RfaH (discussed in Section 1.2.5.1).

1.2.3.3 Termination

The last stage of transcription is termination and involves the disociation of RNAP from the DNA template and release of the nascent RNA, which can occur either by intrinsic or rho-dependent termination in bacteria.

Rho-dependent termination occurs for transcripts that lack the secondary structures associated with intrinsic termination, and rather rely on the contribution of an ATP-dependent helicase known as Rho (Lawson and Berger, 2019). For activity six Rho subunits form a homo-hexamer in a circle arrangement and bind to specific cytosine rich sequences known as 'rut' (Rho-utilisation) sites in 3' untranslated regions of nascent RNA. The Rho then translocates in the 5' to 3' direction along the RNA by utilising the hydrolysis of ATP, until the elongation complex is reached (Skordalakes and Berger, 2003; Mitra *et al.*, 2017). The termination site usually lies ~60 – 90 nt downstream from the rut site however, the exact mechanism of termination is not well understood . The dissociation of the elongation complex is believed to be a result of the unwinding of the DNA:RNA hybrid by this helicase, causing the collapse of the transcription bubble (Dutta, Chalissery and Sen, 2008). Rho has largely been studied in *E. coli* and is essential in this organism but also essential for other organisms such as *Salmonella enterica* and *M. tuberculosis*, however it is dispensible in *B. subtilis* and *Streptomyces lividans* (reviewed by Mitra *et al.*, 2017).

Intrinsic termination does not require additional proteins to carry out termination. Instead RNA secondary structures in the transcribed sequence form hairpin loops that cause RNAP to stall and release from the DNA. Hairpins are formed by a GC rich

inverted repeat sequence of ~7-20 nt followed by a run of UTP nucleotides in the RNA (~7-8 UTPs), which are both encoded downstream of the stop codon and required for sufficient termination (Gusarov and Nudler, 1999; Czyz *et al.*, 2014). The hairpin stalls the RNAP on the template DNA, and the weak U:A RNA:DNA hybrid provides further instability of this complex, causing RNAP to disassociate from the DNA and RNA to be released. However, it was noted that a large number of terminators in *Streptomyces* and *Mycobacteria*, lacked a long run of U nucleotides after the identified hairpins, and efficient termination is thought to occur due to a run of contiguous hairpins, but this is still to be confirmed (Deng, Kiesei and Hopwood, 1987; Gupta and Pal, 2021).

1.2.4 Factors affecting RNA polymerase activity

Aside from sigma factors, other proteins such as transcriptional activators or repressors can alter the activity of the RNAP enzyme. Other effectors aside from proteins can too determine the rate of transcription, including the intracellular levels of nucleotides.

1.2.4.1 RNA polymerase binding proteins

1.2.4.1.1 DksA

DksA is a key 17.5 kDa protein that is expressed at a constant level at all growth phases in proteobacteria. This protein is made up of five α helices that comprises three domains: a globular domain, which contains a zinc binding region, a central segment which forms an extended coiled-coil domain and an extended α helix which is formed by the C-terminal end of DksA (Perederina *et al.*, 2004). This protein acts in a similar way to the Gre proteins by insertion of a coiled coil into the RNAP via its secondary channel and is required to enhance both positive and negative effects of (p)ppGpp, the alarmones that carry out the stringent response in bacteria (Section 1.2.4.2). Both DksA and ppGpp bind to the RNAP at promoters and reduce the lifetime of the RPo at all promoters, which can either be beneficial or disadvantageous to transcription depending on the specific promoter itself (Molodtsov *et al.*, 2018). DksA and ppGpp

have also been shown to have sole effects at differing promoters during varying conditions, displaying their ability to alter transcription independently (Łyzén *et al.*, 2016). As well as this, DksA has been associated with the elongation complex of transcription, and has been shown to arrest elongation complexes by preventing backtracking of the RNAP (Zhang *et al.*, 2005).

1.2.4.1.2 CarD

CarD another transcription factor found in many eubacteria, is essential in Mycobacteria yet is not conserved in *E. coli* (Stallings *et al.*, 2009; Flentie, Garner and Stallings, 2016). CarD is suggested to play a role in stress response due to the apparent induction of this regulator by various stresses including DNA damage and both oxidative stress and starvation (Stallings and Glickman, 2011). In addition to this, 193 and 176 genes were upregulated and downregulated >2 fold, respectively, upon the removal of CarD, providing evidence that this transcription factor can both activate and repress transcription at promoters (Stallings *et al.*, 2009). Crystal structures revealed this protein works at the level of transcription initiation by preventing the collapse of the transcription bubble, further stabilising the RPo (Davis *et al.*, 2015; Rammohan *et al.*, 2015; Zhu *et al.*, 2019). However, CarD is unable to bind the genome alone and for this modulation of expression CarD is required to bind to RNAP, with this complex shown to co-localise at promoters dependent on the principal sigma factor in *M. tuberculosis*, σ^A (Srivastava *et al.*, 2013). The kinetic properties of the targeted promoter region are thought to determine the role of CarD, either activation or repression of expression, with the latter suggested to result from inhibition of promoter escape (Zhu *et al.*, 2019).

CarD is composed of two domains, a CTD containing a compact fold consisting of 5 α helices and an NTD which appears as a tudor-like fold (Srivastava *et al.*, 2013). The structure of the CTD bears no resemblance of sequence similarity to any fold described previously, however the NTD has a similar amino acid sequence and structure to the transcription-repair coupling factor, encoded by the *mfd* gene, more specifically the RNAP binding region of this protein (Srivastava *et al.*, 2013). Binding of CarD and RNAP is a result of interactions between the NTD of CarD and the N-termini β-subunit of the RNAP, as well as interactions between the upstream edge of the -10 promoter

element via its CTD, on the opposite side to that of σ, which stabilises the RNAP open complex (RPo) (Stallings *et al.*, 2009; *Srivastava et al.*, 2013).

Within the CarD CTD lies a highly conserved exposed tryptophan residue (W86 and W85 in *M. tuberculosis* and *S. coelicolor*, respectively), that is surrounded by a cluster of basic residues and situated away from the RNAP-binding region of CarD. This tryptophan residue is essential for CarD to function as a transcriptional activator, which is likely due to its role as the most proximal residue in relation to the DNA, facilitating CarD/DNA interactions centring on W86 around the -12 base. This then situates the basic residues of the CTD at the upstream edge of the transcription bubble, around -11 to -15 residues on the minor groove of DNA, stabilising the RPo providing evidence for its importance at the level of transcription initiation (Srivastava *et al.*, 2013).

1.2.4.1.3 RbpA

RNA Polymerase Binding Protein A (RbpA) is a 14 kDa transcriptional activator with significant conservation in the Actinobacteria phylum, and is essential for the normal growth of both *S. coelicolor* and *M. tuberculosis* (Newell *et al.*, 2006; Hu *et al.*, 2014; Hubin *et al.*, 2015). RbpA has four key domains: N-terminal tail (NTT); core domain (CD), basic linker (BL); sigma interaction domain (SID) (Hubin *et al.*, 2015, 2017). The BL contains 4 positively charged amino acids (K73, K74, K76 and R79 in *M. tuberculosis*) that are required for interactions with the DNA, with mutagenesis studies of R79 displaying this conserved residue is required for full function of RbpA *in vivo*, and sufficient transcription initiation *in vitro* (Hubin *et al.*, 2015). The N-terminal tail of RbpA, consisting of 26 amino acids, interacts with the lid of β' and $\sigma^{3.2}$ region by insertion into the active site cleft of RNAP (Boyaci *et al.*, 2018). As expected this SID, which is composed of 2 α-helices, interacts with the specific sigma factor bound to RNAP (discussed above), with deletion of this region ceasing all interaction with σ (Tabib-Salazar *et al.*, 2013).

Upon the deletion of RbpA from *S. coelicolor*, mutants presented a slow growth phenotype with smaller colonies on solid media and a slow exponential phase displayed in liquid media (Newell *et al.*, 2006). This is as a result of RbpA's ability to activate promoters, more specifically those recognised by the principal sigma factors,

σ^{A} and σ^{HrdB} in *M. tuberculosis* and *S. coelicolor,* respectively (Tabib-Salazar *et al.*, 2013; Hubin *et al.*, 2015). RbpA forms a binary complex with principal sigma factors (Hubin *et al.*, 2015) and also can interact with alternative Group 2 factors, but not to Group 3 and Group 4 sigma factors. Structures of RNAP transcription initiation complexes have revealed that RbpA also interacts via its core domain with the β' zinc binding domain (ZBD) and residues flanking the β' zipper subunit (Hubin *et al.*, 2017). Furthermore, the flexible NTT of RbpA winds into the active site and interacts with the sigma finger (Boyaci *et al.*, 2018). In transcription initiation complexes the BL stabilises the RPo complex through interaction with the DNA minor groove at -14 upstream from the transcription initiation site. This is similar to CarD and the two proteins are hypothesised to have an additive effect on initiation at promoters (Rammohan *et al.*, 2016; Jensen *et al.*, 2019). This is supported by the fact there are no steric clashes between these proteins in cryo-EM structures, with both found at initiation complexes containing σA in *M. tuberculosis* (Hubin *et al.*, 2017; Lilic, Darst and Campbell, 2021). Like CarD, RbpA can too activate or repress transcription dependent on the promoter itself (Prusa *et al.*, 2018).

1.2.4.2 (p)ppGpp and the stringent response

The stringent response is a global change in gene expression in response to stress conditions, including amino acid or fatty acid depletion, or nutrient starvation, that is found in all free-living bacteria (Stephens, Artz and Ames, 1975). It is characterised by the rapid downregulation of growth-related genes, as well as a simultaneous upregulation of those involved in the stress response and amino acid synthesis (recently reviewed by Travis and Schumacher, 2022). These changes are coordinated by the highly phosphorylated alarmones, guanosine tetraphosphate and pentaphosphate, (collectively known as (p)ppGpp, but generally referred to as ppGpp henceforth), that are produced during stress by the ribosome-associated RelA protein (Cashel, 1969; Barker *et al.*, 2001; Dalebroux and Swanson, 2012). Upon entry of an uncharged tRNA into the acceptor (A-site) of the ribosome, the RelA protein begins synthesising these alarmones by transferring the β- and γ-phosphates of ATP onto the 3' hydroxy group of GDP (guanosine 5'diphosphosphate) and GTP (guanosine 5'triphosphate) (Haseltine and Block, 1973; Sy and Lipmann, 1973).

In addition to RelA, the SpoT protein present in *E. coli* is also able to act as both a (p)ppGpp synthetase and hydrolase, contributing to regulation of ppGpp. However its synthetase activity is much weaker compared to RelA, and SpoT also produces ppGpp in response to different intracellular signals such as iron or fatty acid limitation (Seyfzadeh, Keener and Nomura, 1993; Vinella *et al.*, 2005). Proteins identified with ppGpp synthetase activity and thus homology to either of these proteins, are grouped into the RelA-SpoT homologue (RSH) family (Atkinson, Tenson and Hauryliuk, 2011).

The stringent response has been well characterised in *E. coli*, with ppGpp binding to two distinct and independently acting sites on RNAP that are separated by 60Å (Magnusson, Farewell and Nyström, 2005). At Site 1, ppGpp directly interacts with the ω and β' subunits to induce global changes in expression (Ross *et al.*, 2013; Zuo, Yuhong, Wang and Steitz, 2013). At Site 2 the above mentioned DksA along with β' residues, bind ppGpp, explaining the requirement for DksA for the stringent response when Site 1 is absent (Ross *et al.*, 2016; Molodtsov *et al.*, 2018). The requirement for DksA, in combination with ppGpp, is both for the activation of amino acid biosynthesis genes and the repression of rRNA promoters (Paul *et al.*, 2004; Paul, Berkmen and Gourse, 2005). In general ppGpp/DksA is thought to decrease the half-life of RPo (Barker *et al.*, 2001), which is particularly effective at suppressing transcription from rRNA promoters that already form unstable open complexes in the absence of ppGpp (Shin *et al.*, 2021). The mechanism of activation might in part be indirect due to an increased availability of RNAP for alternative promoters, due to decreased rRNA transcription, but also the direct activation by ppGpp/DksA (Magnusson, Farewell and Nyström, 2005; Paul, Berkmen and Gourse, 2005; Molodtsov *et al.*, 2018).

Due to a lack of direct ppGpp/DksA interactions with the DNA, the regulation at specific promoters is more complex, and is down to promoter specific kinetics such as RNAP-promoter interactions and energy requirements of strand separation (Haugen, Ross and Gourse, 2008; Gourse *et al.*, 2018). Particular residues of the promoter DNA were highlighted of importance when determining promoters either up- or downregulated by ppGpp/DksA. Promoters inhibited by this complex have been shown to contain both a GC-rich TSS and discriminator region, which is thought to make DNA separation more difficult (Sanchez-Vazquez *et al.*, 2019). Bases A and G

(non-template DNA) at positions -9 and -8, were also shown to be strongly favoured and disfavoured respectively, at inhibited *E. coli* promoters, including those controlling rRNA expression. These promoters also contain a strong preference for a -1 C (non-template DNA) which, in combination with the GC rich discriminator and TSS, causes the RPo to be unstable at these promoters, as discussed above (Barker *et al.*, 2001; Winkelman *et al.*, 2016; Sanchez-Vazquez *et al.*, 2019). In particular rRNA promoters require high concentrations of initiating ribonucleotide (iNTP), ATP in the case of the *rrnB*p1 rRNA promoter, sufficient for the immediate synthesis of RNA and rapid transition into transcription elongation, with the binding of DksA-ppGpp further reducing this transition into RPo formation (Winkelman *et al.*, 2016; Shin *et al.*, 2021).

Stringent response in Gram-positive bacteria

The definition of ppGpp binding sites in proteobacterial RNAP does not translate to Gram-positive bacteria, such as the well-studied *Bacillus* spp. with the identified binding sites at the interface of the β' and ω not conserved in either *Thermo thermophilus* or *Bacillus subtilis* due to evolutionary divergence (Fig 1.2.4.2). Hence an alternative mechanism of stringent control is utilised in *B. subtilis,* as likely other Gram-positive bacteria, where instead this alarmone modulates GTP levels. This control of intracellular GTP, which rapidly decreases during stringent conditions as a result of both consumption for the production of ppGpp, and the further inhibition of GTP synthesis enzymes, serves as importance for rRNA promoters and other growth related promoters that require this nucleotide for transcription initiation (Krásny and Gourse, 2004; Liu, Bittner and Wang, 2015; Travis and Schumacher, 2022). Both GTP biosynthetic enzymes guanylate kinase (GMK) and hypoxanthine phosphoribosyltransferase (HPRT), involved in the conversion of guanine to GMP and hypoxanthine to IMP, respectively, are inhibited at the protein level by ppGpp (Liu, *et al.*, 2015; Corrigan *et al.*, 2016). Structural studies have also confirmed the competitive binding and inhibition of ppGpp at the active site of GMK of several Firmcutes, including *S. aureus* and *B. subtilis* (Liu, *et al.*, 2015).

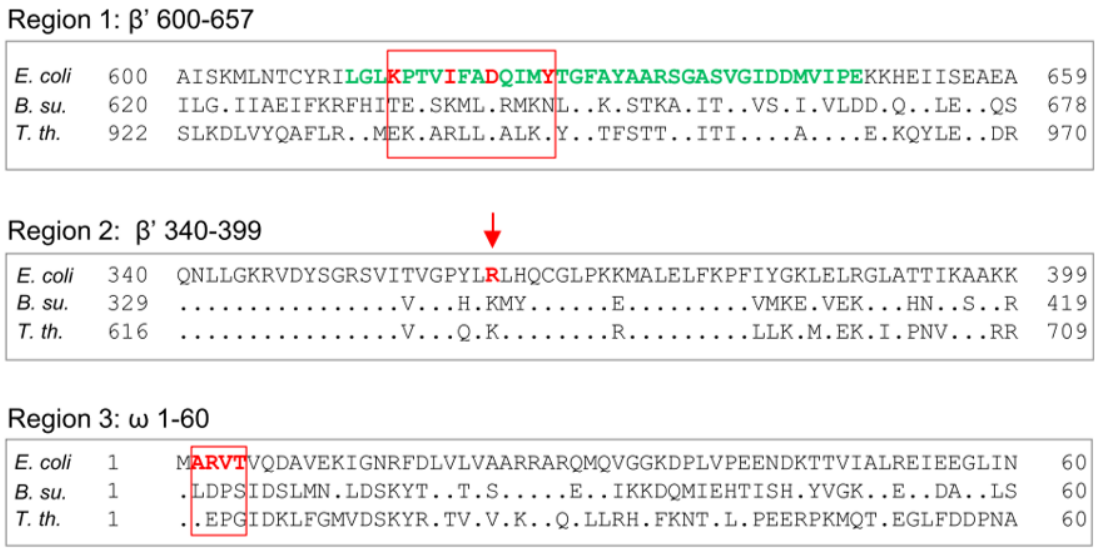


Figure 1.2.4.2: The evolutionary conservation of the ppGpp binding sites in *E. coli, T. thermophilus* and *B. subtilis.* The alignment of three 60 amino-acid regions containing the proposed ppGpp binding sites in the β' and ω subunits of *E.coli* RNAP and counterpart sequences in *B. subtilis* (*B.su.*) and *T. thermophilus* (*T.th.*). Specific residues predicted to bind ppGpp are highlighted in red, with red boxes indicating the residues with the greatest effects on the binding of ppGpp. Cross-linking experiments were used to confirm the binding to RNAP using 6-thio-ppGpp, which were located within the β' subunit and are highlighted in green. Dots represent sequence identity between residues/strains. Figure extracted, with permission, from Ross *et al.,* 2013. License number: 5314970421004.

Stringent response in *Streptomyces* spp.

The response to ppGpp in Actinobacteria including *Streptomyces*, has been well studied but remains poorly understood at the mechanistic level. Early studies indicated that upon nutritional downshift in *S. coelicolor*, by the removal of amino acids from the growth medium, a large decrease in GTP concentration was observed within 5 min of stress, suggesting a similar mechanism to other Gram-positive organisms (see above) (Strauch *et al.*, 1991; Corrigan *et al.*, 2016). This is supported by the fact that GTP biosynthesis is also affected in *S. coelicolor* where GMK is also inhibited by ppGpp. The GMK-ppGpp interaction is conserved in *S. coelicolor* and within other Actinobacteria, suggesting the stringent response is carried out in a similar way to that observed in other organisms such as *S. aureus* and *B. subtilis* (Hesketh *et al.*, 2007a; Ross *et al.*, 2014; Liu, Angela R. Myers, *et al.*, 2015).

ppGpp is also required for the general growth progression and morphological differentiation in *Streptomyces*, and is demonstrated to play a significant role in the

initiation of secondary metabolism, acting as both a positive and negative regulator of antibiotic production (Ochi, 1986; Gatewood and Jones, 2010). The effect of ppGpp on antibiotic production has been long established, with a *relA*Δ mutant constructed in *S. coelicolor* (M570) shown to be incapable of synthesising ppGpp under nitrogen-limiting conditions, as well as both the RED and ACT antibiotics that are usually detectable in this strain (M600) (Section 1.1.3.3) (Chakraburtty and Bibb, 1997). However the relationship between ppGpp and antibiotic production is not straightforward – the antibiotics cephalomycin C and clavulanic acid are overproduced in a *relA*Δ mutant of the *Streptomyces clavuligerus* strain, (Gomez-Escribano *et al.*, 2008).

Aside from the role of ppGpp in regards to antibiotic production, the alarmone has also been shown to inhibit non-competitively polynucleotide phosphorylase (PNPase) in *Streptomyces* (Gatewood and Jones, 2010). PNPase is required for both the synthesis of polyribonucleotides, including the production of poly(A) tails and its exonuclease activity in *Streptomyces* (Sohlberg, Huang and Cohen, 2003; Bralley *et al.*, 2006). The inhibition of this protein by ppGpp resulted in an increased stability of bulk mRNA, which may be expected due to decreased poly (A) tail synthesis by PNPase, resulting in less degradation of RNA. It was proposed that this may provide increased time for translation to complete which may contribute to the production of antibiotics that is observed during the stringent response, however this is still to be proven. It should further be noted that PNPase activity was unaffected in *E. coli* indicating that this response might be limited to *Streptomyces* and perhaps other Actinobacteria (Gatewood and Jones, 2010).

1.2.5 Transcriptional pausing

Transcriptional pausing, characterised by a transient halt of RNAP can occur during any stage of transcription (initiation, elongation or termination), and can result in either the termination of transcription, or provide short time periods where further regulatory events can occur (Kang *et al.*, 2019). These pause events are a common and usually reversible, with RNAP returning to active transcription. Pausing is crucial for dynamic regulation of transcription and is sensed directly or indirectly by RNAP, with the additional time provided allowing for such things as the binding of signalling molecules to riboswitches, which can determine termination decisions, or for protein

loading that can further enable recruitment of ribosomes for sufficient translation (Artsimovitch and Landick, 2002; Yakhnin *et al.*, 2020).

Usually pauses are <10 seconds (known as short elemental pauses), however the extent of these pauses is often determined by the sequence in the DNA template with sequence-specific pauses often longer (>20 seconds) (Herbert *et al.*, 2006; Gajos *et al.*, 2021). Elemental pauses can also give rise to prolonged pause states by either backtracking of the RNAP on the nascent RNA, formation of RNA secondary structures such as hairpin loops, and/or the interaction of transcriptional regulator proteins that can further stimulate (or repress) this pausing action (Zhang and Landick, 2016; Landick, 2021). RNA structures in particular can further increase pause time by invading the RNA exit channel and stabilising the open clamp conformation of RNAP, which can subsequently affect the success of coupled translation. This can include blocking access to the Shine-Dalgarno sequence, preventing ribosomal binding, as well as pausing of RNAP on these hairpin structures, allowing the sufficient time for translation initiation and coupling of transcription to translation (Hein *et al*., 2014; Yakhnin *et al.*, 2019). This coupling allows the release of the RNAP when the pause hairpin is disrupted by the translating ribosome, allowing synchronised transcription and translation to continue, which can further determine the transcriptional effects of attenuator mechanisms present in downstream leader sequences (Zhang and Landick, 2016). Nevertheless, it should be noted that the formation of hairpins does not necessarily mean an increase in pause dwell time as RNA structures can also promote RNAP elongation repressing this pausing action (Artsimovitch and Landick, 2000).

The interplay of certain protein regulators has also been shown to occur at these sites stabilising and enhancing pause signals and/or assisting pause hairpin formation. This includes the elongation factors NusA, RfaH and NusG (discussed in Section 1.2.5.1), as well as GreA/B (Section 1.2.3.2). However, there are only a few studies focussing on pause events by bacterial RNAP, with at present no work carried out in the *Streptomyces* genus (Landick, 2021).

1.2.5.1 Nus factors

NusG

NusG, (N-utilization substance G,) is a transcription factor that is conserved in all three domains of life, however its effects on transcriptional pausing differ between organisms (Yakhnin *et al.,* 2016). *E. coli* $NusG_{EC}$ largely increases the rate of RNA biosynthesis by increasing the rate of transcription elongation and reducing the amount of pausing that can eventually lead to termination (Artsimovitch and Landick, 2000). This occurs by the NTD of NusG interacting with the β' subunit of elongating RNAP, stabilising the closed formation of the RNAP clamp, preventing pausing. This protein also has the ability to bind to the ribosomal protein S10 (NusE) promoting transcription-translation coupling, as well as directly interacting with the Rho terminator suggesting a further role in transcription termination (Burmann *et al.*, 2010; Saxena *et al.*, 2019). In contrast to *E. coli*, *B. subtilis* $NusG_{BS}$ is seen to stimulate pausing by binding to the RNAP and interacting with specific residues, predominantly T residues, within the paused transcription bubble on the non-template DNA strand (Yakhnin *et al.*, 2008; Yakhnin *et al.*, 2016; Yakhnin *et al.,* 2019). This sequence specific pausing was seen to be apparent for both the *trpEDCFBA* operon encoding genes required for tryptophan biosynthesis, and the *tlrB* gene encoding the methylase required for tylosin resistance, with the substitution of these residues leading to a much shorter pause half-life of the complex (Yakhnin, Murakami and Babitzke, 2016; Yakhnin *et al.,* 2019). *Mycobacterium tuberculosis* $NusG_{MTb}$ was also seen to be deficient in increasing the elongation rate of RNAP, and in its ability to bind the Rho termination factor, also a characteristic of $NusG_{BS}$, however the *B. subtilis* protein had retained the ability to interact with the NusE (ribosomal S10) protein (Sudha Kalyani *et al.,* 2015). As well as this NusG tested from *M. bovis* was additionally seen to stimulate hairpin dependent transcription termination suggesting a similar negative impact on gene expression as seen with the $NusG_{BS}$ (Czyz *et al.*, 2014).

These differences in NusG action in model organisms *E. coli* and *B. subtilis*, as well as the *Mycobacterium* genus, may be due to the structural differences of this protein in the differing organisms (Sudha Kalyani *et al.*, 2015). *E. coli* NusG, a 21kDa monomeric protein, consists of two domains; the structurally conserved NTD, consisting of a α/β sandwich, required for RNAP interaction, and the CTD formed by

a small β-barrel, connected by a flexible linker (Burmann *et al.*, 2010, 2011). $NusG_{MTb}$, as well as several other mycobacteria homologues, are seen to have an additional N-terminal region compared to the same protein from *E. coli* (Sudha Kalyani *et al.*, 2015). The $NusG_{M.Tb}$ also displayed a larger linker between the C-terminal and N-terminal domain, suggesting further flexibility of these domains relative to an axis *(Strauß et al*., 2015).

NusG likely plays an essential role in *Streptomyces,* however little is known about the genes/operons it controls (Puttikhunt *et al.,* 1993). A recent study by Peng and colleagues (2018) showed that the overexpression of NusG (from *S. coelicolor*) increased the expression and production of antibiotics in *S. lividans*. This may indicate an unseen anti-termination role of NusG in the expression of genes associated with antibiotic production, however, it may also be due to NusG increasing pausing and termination at other growth-related gene sequences, meaning more RNAP available for binding at other promoter regions.

NusA

NusA, another universal bacterial transcription elongation factor, is also shown to be multi-functional in its ability to both stimulate and repress transcriptional pausing in *E. coli*. This protein contains an NTD able to interact with RNAP, as well as 3 conserved RNA-binding domains, S1 and K-homology domains KH1 and KH2 (Gopal *et al.*, 2001).

E. coli $NusA_{EC}$ is necessary for the anti-termination of ribosomal RNA (*rrn*) synthesis, as well as increasing pause half-life and both intrinsic and rho-dependent termination (Vogel and Jensen, 1997). This anti-termination of the *rrn* operon in *M. tuberculosis* was also seen to be reliant on $NusA_{MTb}$, with the first NusA-RNA complex determined in this organism *(Arnvig et al*., 2004; Beuth *et al.,* 2005). This may further indicate that this anti-termination mechanism is universal among bacteria. However, both hairpin-dependent pausing and termination are also enhanced by NusA. NusA caused a ~4 fold increase in termination of transcription of the *tlrB* gene that encodes the tylosin resistance methylase RlmA in *B. subtilis* ($NusA_{BS}$) *(Yakhnin et al.*, 2019).

To increase pause half-life of RNAP, NusA is thought to stabilise RNA hairpins, by the direct interaction of its NTD with the hairpin at the exit channel mouth, as well as interacting with the flap tip covering this exit channel present on the β-subunit of

RNAP (Toulokhonov, Artsimovitch and Landick, 2001). $NusA_{BS}$ showed the capacity to produce similar pausing when compared to $NusA_{EC}$, suggesting related mechanisms of action (Yakhnin *et al.*, 2008). However, little is known about the specificity of NusA for binding sequences and the effect RNA sequences have on the stabilisation of this protein in differing organisms. When compared to the more selective elongation factor NusG, NusA is thought to cause more general pausing (Yakhnin *et al.*, 2008; Zhang & Landick, 2016).

RfaH

RfaH, also belonging to the NusG/Spt5 family, is a paralog of NusG and is able to reduce pausing caused by both backtracking of the RNAP and RNA hairpin-stabilised pauses. Structurally, RfAH contains the conserved NTD observed in NusG, however its CTD is composed of a helical α-hairpin which blocks the RNAP binding site (on the NTD) and causes RfaH to exist in an autoinhibited state (Belogurov *et al.*, 2007). This state of RfaH prevents it from spontaneously binding to elongating RNAP, and binding only to this enzyme when the complex is paused at a specific DNA sequence known as *ops* (*operon polarity suppressor*). This displays the increased specificity of this elongation factor, with the *ops* sequence able to form a DNA hairpin on the non-template strand which acts as a recruitment partner for RfaH (Bailey *et al.*, 1996;Kolbi 2014). This further causes displacement of the CTD from the RNAP binding site on the NTD in RfaH, facilitating the binding of this protein to the β and β' subunits of RNAP (Kang *et al.*, 2018; Galaz-Davison, Román and Ramírez-Sarmiento, 2021). However, compared to NusG, RfaH is non-essential and is not capable of binding to the Rho terminator (Burmann *et al.*, 2012), with homologs of this protein not apparent in the *Streptomyces* genus.

1.2.5.2 Reiterative transcription

Reiterative transcription (RT), also known as pseudo-templated transcription, transcriptional slippage or RNAP stuttering, is characterised by the slipping of the RNAP enzyme on particular homopolymeric tracts in DNA. This causes movement of the nascent RNA upstream or downstream, relative to the template DNA in the active site of the enzyme, resulting in either the repeat or loss of nucleotides in the nascent RNA, respectively (Fig 1.2.5.1.1) (Jacques and Kolakofsky, 1991; Xiong and

Reznikoff, 1993). This phenomenon occurs due to the weakening of the DNA:RNA hybrid that occurs on these tracts, with these slippage events often occurring multiple times leading to the loss of nucleotides, or a single base in the DNA template specifying several residues in the RNA transcript, which has been shown to play a key regulatory role of gene expression in several differing organisms (Han and Turnbough, 2014). Tracts of three of the same nucleotide or more are required for this mechanism to occur during transcription initiation, with homopolymeric tracts of at least 9 nt required for slippage events to occur during elongation or termination (Fig 1.2.5.1.2) (Cheng, Dylla and Turnbough, 2001; Zhou *et al.*, 2013; Molodtsov V *et al.*, 2014). These slippage events are more commonly observed during initiation, however, slippage of an elongating RNAP can result in frame-shift mutations in the nascent RNA causing the expression of alternate reading frames and the beneficial reduction in the expression of certain genes (Larsen *et al.*, 2000; Anikin *et al.*, 2010).

RT is observed to occur on tracts consisting of any of the nucleotides (A, T, G or C), and has also be observed for runs of dinucleotide repeats that cause two-base shifts in transcripts, however, this mechanism is more commonly identified and more efficient for tracts which allow the addition of U or A in to the nascent RNA transcript, likely due to a weaker DNA:RNA hybrid (Qi *et al.*, 1996; Pal and Luse, 2002). Resolution of RT can include either the release of the RNA transcript from the RNAP complex, resulting in transcription termination, or can proceed with canonical transcription allowing for the full RNA transcript to be produced (Zhou *et al.*, 2013; Molodtsov *et al.*, 2014; Vvedenskaya *et al.*, 2015).

This pausing/stuttering of the RNAP on these homopolymeric tracts results from the sensing of nucleotide concentrations, usually those associated with the homopolymeric tract and/or the following nucleotide, and can allow gene expression to be altered in accordance with intracellular conditions (Turnbough, 2008). This is particularly important for the expression of genes associated with the production of nucleotides themselves (see below).

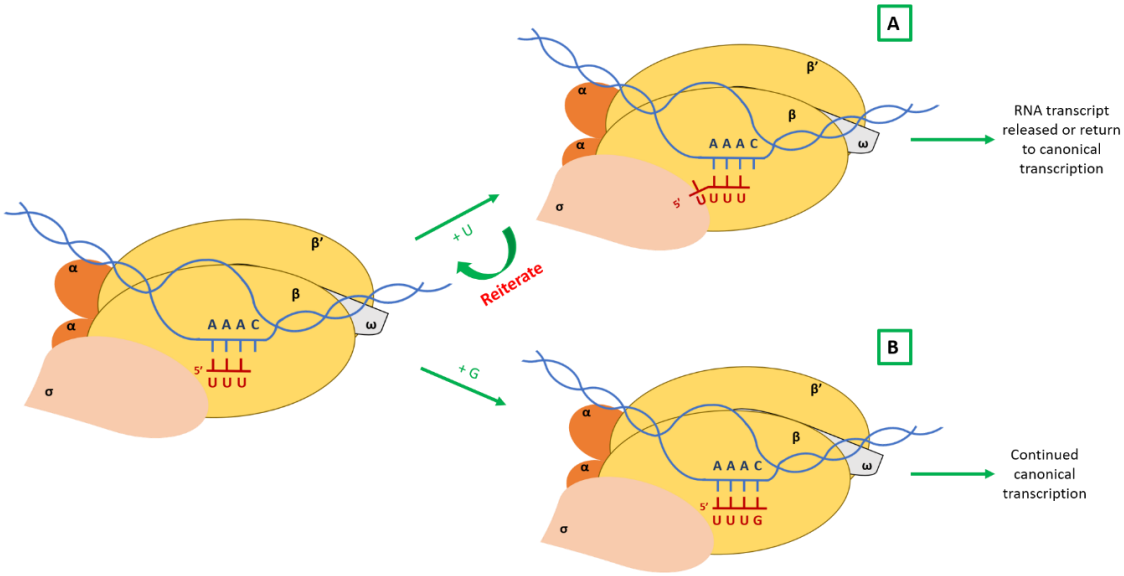


Figure 1.2.5.1.1: Reiterative transcription during initiation. The RNAP subunits are indicated on the enzyme with the nascent RNA and DNA template shown in red and blue, respectively. A displays the DNA sequence resultant from one reiterative transcription event on a homopolymeric tract consisting of ATP. B displays the addition of the correct +4 G nucleotide (after the tract) required for commitment to transcription.

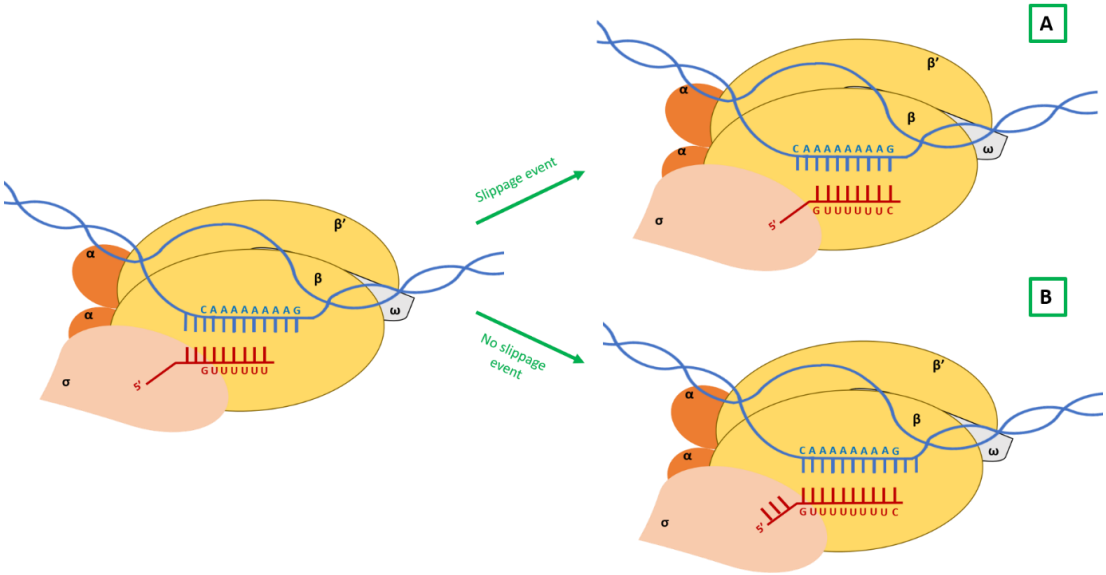


Figure 1.2.5.1.2: Reiterative transcription during elongation. The RNAP subunits are indicated on the enzyme with the nascent RNA and DNA template shown in red and blue, respectively. A displays the loss of U nucleotides resultant from slippage on an ATP encoded homopolymeric tract. B displays the correctly encoded nascent RNA produced in the absence of reiterative transcription.

Reiterative transcription during transcription initiation

Full length transcripts that result from RT events during initiation exhibit altered 5' ends, displaying either additional or loss of nucleotides, as a result of RNAP slipping. The factors that affect the outcome of RT are largely unstudied, however the distance between the homopolymeric tract and both the -10 element and TSS likely affects the extent of slippage events and outcome of transcription (Han and Turnbough, 1998). It has already been noted that transcripts subjected to UTP RT events are generally released from transcription initiation complexes (Turnbough and Switzer, 2008).

Some of the first examples of this regulation in bacteria was in the control of pyrimidine biosynthesis genes, regulated at the level of transcription initiation. Both control of the *pyrBI*, and the *carAB* operons were seen to be controlled by UTP-sensitive RT in *E. coli*, with > 30 additional U residues seen to be encoded in the nascent RNA transcript *in vitro* (Liu, Heath and Turnbough Jr., 1994). The *pyrBI* operon, encoding the aspartate transcarbamylase enzyme that catalyses the first step for de novo pyrimidine synthesis, contains an 5' **A**ATTTG 3' sequence (non-template DNA; TSS in bold), with this T-tract liable to slippage in high UTP levels, causing one or more additional UTP residues to be incorporated, ultimately leading to the release of the transcript and transcription termination (Liu, *et al.*, 1994). However, at low UTP levels, where expression is required, slippage is less likely to occur and the addition of the following GTP nucleotide to the 3' end of the transcript prevents slippage and increases the likelihood of entry into transcription elongation. This is similar to the control elicited at the p1 promoter for the *carAB* operon that encodes the two subunits of the carbamoyl phosphate synthetase enzyme, required for the formation of carbamoyl phosphate, an intermediate for both pyrimidine and arginine synthesis. An increase of RT is again seen on a homopolymeric T-tract (5' **G**TTTG 3', non-template DNA, TSS in Bold) upon increased UTP levels with these slippage events again causing the release of this transcript (Han and Turnbough, 1998).

Slippage events that occur during initiation do not always lead to the downregulation of transcription as seen with the CTP-sensitive RT that is implemented in the control of the *pyrG* gene in *B. subtilis* (Meng, Turnbough and Switzer, 2004). Transcription of the *pyrG* gene, encoding the CTP synthetase, is seen to initiate on a tract of G's, (5'- **G**GGC, TSS in bold), which itself is liable to slippage events dependent on the

availability of the following CTP nucleotide that is required to commit to canonical transcription. When CTP levels are low, slippage events can occur on this G-tract with up to 5 additional (8 nt in total) G nucleotides encoded at the 5' end of the transcript. These additional nucleotides allow the formation of an anti-terminator hairpin that disrupts the downstream formation of a terminator hairpin, usually formed in high CTP levels and able to prevent transcriptional read-through into the *pyrG* gene (Fig 1.2.5.1.3). Control of the *pyrG* operon therefore provides an example of how slippage can alter secondary structure formation to control downstream expression, not seen for the control of the *pyrBI* and *carAB* operons.

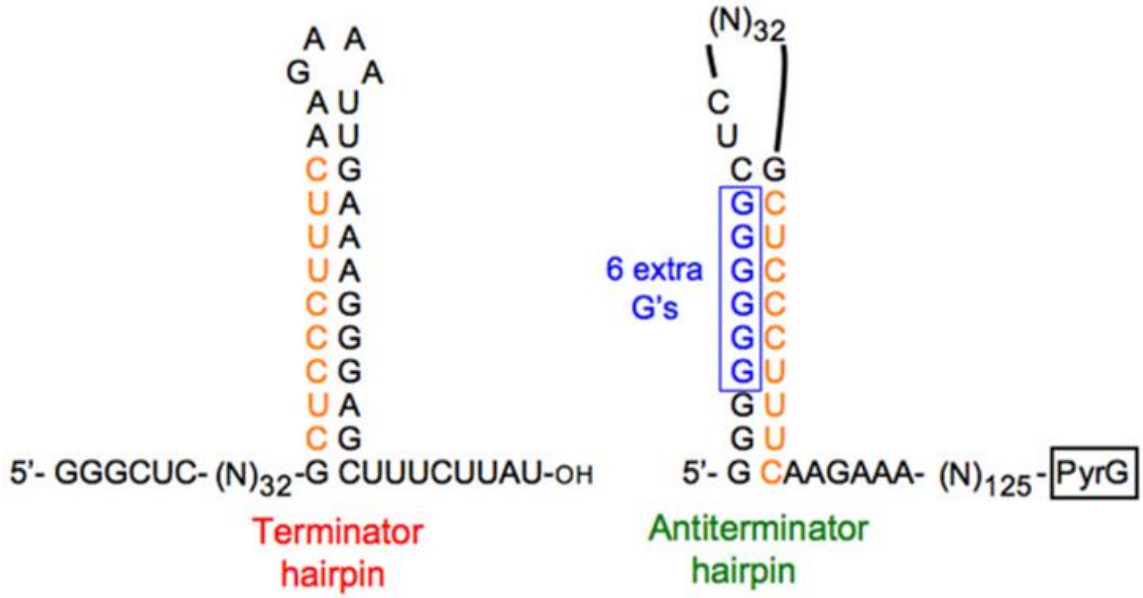


Figure 1.2.5.1.3: The fates of *pyrG* transcription as a result of reiterative or canonical transcription. The formation of a terminator as a result of canonical transcription (left) or an anti-terminator hairpin formed from 6 slippage events resulting in 6 additional GTP included in the nascent RNA (blue). The sequence required for the formation of both hairpins is seen in orange. Figure extracted, with permission from Turnbough (2011), license number: 5196561126466.

Although relatively few characterised examples of RT exist, evidence is accummulating that it is unlikely to be a rare phenomenon. For example, a recent study in the *Streptococcus agalactiae* organism showed that up to 15% of TSS were subjected to some form of slippage, (with the majority of these slippage events associated with repeated A incorporation) (Rosinski-Chupin *et al.*, 2015). Although this mechanism of expression control is yet to be identified in the *Streptomyces* genus, it has recently been characterised within the Actinobacteria phylum. RT was identified to control the expression of the *gyr* operon encoding DNA gyrase in the related *M.*

tuberculosis bacteria (Turnbough, 2011; Jha, Tare and Nagaraja, 2018). This promoter region of the operon encoding the *GyrA* and *GyrB* genes is seen to contain multiple promoters including the principal promoter (P_{gyrB1}) and an unusual promoter (P_{gyrR}) that is seen to be in the reverse orientation to the operon and overlapping the principle promoter. The P_{gyrR} promoter is seen to aid in control of gene expression by competing with the principal promoter for RNAP, as well displaying sensitivity to RT events that further increases the occupancy of RNAP at this promoter. RT occurs at the TSS of this promoter consisting of 5 T residues, with slippage at this site also sensitive to the supercoil state of the DNA, likely representing DNA gyrase availability in the cell (Gellert *et al.*, 1976). When the DNA is supercoiled, and thus DNA gyrase is not required by the cell, an increased amount of slippage and thus promoter occupancy of the reverse promoter P_{gyrR} was observed. This mechanism is thought to be due to an increased and reduced binding of RNAP to this P_{gyrR} and P_{gyrB1}, respectively, when template DNA is supercoiled, and vice versa when DNA is in the relaxed form (Jha, Tare and Nagaraja, 2018). These promoters were also sensitive to ppGpp levels, which can further be related to the growth phase of the bacteria, with expression from the P_{gyrB1} seen to decrease in the presence of ppGpp. This demonstrates the interplay and collaboration of several mechanisms, including growth phase, nucleotide concentration and supercoiled state of DNA on the extent of RT, for further control of overall expression.

Structural studies of a reiteratively transcribing RNAP

Recent studies have used x-ray crystallography to understand further how this phenomenon takes place, with an alternative mechanism of reiterative RNA extension seen to occur when compared to standard initiation complexes. Crystal structures of the RT complex at the *pyrG* promoter region detailed that the 5’ end of the slipped nascent RNA extends towards/ into the main channel of the RNAP, rather than the usual designated RNA exit channel (Murakami *et al.*, 2017). This is thought to be due to the flipping of the -4 (G) slipped residue into the characterised rif pocket in the β-subunit of RNAP, that is likely caused by the base-stacking of the -1 A nucleotide conserved in the template DNA (Shin, Hedglin and Murakami, 2020). After the maximal 8 GTP residues are encoded at the 5’ end, which is thought to be due to the narrowness of the main channel preventing further nucleotide addition, there is a likely

pausing event that allows CTP to be inserted, a switch to canonical transcription and expression of the downstream *pyrG* gene.

More recent crystal structures of reiteratively transcribing complexes confirmed this alternative exit pathway with RNA slippage products of ~6-11 nt extending across the transcription bubble, and >11 nt leaving the active site via the channel in which non-template DNA exits the RNAP during standard transcription initiation (Fig 1.2.5.1.4C) (Liu *et al.*, 2022). These structures also revealed that it is in fact the RNA that slips, and that the DNA-RNA hybrid doesn't extend beyond 3bp in length; the hybrid alternates between a 3 bp pre-slippage state and a 2 bp post slipped state, differing to standard transcription initiation where RNA extension increases up to ~10 bp before promoter escape (Liu *et al.*, 2022). As a result of this, only one translocation event of the non-template DNA strand, relative to the RNAP active centre occurs, regardless of the number of slippage events, and thus only 1 bp of DNA scrunching is observed, as reported in previous models (Fig 1.2.5.1.4B) (Shin, Hedglin and Murakami, 2020).

Initial work suggested that longer nascent RNA transcripts (>8nt) may likely cause the displacement of the sigma factor from the RNAP complex and the termination of transcription, similar to that observed for the *pyrBI* operon (Murakami *et al.*, 2017). However, a small conformational change in the tip of the σ-finger, characterised by a folding back of this protein onto itself, was observed, allowing longer RNAs to exit via this alternate pathway, and further preventing promoter escape (Liu *et al.*, 2022). It was further noted that the slipped RNA spanning 6-11 nt interacts with specific residues of both the RNAP β subunit and σ by either hydrogen bonds or salt-bridged interactions with the sugar-phosphate backbone of this RNA. All of the identified β residues, and the majority of the σ residues, are highly conserved or invariant across both Gram-positive and Gram-negative bacteria, thus this mechanism of slippage extension across the transcription bubble, is likely widely observed across all bacterial species (Turnbough and Switzer, 2008; Liu *et al.*, 2022).

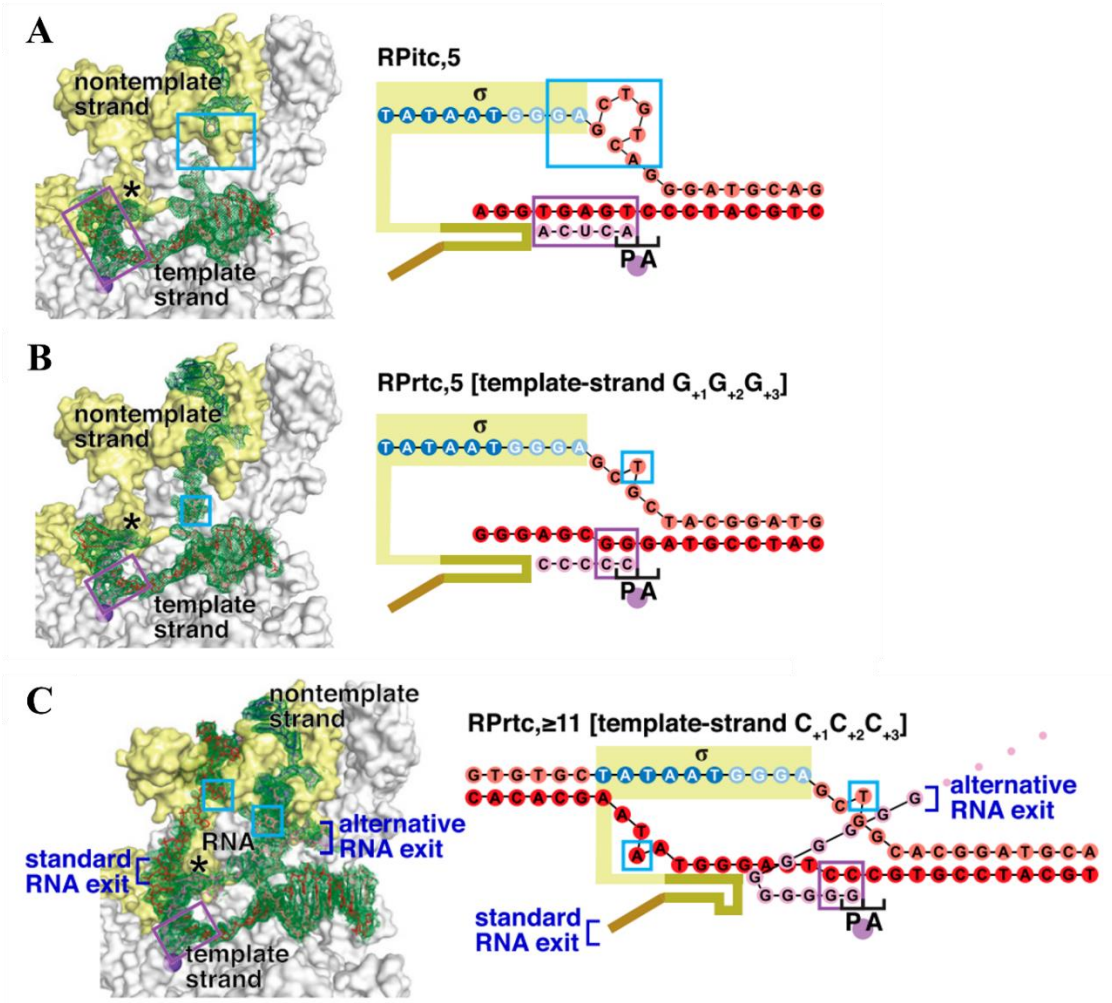


Figure 1.2.5.1.4: Structures of standard and reiterative transcribing initiation complexes. A and B display the crystal structures of standard initiation (RPitc) and reiteratively (RPrtc) transcribing complexes containing a 5 nt RNA product. The σ factor (yellow), -10 element (dark blue), discriminator element (light blue), the rest of the non-template DNA strand (orange), DNA template strand (red) and RNA product (magenta) are highlighted. Cyan rectangles also represent the ordered scrunched nucleotides seen in complexes. C displays the alternative RNA exit for reiteratively transcribing complexes using cryo-EM density and atomic model, with RNA leaving via the main active-site cleft for RNAP, the typical path for the scrunched non-template DNA during standard transcription initiation. The same annotations are highlighted for C as seen for A and B. Figure edited and extracted from Liu *et al.*, 2022, with permission (PNAS, CC-BY-NC-ND).

1.3 Project aims

The focus of this project was to further understand how the levels of RNAP are controlled in *S. coelicolor*, with a specific focus on the transcriptional control of the *rpoBC* operon which encodes the large and most likely rate limiting subunits of RNAP, β and β'. *Streptomyces* are well known for their ability to produce a great plethora of useful secondary metabolites, encoded within their genome, with genome sequencing of multiple strains detailing several large clusters of genes that are not expressed in standard growth conditions (Bentley *et al.*, 2002). This further understanding of how RNAP is controlled may further allow for increased level of enzyme, increasing the likelihood of binding and expression from these silent gene cluster promoters. Previously mutations in *rpoB* were shown to both increase the production of antibiotics, as well as stimulating expression from previously silent gene clusters, resulting in the synthesis of novel antibiotic compounds (Hosaka *et al.*, 2009; Gomez-Escribano and Bibb, 2011; Gomez-Escribano *et al.*, 2012). The aims of this thesis are outlined below:

1.) To identify promoters controlling *rpoBC* expression

2.) To confirm activation at *rpoBC* promoter(s) *in vitro*

3.) To identify areas of control that are elicted at the level of transcription for *rpoBC*

4.) To determine the effects of the stringent response on the expression of *rpoBC*

5.) To determine the effects of oxidative stress on the expression of *rpoBC*

6.) To investigate the effect of changes in *rpoBC* expression on antibiotic production

Chapter 2: Materials and Methods

2.1 Materials

2.1.1 Chemicals and reagents

- Acid Phenol:Chloroform:IAA (125:24:1) (Fisher Scientific)
- Acrylamide/Bis 19:1 40% (w/v) (Ambion®, Fisher Scientific)
- Agar (Sigma-Aldrich)
- Agarose (Melford)
- Ammonium Persulfate (APS; Sigma-Aldrich)
- Ampicillin (Melford Laboratories)
- Antifoam (Sigma-Aldrich)
- Apramycin (Duchefa Biochemie)
- Bromophenol Blue (Amersham Biosciences)
- Casamino Acids (Difco)
- Chloramphenicol (Melford Laboratories)
- Deoxyribonucleotide phosphates (dNTPs; New England Biolabs)
- Diamide (Sigma-Aldrich)
- Dimethylsulphoxide (DMSO; Sigma-Aldrich)
- Dithiothreitol (DTT; Melford Laboratories)
- GelRed (Biotin)
- Glycerol (Fisher Scientific)
- HEPES (Sigma-Aldrich)
- Iodoacetamide (Sigma-Aldrich)
- Isopropyl-β-D-thiogalactopyranoside (IPTG; Melford Laboratories)
- Kanamycin (Melford Laboratories)
- Malt extract (Oxoid)
- Mannitol (Sigma-Aldrich)
- MOPS (Fisher Scientific)
- Nalidixic acid (Duchefa Biochemie)
- N-decanal (Sigma-Aldrich)
- N, N-dimethyl-formamide (Fisher Scientific)
- Nutrient agar (Difco)

- Peptone (Bacto)
- Polyethylene glycol 8000 (Sigma-Aldrich)
- Ponceau C solution (Sigma-Aldrich)
- Protease Inhibitor Cocktail tablets (Sigma-Aldrich)
- Radiolabelled nucleotides (Perkin Elmer)
- Rifampicin (Sigma-Aldrich)
- Sigmacote® (Sigma-Aldrich)
- Sodium chloride (Fisher Scientific)
- Sodium dodecyl sulphate (SDS; Fisher Scientific)
- Soya flour (Infinity Foods)
- Sucrose (Sigma-Aldrich)
- TES (N-Tris(hydroxymethyl)methyl-2-aminoethane sulfonic acid; Fisher Scientific)
- Tetramethyl-ethylenediamine (TEMED; Fisher Scientific)
- Thiostrepton (Sigma-Aldrich)
- Trichloroacetic acid (TCA) (Sigma-Aldrich)
- Tris (2-Amino-2-hydroxymethyl-propane-1,3-diol; Fisher Scientific)
- Triton X-100 (Sigma-Aldrich)
- Tryptone (Difco)
- Tween 20 (Sigma-Aldrich)
- X-gal (Melford Laboratories)
- Yeast extract (Difco/Oxoid)

2.1.2 Enzymes

2.1.2.1 Polymerases

- GoTaq® G2 DNA polymerase (Promega)
- Maxima® SYBR Green/ROX qPCR master mix (2X)
- Phusion™ high-fidelity DNA polymerase (NEB)

2.1.2.2 DNA modifying and restriction enzymes

- rSAP (Shrimp Alkaline Phosphatase) (Promega)
- Restriction endonucleases (NEB)
- RNA-to-cDNA™ Kit (Fisher Scientific)
- RNase A from Bovine pancreas (sigma-Aldrich)
- RNAsin® Ribonuclease Inhibitor (Promega)
- S1 nuclease (Life technologies)
- T4 DNA Kinase
- T4 DNA ligase (NEB)
- TURBO DNA-FREE DNase (Invitrogen)

2.1.3 Antibodies

- Anti-σ^{HrdB} polyclonal antibody, from rabbit (P. Doughty)
- Anti- RNAP β subunit monoclonal antibody, from mouse, (Abcam)

2.1.4 Plasmids

Table 2.1.4: List of plasmids used in this study.

Name	Description	Reference
pBluescript II SK (+)	*E. coli* cloning vector, ori f1, bla, AmpR, ori Co1E1, lacZ	Stratagene Ltd. (Atling-Mees and Short, 1989)
pRT802	*S. coelicolor i*ntegrative plasmid: aphII, oriT, lacZα, ΦBT1 attP, Kan^R.	Gregory *et al.,* 2003.
pIJ12738_*rpoB_2kb_ACT*	Containing 1kb of *rpoB* region either side of the mutated TSS (ACT). Containing restriction site for the Meganuclease enzyme (I-SceI) aac(3)IV, lacZα, oriT, AprR	Fernández-Martínez & Bibb, 2014
pSNypet	pRT802 derivative, containing *ypet* gene. Strong terminator ttsbiB cloned either side of MCS and *ypet* gene.	Gregory *et al.*, 2003 Horbal *et al.,* 2018; Huff *et al*., 2010 Dr Sophie Nicod Sven Reisloehner
pIJ5972	*S. coelicolor* integrative promoter-probe plasmid containing TTA codon-free luciferase genes (*LuxAB*) for the determination of promoter expression using promoter:*LuxAB* fusions. aac(3)IV	Paget, pers comms. Aigle *et al.*, 2000
pET28B::SUMO	*E. coli* based expression plasmid. Kan^R Tet, pBR322 ori, N-terminal 6x His and SUMO encoding gene.	This study
pCRISPOmyces plasmids		
pCRISPomyces 2.0	*Streptomyces* optimised CRISPR/cas plasmid for targeted genome editing.	Cobb *et al.,* 2015
pCRISPO_Sp1	pCRISPOmyces 2.0 containing the spacer sequence alone – no recombination template – control plasmid.	This study
pCRISPO_Sp1_ACT	pCRISPOmyces 2.0 containing Spacer 1 and *rpoB* HR template with mutant ACT	This study

	TSS (No spacer mutation) – control plasmid.	
pCRISPO_Sp1_WT	pCRISPOmyces 2.0 containing Spacer 1 and *rpoB* HR template with wildtype TSS (No spacer mutation) – control plasmid.	This study
pCRISPO_WT_sp1*	pCRISPOmyces 2.0 containing spacer 1 and *rpoB* HR with wildtype TSS and mutation of Sp1 PAM site – control plasmid.	This study
pCRISPO_ACT_sp1*	pCRISPOmyces 2.0 containing spacer and *rpoB* HR template with mutant ACT TSS and mutation of Sp1 PAM site – test plasmid.	This study
pCRISPO_HR_WT	pCRISPOmyces 2.0 (no spacer) containing *rpoB* HR template with wildtype TSS but no Spacer1 mutation (PAM) included – control plasmid.	This study
pCRISPO_HR_ACT	pCRISPOmyces 2.0 (no spacer inc.) containing *rpoB* HR template with mutant ACT TSS but no Spacer1 mutation (PAM) included – control plasmid.	This study

2.1.5 Primers

Table 2.1.5.1: List of primers used for general cloning in this study.

No	Name	Sequence (5' to 3')	Gene	Use:	Chapter used in:
	Luciferase cloning				
1	*RpoBp_P1_UP_EcoRI_FP.*	GGGAATTCGTGATCTTCGTCGTGCCGTC	SCO4654 (*rpoB*)	Used for the amplification of the *rpoBp1* and *rpoBp1U* promoter regions	3.2.3, 4.4.10, 4.4.12
2	*RpoBp_P*1_UP_DOWN_bglII_*RP*	CCAGATCTGAGGGCTCGGACTCACTAC	^	Used for the amplification of the *rpoBp1U* promoter region (luciferase cloning)	3.2.3, 4.4.10, 4.4.12
3	*P1_UP Reverse_BglII_WT*	CCAGATCTGGCAGCGCAAAGGGTCAG	^	Used for the amplification of the *rpoBp1* and *rpoBp1* promoter region	3.2.3, 4.4.10, 4.4.12
4	*P1_UP Reverse_BglII_ACT*	GCAGATCTGGCAGCGCAGTGGGTCAG	^	Used for the amplification of the mutant ACT *rpoBp1* promoter region	3.2.3, 4.4.10, 4.4.12
5	*rpoBp_EcoRI_P2_FP*	CGGAATTCGTGGTCGCTCTTCGGCGTATC	^	Used for the amplification of the *rpoBp1-2U* promoter regions	3.2.3, 4.4.10, 4.4.12
6	rpoBp_EcoRI_interP_FR	CCGAATTCGACCGCACCGGCACCACGGAT	^	Used for the amplification of the *rpoBp1-3U* promoter regions	3.2.3, 4.4.10, 4.4.12

	***ypet* reporter cloning**				
7	*RpoBp_P1_XhoI_FP*	CCCTCGAGGTGATCTTCGTCGTGCCGTC		Amplification of *rpoBp1* promoter regions	3.2.4
8	*P1_HindIII_WT*	CCAAGCTTGGCAGCGCAAAGGGTCAG	^	Amplification of *rpoBp1* promoter regions	3.2.4
9	*P1_HindIII_ACT*	GCAAGCTTTGGCAGCGCAGTGGGTCAG	^	Amplification of the mutant ACT *rpoBp1* promoter regions	3.2.4
10	*RpoBp_HindIII_RP*	CCAAGCTTGAGGGCTCGGACTCACTAC	^	Amplification of *rpoB* promoter regions	3.2.4
11	*rpoBp_XhoI_P2_HrdB_FP*	CCCTCGAGGTGGTCGCTCTTCGGCGTATC	^	Amplification of *rpoBp1-2U* promoter regions	3.2.4
12	rpoBp_XhoI_FP1_interP	CCCTCGAGGACCGCACCGGCACCACGGAT	^	Amplification of *rpoBp1-3U* promoter regions	3.2.4
13	pRT802_Seq_F	GCTGAAGGAGGAAGACGAAG	-	Present in the pRT802 plasmid	3.2.4
14	Y_Pet_Up	AAGAAGTCGTGCTGCTTCATGT	*ypet*	Present in the pRT802 plasmid	3.2.4
	pSX400 cloning				
15	Rpo_DS_BamHI	CCGGATCCGTCGAGCAGCGCGTCGTCCTG	SCO4654 (*rpoB*)	Amplification of *rpoB* gene (~1kb downstream)	3.3.2, 4.4.1.1
16	Rpo_USDraIII2	CCCACGGCGTGTCCGAATCCGAGGCATGGATC		Amplification of *rpoB* gene (~170 bp of start codon)	3.3.2, 4.4.1.1

17	Deg-tag Forward	TCCAGATGGAGTTCTGAGGTC	-	Confirming integration of pSX400 plasmid into *S. coelicolor* genome	3.3.2
18	Deg-tag Reverse	TTCGGGTTGAAGTAGAGGTTC	SCO4654 (*rpoB*)	Confirming integration of pSX400 plasmid into *S. coelicolor* genome (binding within *rpoB* gene)	3.3.2
19	*TSR_FP*	ATGACTGAGTTGGACACCATC	*tsr*	To confirm the presence of the thiostrepton resistance gene in the psx400 plasmid.	4.4
20	*TSR_RP*	TTATCGGTTGGCCGCGAGATTC	*tsr*	^	4.4
21	ACTC mut F Inverse	ACTCCGCTGCCTGTTAGCTGC	SCO4654 (*rpoB*)	Used with Inverse mut R to introduce an ACTC mutation at *rpoBp1* TSS.	4.4.11
22	TTTC mut F Inverse	TTTCCGCTGCCTGTTAGCTGC	^	Used with Inverse mut R to introduce an TTTC mutation at *rpoBp1* TSS.	4.4.11
23	Inverse T-tract mut R	GGGTCAGTGTAGCCACTTGGC	^	Used to introduce mutations in *rpoBp1* TSS.	4.4.11
	pCRISPOmyces cloning				
24	Spacer_1_rpoB_F	ACGCGTCCAGTGCGGTTCTCGGGG	^	Oligo used to create the target protospacer sequence for cutting by Cas9.	4.4.3
25	Spacer_1_rpoB_R	AAACCCCCGAGAACCGCACTGGAC	^	Oligo used to create the target protospacer sequence for cutting by Cas9.	4.4.3

26	pCRISPO_RpoB_2kb_For_gib.	CTCGGTTGCCGCCGGGCGTTTTTTATCTAGA AAGCTCGTCCGCACCGTGGA	^	Amplification of *rpoB* sequence, including homology to pCRISPOmyces plasmid and XbaI site	4.4.3
27	pCRISPO_RpoB_2kb_Rev_gib.	CGCGGCCTTTTTACGGTTCCTGGCCTCTAGATAGATGTCGAGCAGCGCGTC	^	Amplification of *rpoB* sequence, including homology to pCRISPOmyces plasmid and XbaI site	4.4.3
28	*US-rpoB-mut1-r*	GGCAGCTAACAGGCAGCGCAGTGGGTCAGTGTAGCCA	^	Primer containing the ACT mutation within its *rpoB* coding sequence.	4.4.3
29	DS-rpoB-mut1-f	TGGCTACACTGACCCACTGCGCTGCCTGTTAGCTGCC	^	Primer containing the ACT mutation within its *rpoB* coding sequence.	4.4.3
30	Sp1_Mut_Forward_CRISPO	GGAGCGAACACGGTCTCCGAGAACCGCACTG	^	Primer used for the mutagenesis of the PAM site in the *rpoB*	4.4.3
31	Sp1_Mut_Reverse_CRISPO	CAGTGCGGTTCTCGGAGACCGTGTTCGCTCC	^	Primer used for the mutagenesis of the PAM site in the *rpoB*	4.4.3
32	*RpoB*_Internal_seq	GTTCTTGGGAGGCTCGAAG	^	Primer for the amplification within the *rpoB* gene to allow full sequencing coverage.	4.4.3

Table 2.1.5.2: List of primers used in the present study for site-directed mutagenesis of the 5'UTR.

No	Name:	Sequence (5' to 3')	Gene	Use:	Chapter used in:
33	Atten _Mut 1 Forward	AGCCATTAGCTGCCCCCTGCCCGTC	SCO4654 (*rpoB*)	Used to clone the A1 mutation in the 5'UTR of *rpoBp*. Contains half of the NheI restriction site.	3.4.2
34	Atten_mut 1 Reverse	AGCTGAAAGGGTCAGTGTAGCCACTTG	^	Used to clone the A1 mutation in the 5'UTR of *rpoBp*. Contains half of the NheI restriction site.	3.4.2
35	Atten _Mut 2 Forward	AGCCACCTGCCCGTCACCAGGGGTC	^	Used to clone the A2 mutation in the 5'UTR of *rpoBp*. Contains half of the NheI restriction site.	3.4.2
36	Atten_mut 2 Reverse	AGCTGCAGGCAGCGCAAAGGGTCAG	^	Used to clone the A2 mutation in the 5'UTR of *rpoBp*. Contains half of the NheI restriction site.	3.4.2
37	Atten _Mut 3 Forward	AGCCAACCAGGGGTCTACCCTCG	^	Used to clone the A3 mutation in the 5'UTR of *rpoBp*. Contains half of the NheI restriction site.	3.4.2
38	Atten_mut 3 Reverse	AGCTGGGGCAGCTAACAGGCAGC	^	Used to clone the A3 mutation in the 5'UTR of *rpoBp*. Contains half of the NheI restriction site.	3.4.2
39	Atten _Mut 4 Forward	AGCCATACCCTCGCCCGAGCACTGAC	^	Used to clone the A4 mutation in the 5'UTR of *rpoBp*. Contains half of the NheI restriction site.	3.4.2
40	Atten_mut 4 Reverse	AGCTGGACGGGCAGGGGGCAGCTAAC	^	Used to clone the A4 mutation in the 5'UTR of *rpoBp*. Contains half of the NheI restriction site.	3.4.2
41	Atten _Mut 5 Forward	AGCCACGAGCACTGACGACCAGAG	^	Used to clone the A5 mutation in the 5'UTR of *rpoBp*. Contains half of the NheI restriction site.	3.4.2
42	Atten_mut 5 Reverse	AGCTGGACCCCTGGTGACGGGCAG	^	Used to clone the A5 mutation in the 5'UTR of *rpoBp*. Contains half of the NheI restriction site.	3.4.2

43	Atten _Mut 6 Forward	AGCCACGACCAGAGCATGTCTGAC	^	Used to clone the A6 mutation in the 5'UTR of *rpoBp*. Contains half of the NheI restriction site.	3.4.2
44	Atten_mut 6 Reverse	AGCTG GGCGAGGGTAGACCCCTG	^	Used to clone the A6 mutation in the 5'UTR of *rpoBp*. Contains half of the NheI restriction site.	3.4.2
45	Atten _Mut 7 Forward	AGCCAATGTCTGACCTGGCTCTTTAG	^	Used to clone the A7 mutation in the 5'UTR of *rpoBp*. Contains half of the NheI restriction site.	3.4.2
46	Atten_mut 7 Reverse	AGCTGTCAGTGCTCGGGCGAGGGTAG	^	Used to clone the A7 mutation in the 5'UTR of *rpoBp*. Contains half of the NheI restriction site.	3.4.2
47	Atten _Mut 8 Forward	AGCCATGGCTCTTTAGCCACATC	^	Used to clone the A8 mutation in the 5'UTR of *rpoBp*. Contains half of the NheI restriction site.	3.4.2
48	Atten_mut 8 Reverse	AGCTGGCTCTGGTCGTCAGTGCTC	^	Used to clone the A8 mutation in the 5'UTR of *rpoBp*. Contains half of the NheI restriction site.	3.4.2
49	Atten _Mut 9 Forward	AGCCAGCCACATCAGGCACCCCCTG	^	Used to clone the A9 mutation in the 5'UTR of *rpoBp*. Contains half of the NheI restriction site.	3.4.2
50	Atten_mut 9 Reverse	AGCTGGGTCAGACATGCTCTGGTCGTC	^	Used to clone the A9 mutation in the 5'UTR of *rpoBp*. Contains half of the NheI restriction site.	3.4.2
51	Atten _Mut 10 Forward	AGCCAGCACCCCCTGTCTCCGTGCAC	^	Used to clone the A10 mutation in the 5'UTR of *rpoBp*. Contains half of the NheI restriction site.	3.4.2
52	Atten_mut 10 Reverse	AGCTGTAAAGAGCCAGGTCAGACATG	^	Used to clone the A10 mutation in the 5'UTR of *rpoBp*. Contains half of the NheI restriction site.	3.4.2
53	Atten _Mut 11 Forward	AGCCATCTCCGTGCACGGAAGAG	^	Used to clone the A11 mutation in the 5'UTR of *rpoBp*. Contains half of the NheI restriction site.	3.4.2
54	Atten_mut 11 Reverse	AGCTGCTGATGTGGCTAAAGAGCCAG	^	Used to clone the A11 mutation in the 5'UTR of *rpoBp*. Contains half of the NheI restriction site.	3.4.2

55	Atten _Mut 12 Forward	AGCCA CGGAAGAGGGGCCGGTAC	^	Used to clone the A12 mutation in the 5'UTR of *rpoBp*. Contains half of the NheI restriction site.	3.4.2
56	Atten_mut 12 Reverse	AGCTG CAGGGGGTGCCTGATGTG	^	Used to clone the A12 mutation in the 5'UTR of *rpoBp*. Contains half of the NheI restriction site.	3.4.2
57	Atten _Mut 13 Forward	AGCCAGCCGGTACGCGCGTAGTGAG	^	Used to clone the A13 mutation in the 5'UTR of *rpoBp*. Contains half of the NheI restriction site.	3.4.2
58	Atten_mut 13 Reverse	AGCTGTGCACGGAGACAGGGGGTG	^	Used to clone the A13 mutation in the 5'UTR of *rpoBp*. Contains half of the NheI restriction site.	3.4.2
59	Atten _Mut 14 Forward	AGCCAGCGTAGTGAGTCCGAGCCCTC	^	Used to clone the A14 mutation in the 5'UTR of *rpoBp*. Contains half of the NheI restriction site.	3.4.2
60	Atten_mut 14 Reverse	AGCTGCCCTCTTCCGTGCACGGAGAC	^	Used to clone the A14 mutation in the 5'UTR of *rpoBp*. Contains half of the NheI restriction site.	3.4.2
61	Atten _Mut 15 Forward	AGCCATCCGAGCCCTCAGATCTG	^	Used to clone the A15 mutation in the 5'UTR of *rpoBp*. Contains half of the NheI restriction site.	3.4.2
62	Atten_mut 15 Reverse	AGCTGGCGTACCGGCCCCTCTTCCGTG	^	Used to clone the A15 mutation in the 5'UTR of *rpoBp*. Contains half of the NheI restriction site.	3.4.2
63	Atten _Mut 16 Forward	AGCCACAGATCTGGGGGGGATCCAC	^	Used to clone the A16 mutation in the 5'UTR of *rpoBp*. Contains half of the NheI restriction site.	3.4.2
64	Atten_mut 16 Reverse	AGCTGCTCACTACGCGCGTACCGGCCC CTC	^	Used to clone the A16 mutation in the 5'UTR of *rpoBp*. Contains half of the NheI restriction site.	3.4.2

Table 2.1.5.3: List of primers used for qRT-PCR quantification used in the present study.

No:	Name:	Sequence (5' to 3')	Gene:	Use	Chapter used in:
65	qPCR-ypet-l	AGGAGCGGACCATCTTCTTC	*ypet*	Used for determination of ypet reporter expression.	3.2.4
66	qPCR-ypet-r	CCGTCCTCCTTGAAGTCGAT	^	^	
67	16S_QF	CACTAGGTGTGGGCAACATTC	16s rRNA	Used for the determination of stable 16s rRNA expression	3.2.6, 3.3.1, 3.3.4, 3.3.5, 4.4.2, 4.4.5-7, 4.4.11
68	16S_QR	GTCGAATTAAGCCACATGCTC	^	^	
69	16s Unstable F	CGAGAGGGTGAGTACAAAGGAAG	16s rRNA	Used for the determination of Unstable 16s rRNA expression	3.3.4
70	16s Unstable R	CGTCCTCGCTGTTGTGTTAC	^	^	^
71	rpoB_qRTPCR_forward	CCGCTACAAGGTGAACAAGAAG	SCO4654 (*rpoB*)	Used for the determination of *rpoB* expression	3.3.1, 3.3.3-3.3.5, 4.4.2, 4.4.5- 4.4.7
72	rpoB_qRTPCR_reverse	GGTACTTGATGGTGGCGATG	^	^	^

73	LuxA_qPCR FP	CTTATGTTGCTGCCGCACAC	*LuxA*	Used for the determination of *luxA* expression	3.2.3, 4.4.9, 4.4.11
74	LuxA_qPCR RP	GATAGCTGCAGTGCCAACGTTG	^	^	

Table 2.1.5.4: List of primers used for protein purification in the present study.

No:	**Name:**	**Sequence (5' –> 3')**	**Gene:**	**Use**	**Chapter used in:**
75	HrdB_SCO_BglII FP	GCAGATCTTCGGCCAGCACATCCCGTACGCTC	SCO5820 (*HrdB*)	Amplification of the *HrdB* gene	4.3.1.1
76	HrdB_SCO_HindIII RP	GGAAGCTTCTAGTCGAGGTAGTCGCGCAG	^	^	4.3.1.1

Table 2.1.5.5: List of primers used for *in vitro* transcription assays in the present study.

No:	Name:	Sequence (5' –> 3')	Gene:	Use	Chapter used in:
77	*rpoBp(1)* IV RP	GTAGACCCCTGGTGACGGGCAG	SCO4654 (*rpoB*)	Amplification of the *rpoBp(1)* DNA template	4.3
78	RpoBp_P1_UP_EcoRI_FP.	GGGAATTCGTGATCTTCGTCGTGCCGTC	^	Amplification of *rpoB* DNA template.	4.3
79	*rpoBp(2)* IV RP	GAACAAAAGCTGGAGCTCCAC	^	Amplification of the *rpoBp(2)* DNA template	4.3
80	SCO4652 Forward	GGGAATTCGCGCCCGGCCCGCTCCGGTCGCCG	SCO4652 (*rpIJ*)	Amplification of *rpIJ* DNA template	4.3
81	SCO4652 Reverse	GGAAGCTTGTCCTCTTTCGAACACACGGCAACG	^	^	4.3

2.1.6 Bacterial strains

Table 2.1.6: List of bacterial strains used in the present study

Name – *E.coli*	**Genotype**	**Reference**
DH5α^{TM}		Invitrogen
ET12567/pUZ8002 (ETZ)	Dam-13::Tn9 , dcm-6, hsdM, ChlR hsdS, hsdR, cat, tet	MacNeil et al., 1992; Paget et al., 1999
ET12738/ pUB307 (ETR)	RP1 derivative. (dam-13::Tn9, dcm-6, hsdM)	Bennett et al. 1977; Flett F et al., 1997
BL21 (λDE3/pLysS)	F-- ompT gal dcm lon hsdSB(rB-?- mB--) λ(DE3) pLysS(cmR)	Miroux & Walker, 1996
***S. coelicolor* strains**		
M145	SCP1^{-} SCP2^{-}	Hopwood *et al.*, 1985.
M571	M145 ΔRelA::hygR	R. Chakraburtty and M. J. Bibb, unpublished.
J1981	M145 Δ*rpoC::rpoC*^{His}	Babcock et al., 1997

2.1.7 Growth media and selection

2.1.7.1. *E. coli* growth media and storage

Lennox broth (L-Broth): A liquid growth medium used for *E. coli*. 10 g Difco Bacto tryptone, 5 g Difco yeast extract, 5 g NaCl, 1 g glucose dissolved in 1 L dH2O. Stored in 100 ml and 500 ml aliquots.

Lennox agar (L-agar): Solid growth medium for *E. coli*. 10 g Difco Bacto tryptone, 5 g Difco yeast extract, 5 g NaCl, 1 g glucose dissolved in 1 L dH2O. Aliquoted (100 ml) into flasks, stoppered with a foam bung and foil and autoclaved.

2.1.7.2 *Streptomyces* growth media and storage

Liquid media

NMMP minimal medium: NMMP was used for the liquid growth of *Streptomyces* cultures used for RNA extraction. 1 g of $(NH4)_2SO_4$, 2.5 g Casamino acids, 0.3 g $MgSO_47H_20$, 25 g of PEG6000 and 0.5 ml of minor elements solution (see below) was dissolved in 400 ml of distilled dH_20 before being aliquoted into 80 ml (in 100ml bottles) and autoclaved. Before use 15ml of NaH_2PO_4/K_2HPO_4 buffer (0.1M, pH 6.8), 2.5ml of a carbon source (20%) (final concentration 0.5%) and any other required growth factors were added to make a final volume of 100 ml NMMP. Note that dH_20 was added to mixture when no additional growth factors were required. NMMP was also made without the inclusion of casamino acids, for the induction of a stringent response.

Minor elements solution – NMMP:
In 500ml of distilled dH_20, 0.5g of $ZnSO_4.7H_20$, $FeSO_4.7H_20$, $MnCl_2.4H_20$ and $CaCl_2$ annhydrous were dissolved before the mixture was autoclaved and ready for use in the NMMP media (see above).

YEME (10% sucrose): YEME was used for the cultivation of *Streptomyces* spp. for genomic DNA extraction, as well as for the growth of strains for storage as mycelial stocks. In 1 L of distilled dH_20 3 g yeast extract, 5 g Difco Bacto-peptone, 3 g Malt extract, 10 g glucose, and 100 g Sucrose was dissolved. After autoclaving and prior to use, 2 ml/ L of 2.5M $MgCl_26H_20$ was added.

YT medium (2x): YT medium is a rich medium used for the heat-shocking of *Streptomyces* spores for conjugation. 16 g Difco Bacto-Tryptone, 10 g Difco Yeast Extract, and 5 g NaCl in 1 L of distilled dH_20. Aliquoted (10 ml) and autoclaved.

2x Pre-germination medium: Used for the pre-germination of *Streptomyces* spores prior to liquid growth. In 50 ml of distilled dH_20, 0.5 g yeast extract and 0.5 g Difco casamino acids were dissolved and autoclaved. After this $CaCl_2$ (5 M), that had been previously autoclaved, was added to this mixture to a final concentration of 0.05 M, before being used for the pre-germination of spores.

Solid media

Bennetts agar: Used for the phenotypic characterisation of *Streptomyces* strains. In 1200 ml of distilled dH_20 1.3 g Beef extract, 12 g of glycerol, 2.4 g Bacto casein, 1.2 g yeast extract and 18 g of agar were added at a final pH of 7.3 before being separated into 200 ml aliquots and autoclaved before use.

Mannitol soya flour (MS) agar: In 1 L of tap water, 20 g of mannitol and 20 g of soya flour were dissolved, before added to 250 ml Erlenmeyer flasks containing 1.5 g agar. Flasks were stoppered using a foam bung and foil and were autoclaved twice. When required, flasks were heated to melt the agar.

Basic minimal medium agar: Used for the phenotypic characterisation of *Streptomyces* strains. In 1L of distilled dH_20 0.5 g of L-asparagine, 0.5 g of K_2HPO_4, 0.2 g of $MgSO_47H_20$, 0.01 g of $FeSO_47H_20$, 10 g of agar, before dividing into 200 ml aliquots and autoclaving. Before use 4 ml of 50% glucose solution was added to the autoclaved 200 ml aliquot.

R2 medium: Used for the phenotypic characterisation of *Streptomyces* strains. Noted to increase actinorhodin production. To 800 ml, 103 g sucrose, 0.25 g K_2SO_4, 10.12 g $MgCl_2$. $5H_20$, 10 g glucose and 0.1 g of Difco casamino acids was mixed, before 80 ml of this was decanted into 250 ml flasks containing 2.2 g of agar and autoclaved. Prior to use, 1 ml of KH_2PO_4 (0.5%), 8 ml $CaCL_2.2H_20$, 1.5 ml 20% L-Proline, 10 ml TES buffer (5.73%, pH 7.2), 0.5 ml Trace element solution (see below) and 0.5 ml NaOH (1M).

Trace element solution (R2 medium): In 1 L of dH_20, 40 mg $ZnCl_2$, 200 mg $FeCl_3.6H_2O$, 10 mg $CuCl_2.2H_2O$, 10 mg $MnCl_2.4H2O$, 10 mg $Na_2B_4O_7.10H_2O$, and 10 mg $(NH_4)_6Mo_7O_{24}.4H_20$.

2.1.7 Antibiotic selection and additives

Table 2.1.7: A list of antibiotics and additives used in this study.

Name	Stock solution	Concentration used in liquid media	Concentration used in solid media
Apramycin (amp)	100 mg/ml dissolved in 60% ethanol	20- 50 µg/ml	25 µg/ml
Carbenicillin (Carb)	100 mg/ml, dissolved in 60% ethanol	100 µg/ml	100 µg/ml
Chloramphenicol (cml)	34 mg/ml dissolved in 100% ethanol	34 µg/ml	25 µg/ml
Hygromycin (hyg)	50 mg/ml dH_20 (filter sterilised)	10 µg/ml	10 µg/ml
Kanamycin (kan)	50 mg/ml dH_20 (filter sterilised)	50 µg/ml	25 µg/ml
Nalidixic acid (nali)	25 mg/ml dissolved in 0.15 M NaOH	25 µg/ml	25 µg/ml
Thiostrepton	50 mg/ml dissolved in DMSO	/	10 µg/ml
X-gal	40 mg/ml, dissolved in Dimethylsulfoxide	/	40 µg/ml
Diamide	0.5 M dissolved in dH_20 (filter sterilised)	0.5 mM	/

2.1.8 Solutions and Buffers

2.1.8.1 Nucleic acid manipulation

Plasmid miniprep

- **P1 Buffer:** 50 mM Tris-HCl pH 8.0, 10 mM EDTA, 100 µg/ml RNaseA
- **P2 buffer:** 200 mM NaOH, 1% SDS
- **N3 buffer:** 4.2 M Gu-HCl, 0.9 M potassium acetate, pH 4.8
- **PE buffer:** 10 mM Tris-HCl pH 7.5, 80% ethanol

gDNA extraction

- **TE25S buffer:** 25 mM Tris-HCL (pH 8.0), 25mM EDTA pH8, 0.3M sucrose.
- **Lysis solution:** TE25S buffer, 2 mg/ml lysozyme.
- **2x Kirby mix:** 2% (w/v) TPNS, 12% (w/v) sodium 4-amino-salicilate, 6% (v/v) phenol (pH 8.0) dissolved in 100 mM Tris-HCL (pH 8.3). Aliquots (25 ml) were stored at -20 °C for long term storage and 4 °C otherwise.

RNA extraction:

- **1x Kirby mix:** 1% (w/v) sodium-triisopropulnaphthalene sulphonate (TPNS), 6% (w/v) sodium 4-amino-salicilate, 6% (v/v) phenol (pH 8.0), dissolved in 50 mM Tris-HCL (pH 8.3). Aliquots (25 ml) were stored at -20 °C for long term storage and 4 °C otherwise.
- **3M sodium acetate (pH 6.0)**
- **Stop solution:** 95:5 ethanol:acid phenol.

2.1.8.2 Protein purification

All buffers filter sterilised and stored at 4°C unless specified.

Ni-NTA purification buffers

- **Binding buffer:** 5 mM imidazole, 0.5 M NaCl, 20 mM Tris-HCl (pH 7.9), with 25 μg/ml PMSF added directly before use. For initial lysis, Pierce™ Protease Inhibitor Tablets (EDTA-free) (Thermofisher Scientific) were added to buffer according to manufacturer's instructions.
- **Wash buffer:** Same as binding buffer, plus either 10 mM or 25 mM imidazole.
- **Elution buffer:** 250 mM imidazole, 0.5 M NaCl, 20 mM Tris-HCl (pH 7.9), with 25 μg/ml PMSF added directly before use.

Gel filtration and Anion-exchange purification buffers

- **Binding buffer:** 50 mM Tris-HCl (pH 7.9), 50 mM NaCl, 5% (v/v) glycerol, 5 mM β-mercaptoethanol added directly before use.
- **Anion-exchange Elution buffer:** 50 mM Tris-HCl (pH 7.9), 1 M NaCl, 5% (v/v) glycerol, 5 mM β-mercaptoethanol added directly before use (stored at room temperature).
- **Storage Buffer (gel filtration):** 10 mM Tris-HCL (pH 7.9), 0.1 M KCL, 0.1 mM EDTA. 50% glycerol added after for storage.

Additional protein purification buffers

- **Coomassie quick stain solution:** 40 mg Coomassie G250 was dissolved in 500 ml dH_20 and stirred for ~ 2 hours before 1.5 ml of concentrated HCL was added. Stained was stored at room temperature in a foil-covered bottle.
- **Destaining solution:** 20% (v/v) methanol, 10% (v/v) glacial acetic acid.
- **PMSF (Phenylmethylsulfonyl fluoride):** 10 mg/ml dissolved in isopropanol and stored at -20 °C.
- **20% Ethanol**

2.1.8.3 Western blotting

- **10x TBS-Tween buffer:** 12 g Tris-HCl, pH 7.6, 40 g NaCl and 5 ml Tween 20™ in 500 ml.
- **Alternative blocking solution:** TBS Tween (1x) +5% BSA.
- **Semi – dry transfer buffer:** Tris-Base 5.82 g, Glycine = 2.92 g, 10% SDS 3.75ml and Ethanol = 200ml. dH_20 up to 1L.

2.1.8.4 *In* vitro transcription assays

2.1.8.4.1 Buffers

2x Transcription buffer: 80 mM Tris-HCL (pH 7.9), 20 mM MgCl2, 1.2 mM EDTA, 40% glycerol. Autoclave and add 0.1 volume of 8 mM KH2PO4 (pH 7.5). Prior to use add 1.5 mM DTT and 0.25 mg/ml BSA.

Formamide loading buffer: 80% (w/v) deionized formamide, 1x TBE buffer, 10 mM EDTA, 0.08% (w/v) xylene cyanol, 0.08% Amaranth. Aliquots (800 µL) was stored at -20 °C.

RNAP dilution buffer: 10 mM Tris-HCL (pH 8.0), 10 mM KCL, 5% glycerol (v/v). Autoclave and add BSA to a final concentration of 0.4 mg/ml and 10 µL Triton X. Prior to use add 10 mM β-mercaptoethanol.

2.1.8.4.2 NTP mixes

All NTP mixes were diluted in RNase-free distilled dH_20, with 0.15 µL of either γ-UTP or γ-CTP radionucleotide.

Standard (15x): 1.5 mM CTP, ATP, UTP and GTP

UTP (15x): 1.5 mM UTP

High UTP (15x): 15 mM UTP, 1.5mM ATP, CTP and GTP

No CTP (15x): 1.5 mM ATP, UTP and GTP

No GTP (15x): 1.5 mM ATP, UTP and CTP

Cold CTP chase (15x): 10 mM CTP

Cold UTP chase (15x): 10 mM UTP

2.1.8.5 Gibson assembly

Isothermal reaction mix (5x): 3 ml 1 M Tris-HCL (pH 7.5), 50 mM $MgCl_2$, 1 mM dGTP, dCTP, dTTP and dCTP, 50 mM DTT, 1.5 g PEG-8000, 5 mM NAD and made up to 6 ml with sterile dH_20. Aliquots (1 ml) were stored at -20 °C for future use.

Gibson master mix: 100 µL 5x isothermal reaction mix, 50 µLTaq ligase (40u/ µL), 2 µL T5 exonuclease (1u/ µL), 6.25 µL Phusion polymerase (2u/ µL) made up to 375 µL with nuclease-free dH_20.

2.2 Methods

2.2.1 Growth and storage of bacterial strains

2.2.1.1 *E. coli* growth and storage

E. coli strains were grown for up to 24 h in either liquid or solid media at 37°C. L-broth and agar was used to grown strains in liquid or solid media, respectively. Antibiotic selection (Table 2.1.7) was included when necessary. For liquid growth, *E. coli* were grown in conical flasks at 250 rpm.

For long term storage of *E. coli,* glycerol stocks were made using equal volumes of liquid media containing *E. coli* and 60% glycerol were combined and stored at -80°. Agar plates containing *E. coli* were kept at 4°C for up to two weeks.

2.2.1.2 *Streptomyces* growth and storage

Streptomyces strains were grown at 30°C on both solid and liquid media. MS agar was most commonly used for the growth of sustainable growth of *Streptomyces,* with strains grown for ~5-7 days.

For liquid growth, *Streptomyces* spores (see below) were pre-germinated beforehand. Spores were standardised to an OD450 of 0.06, washed in 500 µL of TES buffer (0.05M, pH 8.0) before centrifuged for 20 seconds at 13,200 rpm and resuspended in 500 µL TES buffer. Spores were then heated at 50 °C for 10 min, and added to a 50 ml falcon tube containing 500 µl of 2x pre-germination medium (Section 2.1.7.2). Spores were then pre-germinated at 37 °C, 250 rpm for ~2-4 hours, before directly adding to siliconized Erlenmeyer flasks containing stainless steel springs and liquid medium required for growth. Flasks were then placed at 30 °C at 300 rpm in an orbital shaker for *Streptomyces* growth.

Note mycelial stocks were also used for the growth of strains defective in sporulation, but again were standardised to an OD450 of 0.06 before directly inoculated for liquid growth as above.

Storage of *Streptomyces* as spore stocks

Streptomyces strains were stored as spore stocks at -20 °C. Stocks were made by plating a confluent lawn of the strain of interest onto MS media (with antibiotic if required) and grown for 4-5 days allowing for sporulation. Sterile distilled dH_20 (10 ml) was then poured onto plates with spores resuspended in the liquid using a flame-sterilised metal loop. Spores in solution were then collected, vortexed and placed through a sterile syringe containing cotton wool and filtered into a sterile 15 ml falcon tube. Spores were then centrifuged for 4000 rpm for 10 min at 4 °C, and the spore pellet resuspended in 1 ml of 20% glycerol before being transferred to a sterile Eppendorf tube.

Storage of *Streptomyces* as mycelial stocks

Storage of mycelial fragments was used for *Streptomyces* strains defective in sporulation (M571) and for longer term storage. The strain of interest was streaked and resuspended in sterile dH_20 as seen above for storage as spore stocks. Five millilitres of this mycelial suspension was then inoculated into 50 ml of YEME (10%), and grown in a Erlenmeyer siliconized sprung flask for ~1-2 days at 30 °C, with shaking 300 rpm. Once the strain had reached stationary phase (OD450 of ~3+), mycelial fragments were transferred to 50 ml falcon tubes and pelleted by centrifugation (2000 g for 10 min). Pellets were then resuspended in 5 ml of 20% glycerol and placed into sterile Eppendorfs as 500 µL aliquots before being stored at -80 °C for future use.

2.2.1.3 Induction of oxidative stress response during liquid growth of *Streptomyces*

Streptomyces strains of interest were grown as above (Section 2.2.1.2) using either spore stocks or mycelial fragments at a starting OD450 of 0.06 in 60 ml of NMMP medium in Erlenmeyer siliconized sprung flasks. Strains were grown overnight to exponential phase (OD450 of ~ 0.8-1.0) before harvesting 6 ml of culture into a 15 ml falcon tube and centrifuged (4000 rpm for 1 min). Supernatant was removed and pellet was directly resuspended in either 1ml kirby mix (for continuation with RNA extraction protocol, Section 2.2.3.3), or 1 ml of RNA protect (Qiagen, 76506) allowing storage at -20 °C before further RNA extraction later.

For induction of the oxidative stress response, 0.5 mM diamide was quickly added to cultures before flasks were returned to growth at 30 °C. Samples from cultures were then extracted and prepped for RNA extraction using the above method 5, 10, 30 and 60 min after the addition of diamide.

2.2.1.4 Induction of a stringent stress response during liquid growth of *Streptomyces*

Streptomyces strains of interest were grown as above (Section 2.2.1.2) using either spore stocks or mycelial fragments at a starting OD450 of 0.06 in 60 ml of NMMP medium (containing amino acids). Strains were grown overnight to exponential phase (OD450 of ~ 0.8-1.0) before harvesting 8 ml of culture (representing time 0) into a 15 ml falcon tube containing stop solution (5:1 ratio) (Section 2.1.8.1) for 5 min at room temperature. Tubes were then centrifuged (4000 rpm for 1 min), before the supernatant was removed and pellets immediately flash frozen in liquid nitrogen and stored at -80 °C. Pellets were then thawed briefly before being resuspended in kirby mix and RNA extracted following protocol in Section 2.2.3.3 .

Prior to this extraction, NMMP lacking amino acids (Section 2.1.7.2) was pre-warmed to 30 °C in a water bath. After initial extraction of *Streptomyces* at time zero, the rest of the culture (~50 ml) was extracted into a 50 ml falcon tube and centrifuged at 4000 rpm for 1 min. Supernatant was quickly removed and the cell pellet was washed by resuspension in prewarmed NMMP medium (lacking amino acids). Cells were then centrifuged, washed again as above, before finally being resuspended in 50 ml of prewarmed NMMP medium, and transferred to a fresh prewarmed siliconized sprung flask. Flasks were immediately returned to growth at 30 °C and a timer was set. After 5, 10, 15, 20, 30 and 60 min, a further 8 ml was quickly extracted and prepped for RNA extraction as above, before returning flasks to shaking incubator for extraction at subsequent time points.

2.2.2 DNA manipulation

2.2.2.1 Restriction digest of DNA (and dephosphorylation)

Digestion was typically performed on ~1 µg plasmid DNA, 1x restriction buffer and 10 units of high-fidelity restriction enzymes when possible. Reactions were left to digest at 37 °C for ~1-3 hours.

For simultaneous dephosphorylation of DNA, 1 unit of rSAP (Shrimp Alkaline Phosphatase, NEB) was added to digestion mixtures before incubation at 37 °C as above. After reactions, rSAP, and if possible restriction enzymes were heat-denatured according to stated protocols.

2.2.2.2 DNA ligation

Ligations were performed using insert and vector DNA, 1x T4 ligase buffer and 0.2 units of T4 ligase (NEB) in a final volume of 10 µL. A 10:1 or 5:1 insert:vector ratio was used for blunt-end and sticky-end ligations, with ligations carried out a room temperature for 1 hour and 16 °C overnight respectively.

2.2.2.3 Polymerase Chain Reaction (PCR)

All reactions were carried out in 0.2 ml thin-walled reaction tubes using the following components specific to those required with Phusion DNA polymerase (NEB):

Component	**Volume (total = 50 µl)**	**Final concentration**
Nuclease free water	30 µL	
5x Phusion HF or GC buffer	10 µL	1x
10mM dNTPs	1 µL	200 µM/µl
10 µM Forward primer	2.5 µL	0.5 µM/µl
10 µM Reverse primer	2.5 µL	0.5 µM/µl
DMSO (100%)	1.5 µL	3%
Phusion DNA polymerase	0.5 µL	1.0 units/50 µl PCR
Template DNA (colony or 25ng plasmid)	2 µL	1ng/µl

PCR reaction conditions

Step	Temperature (°C)	Time	No. of Cycles
Initial Denaturation	98	30 secs	1
Hot start	80		1
Denaturation	98	10 secs	30
Annealing	(Tm -4°C) (45-72)	30 secs	
Extension	72	30-60 secs per kb	
Final Extension	72	10mins	1
Hold	4-10		

Annealing temperature was determined using the melting temperature (Tm) of primers used in reaction. Extension time was also determined using the length of PCR product expected. For standard purification of PCR products, the QIAquick PCR purification kit was used according to manufacturer's instructions.

2.2.2.4 Inverse PCR

Prior to PCR, primers were phosphorylated using T4 PNK (polynucleotide kinase) following the standard protocol (NEB). Reactions were incubated at 37 °C for 30 min before T4 PNK was heat inactivated at 65 °C for 20 min. Primers were then used in PCR reaction mixtures, which were assembled as above, before added to the thermal cycler for amplification using the following conditions:

Step	Temperature (°C)	Time	Cycles
Initial denaturation	98	1 min	x1
Melting	98	1 min	X10 cycles
Annealing	Tm -4(45-72)	30 seconds	
Extension	72	2.5min per kb	
Melting	98	1 min	X10 cycles
Annealing	Tm - 4°C (45-72)	30 seconds	
Extension	72	3.5min per kb	
Final extension	72	15 mins	X1
Hold	4		

The PCR products were then taken forward for analysis using gel electrophoresis (Section 2.2.2.7).

2.2.2.5 PCR from *S. coelicolor* colonies

For amplification of DNA directly from *Streptomyces*, a colony of interest was picked from an MS agar plate, using a sterile loop, and immediately resuspended in 50 µL of 0.2% SDS in TE buffer (10 mM Tris-HCL, 1mM EDTA pH 8.0) in a sterile Eppendorf tube. Tubes were then placed at 90 °C for 10 min, before being left to cool slightly. Tubes were then centrifuged briefly to pellet cell debris before 2 µL of supernatant was then directly added to PCR reactions and protocol followed as in Section 2.2.2.3.

2.2.2.6 Gibson assembly

Reactions were carried out in 0.2 ml thin-walled reaction tubes using the Assembly master mix as seen in Section 2.1.8.5. To 15 µL of Gibson master mix, 5 µL of DNA was added before reactions were incubated at 50 °C for 1 hour. Five microlitres of reaction was then transformed as normal (Section 2.2.4.3).

2.2.2.7 Gel electrophoresis and gel extraction

Most DNA fragments were separated using gels consisting of 0.8% agarose in 1xTAE. Separation of DNA fragments was typically carried out at 150 V for ~45 – 60 min. Gel extraction was carried out on separated DNA using the Monarch® DNA Gel extraction kit according to manufacturer's instructions.

2.2.2.8 Annealing of oligos

Five microlitres of each oligonucleotide (100 μM) being annealed were added to an 1.5 ml Eppendorf tube containing 90 μL of HEPES (30 mM, pH 7.8). Tubes were then heated to 95 °C for 5 min with tubes left to cool in heat block to room temperature.

2.2.3 Nucleic acid extraction

2.2.3.1 Plasmid purification from *E. coli*

All buffers stated in Section 2.1.8.1, were made and filter sterilised before use. Bacterial pellets from 5 ml of culture were resuspended in 250 μL of P1 buffer and transferred to a 1.5 ml Eppendorf tube. A further 250 μL of P2 and 350 μL of N3 buffers were added to tubes, and mixed by repeated inversion. Tubes were then centrifuged for 10 min at max speed (13,200 rpm) before the supernatant was carefully removed, to avoid disruption of cell debris pellet, and transferred to a miniprep column (~ 750 μL). Columns were then centrifuged at max speed for 1 min. Columns were then washed using 750 μL of PE buffer, before centrifuged again and flow-through discarded. Columns were centrifuged again to remove remaining liquid before column was transferred to a fresh 1.5ml Eppendorf tube. Plasmid DNA was eluted from the column using 25-50 μL of sterile double distilled H_20, with concentrations determined using a nanodrop 1000 (Thermofisher Scientific).

For larger extraction of plasmids, a plasmid midiprep kit (Qiagen) was used as per manufacturer's instructions, with concentrations of plasmid determined as above.

2.2.3.2 Chromosomal DNA purification from *S. coelicolor*

Streptomyces strains were grown as specified in Section 2.2.1.2 in 50 ml of YEME (10% sucrose) supplemented with 0.5% glycine, for 46 h at 30 °C with shaking (300 rpm). Cells were transferred to a sterile 50 ml falcon tube and centrifuged for 3 min at 4000 x g, before washing twice using 10 ml of 10% sucrose solution and centrifuged again as above. After washing, *Streptomyces* pellets were resuspended in 6 ml of lysis solution (Section 2.1.8.1) and incubated at 37 °C for 10 min. After this, 4 ml of 2x kirby mix was added and tubes were vortexed for 1 min, before 8 ml of phenol/chloroform/isoamyalcohol was added and mixtures vortexed for a further 15 seconds. Tubes were then centrifuged for 10 min at 2000 x g, with upper phase extracted to a fresh tube containing 3 ml of phenol/chloroform/isoamyalcohol and vortexed for 1 min. Tubes were then centrifuged again, and the top layer extracted as above into a fresh tube containing 1 volume of isopropanol and 0.1 volume of 3M sodium acetate (pH 6.0), and mixed. Tubes were again centrifuged at 4000 x g for 5 min before the DNA pellet was washed using 5 ml of 70% ethanol. All ethanol was removed, and pellet was air-dried before dissolving in 5 ml of TE buffer containing 10 μg/ml of RNase and incubating at 37 °C for 30 min. Extraction of DNA performed using isopropanol and sodium acetate, and washing of DNA pellet in 70% ethanol was performed again as above, before final air-dried pellet was resuspended in 500-1000 μL of TE buffer and stored at -20°C.

2.2.3.3 RNA extraction from *S. coelicolor*

Pellets of *Streptomyces* mycelia previously harvested were resuspended in 1 ml of Kirby mix before being transferred to a sterile 2ml Eppendorf containing 100 μL of acidified phenol/chloroform mix. Cells were then lysed using sonication on ice in cycles for 2 x 5 seconds and 2 x 3 seconds at 30 kHz for 5 seconds. A further 300 μL of acidified phenol/chloroform was added to the lysed cells for grinding using a mixer mill for 2 min before centrifuging at 4 °C at 13,200 rpm for 5 min. The upper phase was collected and transferred to another sterile 2 ml tube containing 300 μL of phenol/chloroform for another extraction of the nucleic acids; the tubes were vortexed for 2 min before being centrifuged again as above, with 900 μL of the upper phase transferred to another sterile 2ml Eppendorf tube containing 90 μL 3M sodium acetate (pH 6.0) and 900 μL isopropanol, and mixed. Tubes were then stored for 1 h at -20 °C

for the precipitation of RNA before centrifuging for 10 min at 13,200 rpm at 4 °C and discarding the supernatant. RNA pellets were washed with 70% ethanol before the pellet was resuspended in 400 µL of RNase free water for a further phenol: chloroform extraction using 200 µL of phenol: chloroform. Samples were then vortexed and centrifuged as above before 350 µL of upper phase was transferred to a clean tube containing 350 µL isopropanol and 35 µL of sodium acetate and stored at -20 °C for 1 hour, for a second precipitation of the RNA. Samples were then again centrifuged and washed with 70% ethanol as above, before RNA pellets were dried in a sterile fume hood for 10 min and resuspended in 100 µL of RNase-free water. RNA concentrations were measured using the nanodrop, with A260/A280 and A260/A230 values recorded for determination of RNA integrity. All samples were stored at -80°C and further taken forward for subsequent DNaseI treatment and cDNA synthesis (see Section 2.2.5.1).

2.2.3.4 Preparation of RNA for Cappable RNA-seq (Vertis-Biotechnology)

RNA from *Streptomyces* strains M145 and M571 was isolated for further characterisation using the cappable RNA-sequencing method (Ettwiller *et al.*, 2016), to determine the 5' end, and thus TSS, during both exponential growth and stringent conditions. Growth and induction of stringent conditions was induced in biological triplicate, as seen in Sections 2.2.1.2 and 2.2.1.4 respectively, with samples taken before (time 0) and 10 min after the induction of stress. *Streptomyces* pellets stored at -80 °C were briefly thawed before RNA extracted as seen in Section 2.2.3.3 and further purified using the RNA cleanup kit (Monarch®).

Before samples were sent to sequence, RNA purity and integrity, using RIN (RNA integrity number) values was determined using the Aglient RNA 6000 Nano kit and bioanalyser according to manufacturers. Pure RNA was sent to Vertis-Biotechnology according to their instructions. The quality control data can be found in appendix Section 7.1.

2.2.3.5 Determination of DNA and RNA concentrations

The nanodrop 1000 (Thermofisher Scientific) was used for rapid quantification of nucleic acid concentration and purity, with both the A260/A280 and A260/A230 values noted. For more accuracy, Qubit based assays for RNA and DNA were also used with the Qubit 2.0 fluorometer according to manufacturer's instructions (Thermofisher Scientific).

2.2.4 Introduction of DNA into *E. coli* and *Streptomyces*

2.2.4.1 Preparation of chemically competent *E. coli*

E. coli cells were restreaked to single colonies from frozen aliquots (-80 °C) onto fresh L-agar plates and left to incubate overnight at 37 °C. From plates a single colony was used to inoculate a 5 ml of L-broth (with correct selection) in a 50 ml falcon tube and grown overnight with shaking at 250rpm at 37 °C. Two millilitres of overnight culture was then used to inoculate 200 ml of fresh L- broth and grown as previously to an OD600 of ~0.4-0.6. Cultures were then centrifuged at 4000 rpm for 10 min at 4 °C before pellets washed with 40 ml of ice-cold 100 mM $MgCl_2$ + 10% glycerol, and centrifuged again as above. Cell pellets were then resuspended in ice-cold 40 ml of 100 mM $CaCl_2$ + 10% glycerol and left on ice at 4 °C overnight. Cells were pelleted by centrifugation as above and resuspended in 2 ml of 100 mM $CaCl_2$ + 10% glycerol, rapidly placed into Eppendorf tubes as 50-100 μL aliquots, and immediately flash-frozen in liquid nitrogen. Tubes were then stored at -80 °C until required.

2.2.4.2 Preparation of electrocompetent *E. coli*

E. coli cells from a single colony on a fresh L-agar plate were inoculated into 5 ml L-broth, with correct selection, for overnight growth at 37 °C and 250 rpm. Five millilitres of overnight culture was then added to 50 ml of fresh L-broth (+ selection) and grown as above to an OD600 of ~0.4-0.6. Cells were then harvested by centrifugation at 2000 x g for 10 min before the pellet was resuspended in 50 ml of chilled 10% glycerol. Two subsequent washes were carried out on cells with

centrifugation as above, removal of supernatant and resuspensions in 25 ml and 10 ml 10% glycerol consecutively. After a final centrifugation, supernatant was removed and the pellet resuspended in residual volume (~1.5 ml). Cells were quickly placed into tubes in 50 µL aliquots, flash-frozen and stored at -80 °C for future use.

2.2.4.3 Transformation into chemically competent *E. coli*

Cells prepared previously (Section 2.2.4.1) were placed on ice for 5 min to thaw. After this 5 µL of ligation mix or DNA (final concentration 1- 25 ng) was added to 50 µL of cells and left to incubate on ice for 30 min. The mixture was heat shocked at 42 °C for 45-60 seconds in a water bath, before returned to ice for 2 min. L-broth (950 µL) was then added to mixtures and incubated at 37 °C. Cells were pelleted by brief centrifugation at max speed (13200 rpm), supernatant was removed aside from 50 µL of residual broth used to resuspend cell pellet. Cells were plated onto L-agar, containing correct selection, and left to grow overnight at 37 °C.

2.2.4.4 Electroporation into *E. coli*

Cells made previously (Section 2.2.4.2) were placed on ice for 10 min to thaw, alongside the 1 mm electroporation cuvettes required. Cells (50 µL) was mixed with plasmid DNA and left on ice to incubate for 5 min, before transferred to a cooled electroporation cuvette. Electroporation was carried out at 2,500 volts, 400 Ω and 330 µF using the Eppendorf Electroporator 2510 system, with 950 µL of SOC medium added immediately after transformation. Cells were then placed at 37 °C with shaking (250 rpm) for 1 hour to recover, before plated and spread onto pre-warmed L-agar plates containing correct selection as seen in Section 2.2.4.3. Note positive control reactions using Puc19 plasmid DNA were used to confirm DNA uptake by cells.

2.2.4.5 Conjugation of DNA into *Streptomyces*

Sufficient transfer of DNA into the *Streptomyces* genus relies on specific non-methylating *E. coli* strains ETZ or ETR (Section 2.1.6.1), that contain the plasmid DNA of interest (by previous transformation). Overnight cultures of *E. coli* strains were reinoculated into fresh L-broth with correct antibiotic selection and grown at 37 °C with shaking (250 rpm) to an OD600 of ~0.4-0.5; ETR and ETZ strains both

contain the antibiotic for plasmid of interest as well as either Carbenicillin (100 μg/ml) or Kanamycin (50 μg/ml), respectively. Cells were then centrifuged for 3 min at 4 °C for 3 min before washed with 15 ml of L-broth. This wash was again repeated before *E. coli* cells were resuspended in 1ml of L-broth.

For each conjugation reaction, *Streptomyces* spores (10 μL, ~10^7 spores) were added to 500 μL f 2x YT medium (Section 2.1.7.2) and heated for 10 min in a 50 °C water bath, before cooling to room temperature for ~ 10 min.

Equal volumes (500 μL) of spores and *E. coli* were combined, centrifuged at 13,200 rpm for 20 seconds, before supernatant was removed and cell pellet resuspended in residual liquid. Mixtures were then plated on to MS agar plates containing 10 mM $MgCl_2$ made previously, and left to incubate overnight for 16 hours. The next day, plates were overlayed with 1 ml of sterile water containing antibiotic specific to plasmid DNA of interest, as well as Nalidixic acid (0.1 mg/ plate) to kill *E. coli*. Overlay was left to dry into plates for ~ 15 min before returning to growth at 30 °C for around 3-5 days.

2.2.5 Analysis of nucleic acids

2.2.5.1 qPCR preparation and method

RNA extracted was samples was prepared using both a DNAse treatment before conversion to cDNA. Ten micrograms of each RNA sample following extraction (Section 2.2.3.3) was taken forward for DNAseI treatment using the TURBO DNA-free kit (Invitrogen) according to manufacturer instructions, before RNA concentrations were measured using the nanodrop and stored at -80°C for further use. Select treated RNA samples were further subjected to analysis using gel electrophoresis for confirmation of RNA integrity after treatment (Section 2.2.5.2). All DNase treated RNA samples were taken forward for cDNA synthesis using the High-Capacity RNA-to-cDNA kit (Applied biosystems, 4387406) according to instructions. Reactions were also carried out on each RNA sample omitting both the reverse transcriptase enzyme and buffer (-RT control) to determine the amount of

DNA contamination in RNA samples. Both treated (+RT) and untreated (-RT) nucleic acid samples were taken forward for qPCR analysis.

Primers specific to both gene of interest and reference gene used as an internal control were diluted to 1.5nM concentration before being incorporated into the 15 µL reaction mixture used. Reaction mixtures consisted of 7.5 µL SYBR green, 1.5 µL of each forward or reverse primer, 1.5 µL of H_2O and 3ul of each cDNA sample being tested. The qPCR reaction was carried out in 96 well plates using the StepOnePlus^{Tm} software and Real-Time PCR system (Applied biosystems). Standard curves were produced using known DNA concentrations of either genomic or plasmid DNA, for the determination of copy number of both reference and gene of interest. Each sample was averaged from technical replicates, with further analysis carried out using Microsoft excel and GraphPad prism (version 9.2.0) to calculate significant changes in expression.

2.2.5.2 Gel electrophoresis of RNA

Prior to running of gel, all gel apparatus, including tank, gel cassette and comb were pre-soaked for 30 min in 1% SDS and 0.1M NaOH to eliminate contaminating RNAses. To make the gel, 20 ml of 5x MOPS buffer, 72 ml of RNase-free H_20 and 1.3g of agarose were mixed together and melted in microwave, before 20 ml was placed into a separate sterile 50 ml falcon tube. The mixture was cooled to ~55 °C before 3.5 ml of 37% formaldehyde was gently mixed in, and gel mixture added to gel cassette to set. Gel was pre-run in 1x MOPS buffer for 10 min and run for 30 min after the addition of diluted RNA samples (see below) at 100 V. The gel was visualised under high UV after staining with SYBR safe (Thermofisher Scientific) according to manufacturer's protocol.

RNA samples were prepped for electrophoresis by diluting 3 µL in 6 µL of RNA sample buffer (Section) and heated to 60 °C for 5 min. RNA samples were then cooled on ice for 2 min before the addition of 2 µL RNA loading buffer (Section 2.1.8.1).

2.2.6 Protein purification and detection

Note all buffers used for the differing protein purification methods were ice-cold unless stated.

2.2.6.1 Protein expression in *E. coli*

Prior to expression, Rosetta II strains were transformed with expression plasmids (e.g pSUSH) (Section 2.2.4.3) and grown overnight from a single colony in 10 ml L- broth (with selection) at 37 °C with shaking (250 rpm). The next day, 500 µL of culture was used to reinoculate 50 ml of fresh L- broth (with selection) in a 250 ml flask and grown to an OD600 of ~0.6-1.0. To increase protein expression, a cold shock on the strains was carried out by placing flasks into an ice bath for ~ 15 min with intermittent shaking by hand. IPTG (0.5 mM) was then added to *E. coli* and flasks were returned to grow at 37 °C for ~3-4 hours. Note if proteins were more difficult to express, cultures were left overnight at 18 °C. Cells were harvested for protein extraction by centrifugation at 3,500 g for 20 min, washed in ice-cold cell-wash buffer (Section 2.1.8.2) and centrifuged again as previously. Cell pellets were either flash-frozen for storage at 80 °C, or directly taken forward for cell lysis (Section 2.2.6.2).

2.2.6.2 Preparing cell lysate

Cell pellets, were gently resuspended in 1 ml of cold wash buffer, before an additional 9 ml of buffer was added. Cells were then sonicated on ice in pulses (1 second on, 1 second off) at Amp 40% for 1 min, before being transferred to high speed centrifuge tubes and centrifuged at 15, 000 rpm for 50 min, pelleting the cell debris and insoluble protein fraction. The supernatant was then carefully separated into a sterile tube (remaining on ice) before continuing with column purification (Section 2.2.6.3).

Note that 50 µL samples were taken at each step (before and after lysis) for SDS-PAGE analysis to confirm purification of protein of interest (Section 2.2.6.8)

2.2.6.3 Ni-NTA Sepharose affinity chromatography

His-pur™ Ni-NTA resin (Thermofisher Scientific) was placed into a syringe barrel, sealed with glass wool, giving a final column volume (CV) of 2-4 ml. Before use, the column was washed with 10CV of wash buffer. After this, the cap was placed on the

column and the supernatant, containing protein mixture, was added to the resin and mixed before being placed into a sterile falcon tube. Tubes were then incubated on a rotator at 4 °C for 60 min at 20 revs per min. Tubes were then centrifuged for 10 min at 4000 rpm, before supernatant was discarded. The pelleted resin was then washed by resuspension in 10 CV wash buffer and centrifuged again as above. The supernatant was again removed and pellet resuspended in a further 10 CV wash buffer before added to the column. Protein was then eluted from the resin by addition of 5 CV elution buffer and collected in several Eppendorf tubes containing around 1 ml each, and placed immediately on ice. Again 50 μL samples were collected through purification and run on SDS-page for analysis.

For reuse of the column, 10 CV of wash buffer and 5 CV of dH20 was added consecutively to wash the column. 5 CV of 20% ethanol was then added to the column, with half allowed to flow-through, before cap and lid was replaced and the column stored at 4 °C for future use.

2.2.6.4 Anion-exchange chromatography

The Mono QTM 5/50 GL column (GE healthcare) has a CV of 1 ml, a flow-rate of 0.5-3 ml/min and a maximum pressure of 4.2 MPa, and purification was carried out using the AKTA purified box-900 (Amersham). Both pumps were first washed with distilled dH_20, with both then washed either with monoQ binding or elution buffer (Section 2.1.8.2) before the column was connected and washed with 5 CV of binding buffer. Simultaneous to this, the injection loop was washed with 2CV of dH_20 and then 2CV of buffer, before sample was added to the loop. Protein was then injected on to column before eluted on a gradient from 30 – 100% elution buffer. Fractions were collected (1 ml) and samples with significant 280 nm peaks visualised using SDS-PAGE (Section 2.2.6.9).

2.2.6.5 Gel filtration chromatography

The HiLoad™ 16/600 superdex™ 75 pg gel filtration column (Cytivia™) has a CV of 120 ml, and a maximum flow-rate and pressure of 1 ml/min and 0.3 Mpa respectively. The column and pumps were washed with 1 CV of storage buffer (or gel filtration buffer) with the injection loop also washed with 1 CV of dH_20 and then 1 CV storage

buffer. Protein was then injected and loaded onto column before collected as 2 ml elutions using a flow-rate of 1 ml.

2.2.6.6 SUMO/Ulp protease digestion

Using a previously purified stock of SUMO protease (~3.6 mg/ml, Hare, Pers comms), protein of interest was cleaved from the SUMO protein by adding 1 mg of protease for every 30 mg of protein and a final concentration of 2 mM DTT. Mixtures were then incubated on a rotator at 4 °C overnight at 20 revs per min for sufficient cleavage.

2.2.6.7 Protein buffer exchange

To exchange buffers of protein samples SnakeSkin™ dialysis tubing with a 3.5K MWCO (Thermofisher Scientific) was used. The dialysis membrane was initially cut to the right size and hydrated by placing into a large beaker containing 500 – 1000 ml of buffer (to be exchanged into) for ~ 5 min with stirring. After this, a sterile clip was placed at the bottom of the tubing tightly to ensure no leakage of protein. Protein sample was then added inside the tubing and sealed by placing another clip at the top of tubing to completely seal sample within the membrane. A foam bung was added to the top clip, and the mixture placed into the large beaker of buffer. Protein was left to dialyse overnight at 4 °C with stirring at a low speed, before protein was carefully removed from tubing and placed into a sterile falcon tube on ice.

2.2.6.8 Concentration of protein samples

Concentration of protein was carried out using Pierce™ Protein concentrators, 3K MWCO (20 ml) (Thermofisher Scientific) according to manufacturer's protocol.

2.2.6.9 Protein sample analysis by SDS-PAGE separation

For the separation and analysis of protein purification samples by SDS-PAGE, 4-12% NuPAGE gradient Bis-Tris polyacrylamide gels (Invitrogen) were loaded and run for 1 h 20 min at 120 V. Before loading protein samples were diluted in 1x NuPAGE sample buffer according to manufacturer's instructions.

Bands were then visualised by Coomassie staining, by removing gel from plastic cassette and placed into a small container filled with dH_20 to cover the gel, before

heating in the microwave for 30 seconds and discarding the dH_20, retaining the gel in the box. This was repeated 3 more times, before covering the gel with Coomassie quick stain (Section 2.1.8.2) and heated again for 30 seconds. Stain was discarded into a waste bottle, and gel rinsed in dH_20. Gel was either examined immediately, or submerged in destainer solution (Section 2.1.8.2), before viewing bands on a light box.

2.2.6.10 Determination of protein concentration

Both the Pierce™ BCA protein Assay kit (Thermofisher Scientific) and Qubit Protein broad range assay kit (Invitrogen) were used to determine protein concentration and carried out according to manufacturer's protocol.

2.2.7 Western blotting

Protein samples, required for blotting, were run using SDS-PAGE gel electrophoresis (Section 2.2.6.8), before lanes and additional gel were cut and removed. Prior to this, Protran® nitrocellulose membrane (0.2 μm, Whatman) was cut to a size slightly larger than that of the gel, and pre-soaked in chilled transfer buffer (Section 2.1.8.2), between two sheets of filter paper (3 mm) of similar size, for ~ 5 min. After this, filter paper was rolled out, to remove any excess liquid and the gel and membrane were assembled in the order, filter paper, gel, membrane, filter paper, before placed into the semi-dry cassette with the gel on top of the membrane (transfers down). Transfer was then carried out for 30 min at 100 V, using the Trans-Blot Turbo Transfer System (Biorad), before membrane was washed with deionized water, and stained with panceau for ~ 1 min to confirm correct transfer. Membrane was then cut using stained bands as a guide, before incubating in blocking solution at room temperature for ~1 hour with rocking. Membranes were then washed x 3 in 1x TBS Tween (Section 2.1.8.3) for 10 min, incubated overnight at 4 °C with primary antibody of correct dilution, washed again as above, before incubation with secondary antibody for 1 hour at 4 °C. For ECL detection, membranes were again washed as above, before the addition of 1 ml ECL working solution (Pierce™ ECL Western blotting substrate, Thermofisher Scientific) to the membrane for 1 min, before excess solution removed, membrane wrapped in saran wrap and exposed for 5 min using gel capture software (DNr Bio-Imaging Systems F-chemiBls 3.2 M machine).

2.2.8 *In vitro* transcription assays

2.2.8.1 Preparation of denaturing urea-acrylamide gels

Both 8% and 24% gels urea-acrylamide based gels were made using up the 40 % acrylamide/Bisacrylamide (19:1) (Thermofisher Scientific) and fresh urea as follows:

	8%	24%
40% Acrylamide/ Bisacrylamide (19:1)	10 ml	30 ml
Urea	21 g (7M final)	21 g (7M final)
TBE buffer (10X)	5	5
ddH_20	Up to 50 ml	Up to 50 ml

All components were mixed using a magnetic stirrer, with slight heat (~37 °C) to ensure urea was dissolved. Mixtures were then filtered before the addition of 50 µL TEMED and 500 µL fresh 10% APS and immediate pouring into previously assembled glass plates. After this, the comb was quickly added and the cassette placed flat to allow gel to set for ~ 2 hours. Gels were pre-run at 600 V 50 W for 30 min at 55 °C using the AE-6141E MiniQuencer (Atto), before gel run using same voltage for either 1 hour 20 min or 2 hours 40 min for 8% and 24% gels respectively. Note that 8% gels were then transferred to Whatmann 2 MM paper and dried at 80 °C for 45 min before leaving to expose to plates for phosphorimaging. The 24% gels were not dried due to cracking, and remained on one of the glass plates when exposing to phosphoimager plates. Exposure time was determined by age/radioactivity of the radionucleotide, but ranges from 2 hours – 24 hours at room temperature and visualised using a typhoon scanner. Gel images were edited and annotated using either ImageJ or Image Studio Lite.

2.2.8.2 *In vitro* transcription assays

Prior to experiments, DNA templates were amplified using PCR (Section 2.2.2.3), purified using gel extraction (Section 2.2.2.7) and diluted to a final concentration of 100 nM, under RNAse free conditions. rNTPs were also diluted to the correct concentrations required (Section 2.1.8.4.2) and water bath pre-heated to 30 °C prior to assay. Initially σ^{HrdB} and RbpA were combined and incubated on ice for 15 min, after which both RNA polymerase (holoenzyme) and CarD were added to the mixture and reactions incubated for a further 15 min. DNA template was then added to tubes, incubated for 15 min, before the rest of reaction components were added (aside from dNTPs); 2x transcription buffer, water and RNase inhibitor (RNAsin, Promega). Reactions were then pre-incubated at 30 °C for 10 min, before the addition of dNTPs, including the radionucleotide (either α-UTP or α-CTP), to initiate transcription. Assays were carried out for 15 min before the addition of 1 µL cold NTP chase (UTP or CTP), and reactions left for a further 10 min at 30 °C, before terminated by the addition of equal volume of formamide loading buffer. Samples were then heated for 80 °C for 5 min before adding to a gel (Section 2.2.8.1). Note all proteins were diluted to correct concentration using RNAP dilution buffer. A standard 15 µL reaction mixture is as follows:

	Volume (µL)	**Final concentration**
Transcription buffer (2x)	7.5	1x
DNA (100 nM)	0.75	5nM
CarD (7.5 µM)	1	500 nM
RbpA (37.5 µM)	1	2.5µM
HrdB (3.75 µM)	1	250 nM
Holoenzyme (50 nM final) (750 nM)	1	50 nM
Hot NTP mix	1.15	-
RNase inhibitor (RNasin)	0.5	50 U
Water (nuclease-free)	1.1	-
Total -	15	-

2.2.9 Luciferase assays

Strains containing correct pIJ5972 fusions were grown to an OD450 of ~0.8-1.0 in biological triplicate unless specified, with 100 μL of culture extracted in triplicate and placed into tubes compatible with the luminometer (Lumat L-9507, Berthold). Light production was determined by the addition of 100 μL of 1% n-decanal (in 9% ethanol) substrate, with readings recorded over 20 seconds, 5 seconds after the addition of substrate to cultures. RLU values were averaged from the technical triplicates and normalised using OD450 of original culture (Aigle *et al.*, 2000).

2.2.10 Replica plating

Sterilised velvet was pushed on to a previously grown MS agar plate containing around 300 sporulating *Streptomyces* colonies, to allow transfer of spores. The velvet was then pushed onto two Difco nutrient agar plates, firstly onto a plate containing antibiotic selection, the second lacking this selection. Spores were placed initially onto selective plates to ensure any spores not present are due to presence of antibiotic rather than inefficient transfer of spores. Plates were then left to incubate for 1- 2 days at 30 °C, before sensitive and resistant colonies, on original plate were determined.

Chapter 3: Transcriptional organization and stress regulation of the *rpoBC* operon

3.1 Overview

In all bacteria the large β and β' subunits of RNAP are encoded by the *rpoBC* operon, and in most cases it is thought that there is co-transcription with upstream ribosomal protein (r-protein) genes, thereby providing a mechanism to coordinate ribosome and RNA polymerase production. This is the case for several organisms including the model bacterium *E. coli* (Yamamoto and Nomura, 1978; Newman, Linn and Hayward, 1979). However, in *Streptomyces*, work in *S. griseus* suggested that *rpoBC* is expressed independently of upstream r-protein genes and that transcription likely initiates from a single promoter in a large intergenic region (Küster, Piepersberg and Distler, 1998). Since the production of the large subunits is thought to govern the production of RNAP in general (Dykxhoorn, St. Pierre and Linn, 1996), an understanding of how *rpoBC* is controlled may allow further determination as to how *Streptomyces* controls RNAP levels in response to cellular need, and provide insights into how this might be manipulated. This is relevant to understanding secondary metabolite production if mutations in RNAP that stimulate antibiotic production (Lai *et al.*, 2002; Hosaka *et al.*, 2009; Ochi, Tanaka and Tojo, 2014), affect the global transcriptome by decreasing the activity of RNAP at highly expressed growth-related genes, providing a higher level of free RNAP for the expression of genes involved in secondary metabolism (see Section 1.1.3.3). Therefore, the cellular levels of RNAP might play a crucial role in the activation of secondary metabolic gene clusters. However, the mechanisms that underlie RNAP control are poorly understood in most bacteria, and this is especially the case for *Streptomyces*.

This chapter focusses on the transcriptional organisation of the *rpoBC* operon, using available genomic information along with high-resolution promoter mapping data to identify all promoters. The activity of the promoter region is analysed under standard growth conditions as well as in response to both oxidative stress and amino acid starvation. This chapter also studies the 5'UTR of *rpoBC* and investigates its potential regulatory role using mutagenesis.

3.2 Transcriptional organisation of the *rpoBC* region

In *S. coelicolor*, *rpoB* (SCO4654) *and rpoC* (SCO4655) encode the β and β' subunits of ~128.5kDa (1161 residues) and ~144.6kDa (1299 residues), respectively. As is the case in most other bacteria, including *E. coli*, the genes are situated within a cluster of genes, sometimes referred to as the *rif* cluster, that also encodes the genes necessary to transcribe the ribosomal proteins *rplKA* and *rplJL* (Küster, Piepersberg and Distler, 1998); the *rif* cluster name relates to previously identified mutations in this region that confer resistance to the antibiotic rifampicin, however most of these mutations are located in *rpoB* (Xu *et al.*, 2002). The *S. coelicolor rpoBC* operon is situated downstream from *rplJL*, however is believed to be transcribed independently from within the 590bp intergenic region between these genes (Fig 3.2.1).

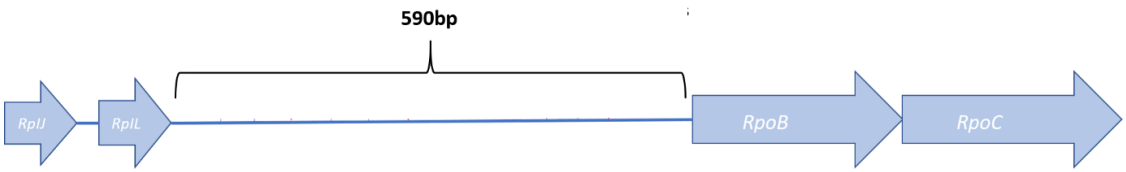


Figure 3.2.1: The genetic organisation of the *rplJL* and *rpoBC* genes in the *Streptomyces* genome (not to scale).

3.2.1 Identification of transcription start sites upstream of *rpoBC*

Previously, in *S. griseus rpoBC* was hypothesised to be under the control of a single promoter located in the large *rplJ-rpoB* intergenic region (Küster, Piepersberg and Distler, 1998). Global mapping of transcription start sites (TSS) in *S. coelicolor* (Jeong *et al.*, 2016a) revealed two start sites (*rpoBp1* at position 5077883 and *rpoBp2* at 5077619, 178bp and 442bp from *rpoB* start codon, respectively), with transcript abundances of ~60% and 40%, respectively.

Since the *rpoB* upstream region is highly conserved and the regulation of *rpoBC* expression is likely to be shared among *Streptomyces*, TSS- mapping sequencing data from *S. venezuelae* was also used to independently analyse the region (ArrayExpress,

E-MTAB-10690; StrepDB). Bush and colleagues (M. Bush and M. Buttner, pers comm) used Cappable-seq (Ettwiller *et al.*, 2016) to identify 5' TSS in RNA isolated after 10 h, 14 h, 18 h and 24 h growth, which represent vegetative, pre-sporulation, onset of sporulation and mid/late sporulation, respectively. Analysis of 18 h data confirmed the presence of *rpoBp1* and *rpoBp2*, along with a third start site, *rpoBp3*, located 522bp upstream from the *rpoB* start codon (Fig 3.2.1.1). As was the case with *S. coelicolor* (Jeong *et al.*, 2016), *rpoBp1* appeared to be the strongest promoter in *S. venezuelae* at 18 h , with ~94% of total transcripts detected initiating at p1.

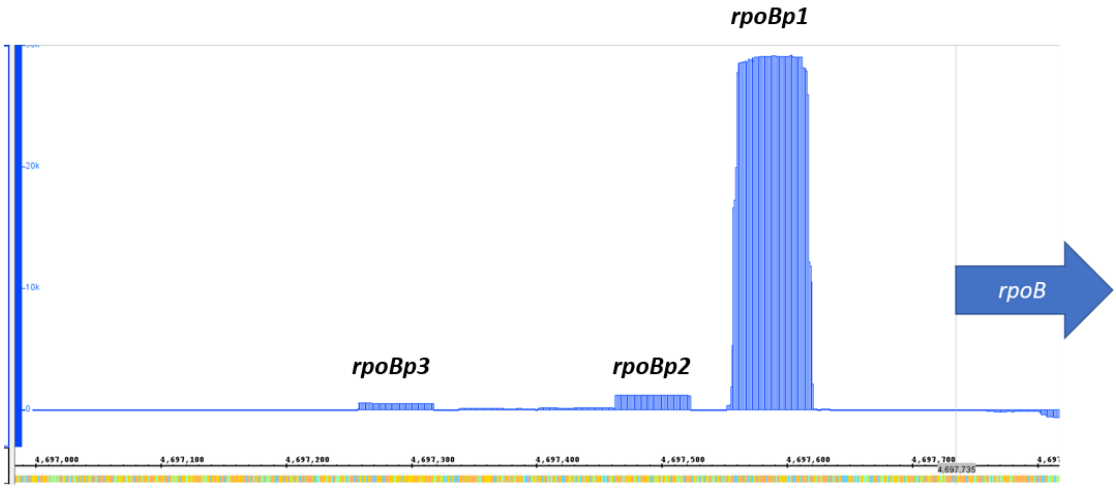


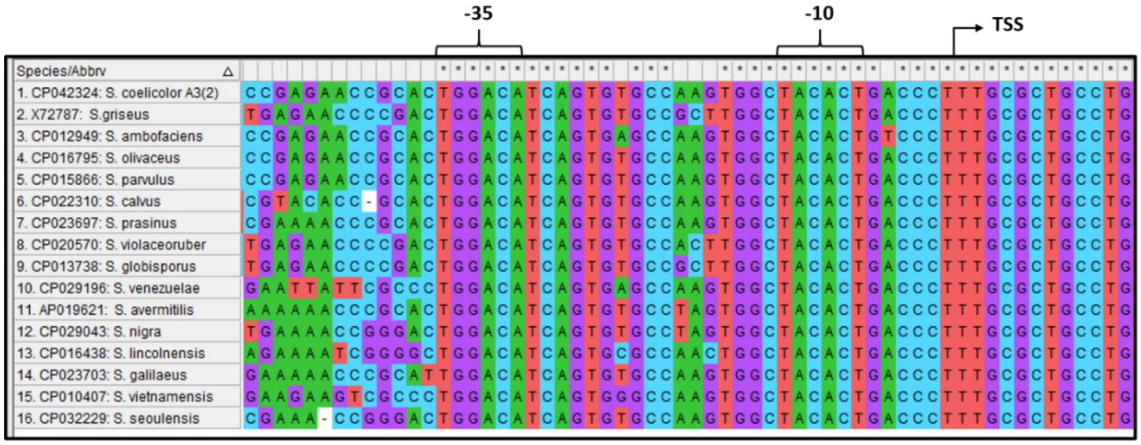
Figure 3.2.1.1: Analysis of TSS mapping reads for the *rpoBC* promoter region. The cappable-seq data for the 18h time point was obtained from Matt Bush (pers comms) and aligned to the *S. venezuelae* genome (NZ_CP018074.1, ~4697000 – ~469780bp) using Integrated Genome Browser (IGB) (version 9.1.8). Blue reads/"peaks" represent RNA aligned to particular area of genome sequence. The start of the *rpoB* reading frame is indicated. Transcription initiates at the upstream extremity of the *rpoBp1*, *rpoBp2* and *rpoBp3* "peaks".

3.2.2: Conservation of the *rpoBC* promoter regions

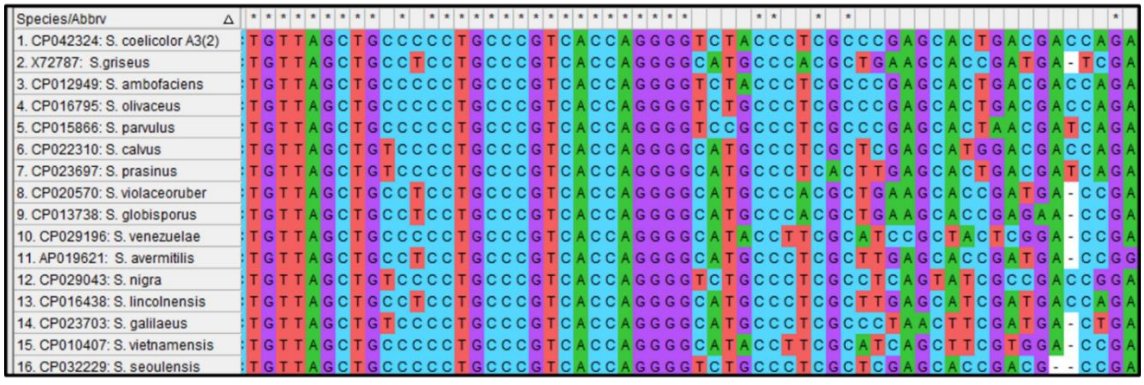
To confirm as well as investigate the conservation of the promoters that control *rpoBC*, the equivalent DNA sequences from 17 *Streptomyces* strains were compared and aligned where possible.

rpoBp1 The *rpoBp1* initiates transcription at a +1T located 178 bp upstream from *S. coelicolor rpoB* and the untranslated region (designated p1U) is conserved in all 17 strains (Fig 3.2.2.1). A very high level of sequence conservation is observed at the +1

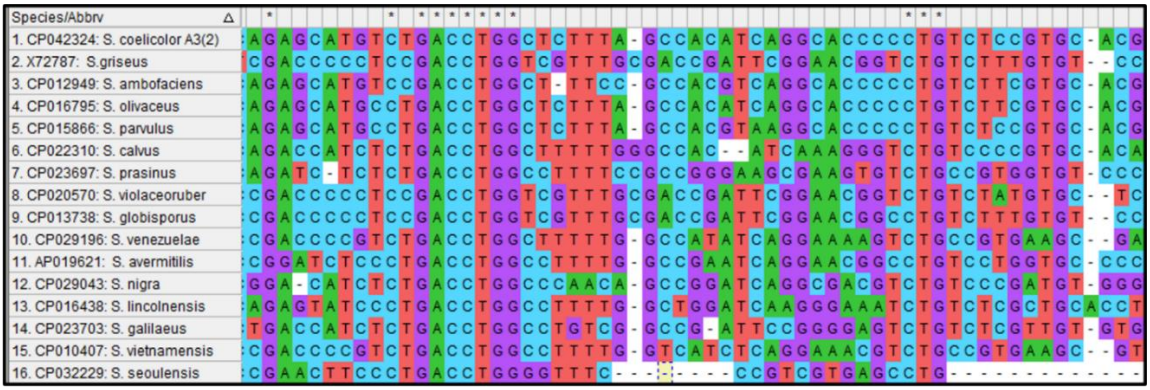
TSS as well as the -35 and –10 motifs, previously identified as 5'TGGACA and 5'TACACT, respectively in *S. griseus* (Küster, Piepersberg and Distler, 1998). These promoter elements are similar to the consensus sequences for vegetative promoters, and hence it may be assumed that the σ^{HrdB} holoenzyme recognises the promoter (Figs 3.2.2.1 and 3.2.2.2) (Jeong *et al.*, 2016a). Indeed, *p1* is transcribed by the σ^{HrdB} holoenzyme, confirmed using *in vitro* transcription assays (Tabib-Salazar *et al.*, 2013). Other notable features of this promoter include an A/T rich region located at ~-45, which might represent an UP-element for α-CTD (Jensen and Galburt, 2021) or a possible binding site for a WhiB-like (Wbl) transcriptional activator (Lilic, Darst and Campbell, 2021), and the highly unusual T-tract at +1, which is the main subject of this work.



1-60bp



58-118bp



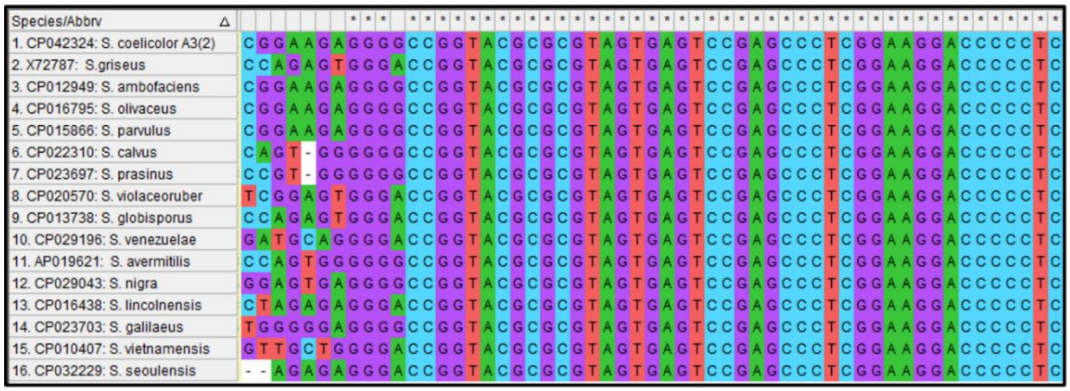
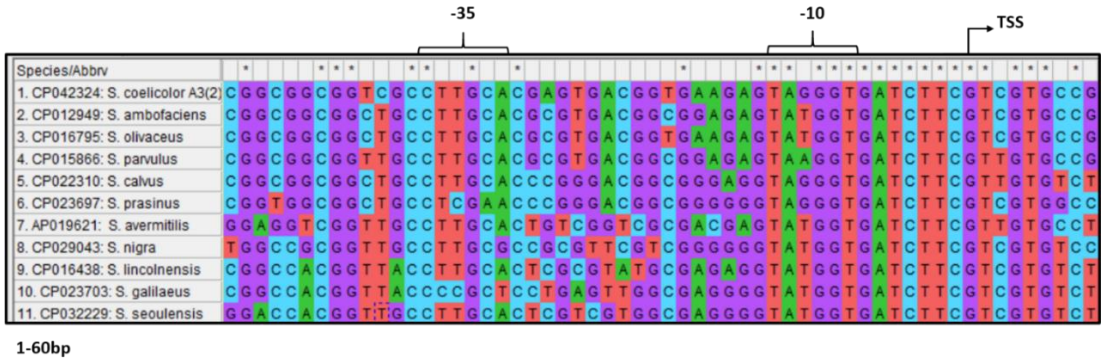


Figure 3.2.2.1: Alignment of the *rpoBp1* promoter and downstream 5'UTR region in the *Streptomyces* genus. The *rpoBp1* promoter of 17 organisms belonging to the *Streptomyces* genus were aligned using CLUSTALW and MEGA-X; *S. coelicolor*, *S. griseus*, *S. ambofaciens*, *S. olivaceus*, *S. parvulus*, *S. calvus*, *S. prasinus*, *S. violaceoruber*, *S. globisporus*, *S. venezuelae*, *S. avermitilis*, *S. nigra*, *S. lincolnensis*, *S.galilaeus*, *S. vietnamensis* and *S. seoulensis*. Accession numbers are indicated. The conserved -35 and -10 promoter elements are highlighted. *S. coelicolor* sequences from 5077849 to 5081724 was used as a basis for alignment (directly upstream of *rpoB* start codon). The number of bases covered in individual alignment boxes are indicated.

rpoBp2



rpoBp3

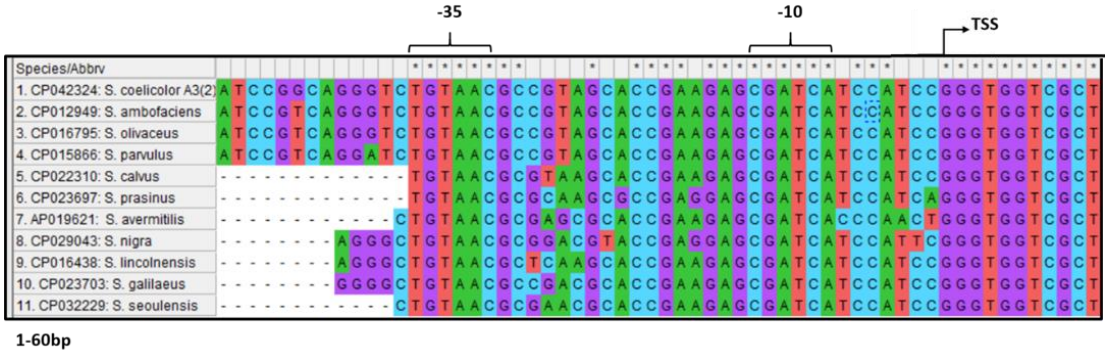


Figure 3.2.2.2: Alignment of the *rpoBp2* promoter and hypothesised *rpoBp3* promoter in *Streptomyces* genus. The *rpoB* promoter region of organisms belonging to the *Streptomyces* genus were aligned using CLUSTAL W and MEGA-X; *S. coelicolor*, *S. ambofaciens, S. olivaceus, S. parvulus, S. calvus, S. prasinus, S. avermitilis, S. nigra, S. lincolnensis, S. galilaeus* and *S. seoulensis.* Accession numbers are indicated. . The conserved -35 and -10 promoter elements are highlighted. The *S. coelicolor* sequence from 5077849 to 5081724 (directly before *rpoB* start codon) was used as a basis. The number of bases covered in individual alignment boxes are indicated. Strains used previously (Fig 3.2.2.1) that did not align were removed from figure.

rpoBp2 The *rpoBp2* promoter region is less conserved than *rpoBp1,* with several strains presenting either SNPs (compared to *S. coelicolor*) or apparently lacking the promoter elements. For those strains that contain *rpoBp2,* the -10 elements are similar to those recognised by σ^{HrdB} (Fig 3.2.2.2). The *S. griseus* strain, amongst others, did not align to this second promoter, which may explain the lone *rpoBp1* promoter identified previously (Küster, Piepersberg and Distler, 1998).

rpoBp3 In contrast to *rpoBp2*, the *rpoBp3* promoter appears to be well conserved in *Streptomyces*. However, the -10 and -35 promoter elements resemble those recognised by ECF σ factors (Otani *et al.*, 2013) rather than the principal σ (Paget & Helmann,

2003). In particular, the *rpoBp3* resemble those recognised by σ^{ShbA}. This ECF σ factor was identified as the main σ factor that is required for transcription of *hrdB* and is therefore required for normal growth (Otani *et al.*, 2013). The *hrdB* gene however is also transcribed by σ^{R}, which might explain why *shbA* mutations are not lethal (Kim *et al.*, 2012). An alignment of the *rpoBp3* consensus sequence with the σ^{ShbA}-dependent *hrdB* promoter (-35, 5' CGTAAC; and -10, 5'CGATGA) revealed 5/6 and 5/6 conservation in the elements in *S. coelicolor* as well as other *Streptomyces* organisms. It was noted that the same strains that do not contain *rpoBp2*, also do not appear to possess a *rpoBp3* promoter (Figs 3.2.2.1 and 3.2.2.2). Further work is required to investigate the role, if any, of σ^{ShbA} in *rpoBC* expression, which might reveal a more general role of this ECF σ factor in RNAP production.

It was noted that both *rpoBp2* and *rpoBp3* did not align for the *S. venezuelae* gene, where the three TSS were initially identified, which may be due to a large insertion/deletion preventing alignment to the *S. coelicolor* sequence. However, the exact *rpoBp3* promoter and TSS sequence was identified using mapped reads from this organism (Fig 3.2.1.1) suggesting the parameters for the alignment were too rigorous and that these promoters are more common amongst *Streptomyces* than initially thought.

The *rpoBp1-3* promoters in *S. coelicolor*, identified in Section 3.2.1, are displayed in Figures 3.2.2.3 and 3.2.2.4, and a full alignment is presented in the Appendix (Section 7.2). Transcription from each of the three promoters was also confirmed using the cappable TSS mapping carried out in *S. coelicolor* during exponential phase (Section 2.2.3.4) (Ettwiller *et al.*, 2016). An intext (grep) search (using standard parameters) on the galaxy platform (Afgan *et al.*, 2018) using raw data, confirmed previously hypothesised TSS for each promoter (data not shown); 20bp of DNA sequence beginning with the hypothesised TSS for *rpoBp1*, *rpoBp2* and *rpoBp3*, was used to search for sequences that were identical.

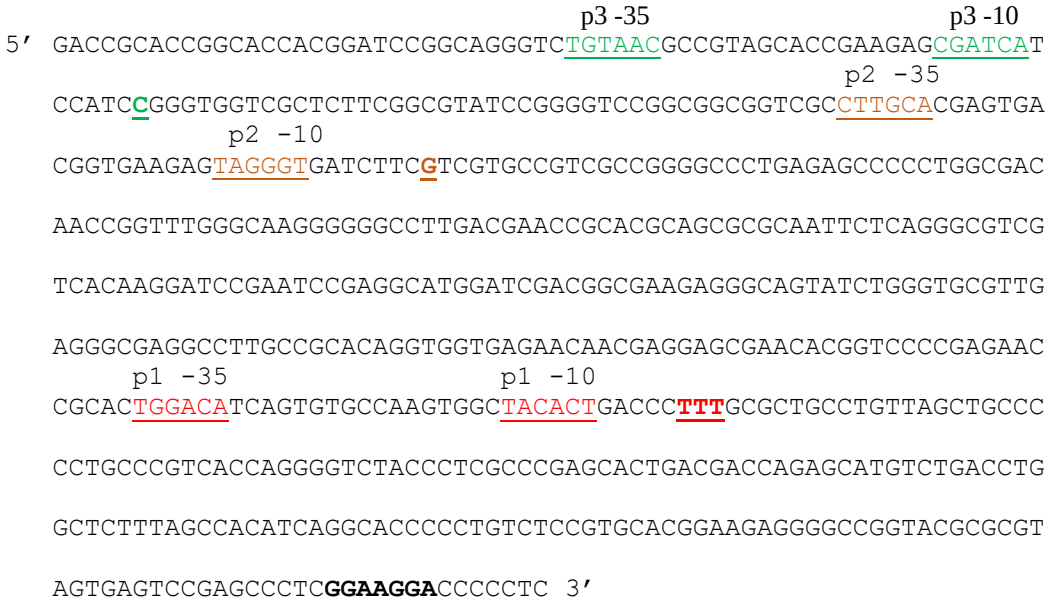


Figure 3.2.2.3: The promoter elements and DNA sequence of the *rpIL-rpoBC* intergenic region. Sequences underlined indicate promoter elements. Bases in bold and underline indicate TSS. The bold 5' GGAAGGA 3' sequence indicates the *rpoB* ribosomal binding site (RBS). p1 and p2 represent the *rpoBp1* and *rpoBp2* promoters that are activated by σ^{HrdB} (red and orange respectively). p3 represents the *rpoBp3* promoter hypothesised to be activated by the σ^{ShbA} sigma factor (green).

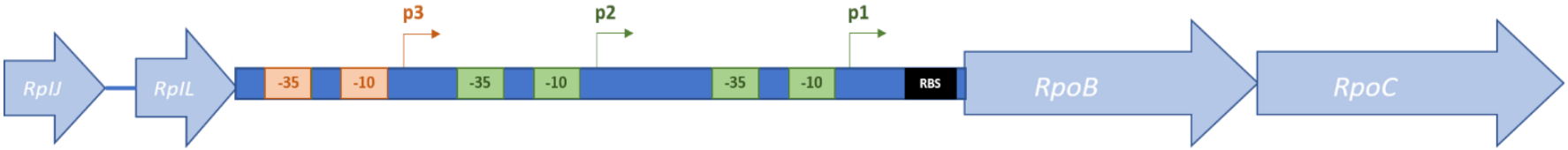


Figure 3.2.2.4: A schematic diagram of the *rpoBC* promoter region. P1, P2 and P3 represent the 3 promoters identified upstream of the *rpoBC* operon. The P1 and P2 are activated by RNAP containing the σ^{HrdB}. P3 represents the promoter region hypothesised to be activated by the σ^{ShbA} sigma factor. The -10 and -35 elements are provided as well as the Transcriptional Start Sites (TSS). Figure is not to scale.

3.2.3: Analysis of basal *rpoB* expression using *luxAB* transcriptional fusions

To investigate the contribution of the three *rpoBC* promoters and the *rpoBp1* 5'UTR in controlling *rpoBC* expression, a series of *luxAB* transcriptional fusions were constructed (Fig 3.2.3.1). The *Vibrio harveyi luxAB* system allows rapid assessment of gene expression via the quantification of light production, following supply of the substrate n-decanal (Belas *et al.*, 1982). The ϕC31-based integrative *luxAB* reporter plasmid pIJ5972 was used, where the *luxAB* genes have been engineered to be free of rare TTA codons, and also lacks a naturally occurring EcoRI site (Aigle *et al.*, 2000; M. Paget, pers comm), allowing promoters to be cloned directionally into EcoRI/BamHI- digested pIJ5972. Initially four transcriptional fusions were constructed: *rpoBp1U*, *rpoBp1-2U*, *rpoBp1-3U* and *rpoBp1*. *rpoBp1U* included *rpoBp1* along with the 5'UTR up to but not including the *rpoB* RBS, and so relied on the *luxAB* RBS for translation initiation. The *rpoBp1-2U* and *rpoBp1-3U* fragments extended to different positions upstream to include *rpoBp2* and *rpoBp3*. The *rpoBp1* promoter fragment included the *rpoBp1* promoter as far as +10, but lacked the remaining 5'UTR (Fig 3.2.3.1). Promoter regions were amplified using primer sets 1 and 3, 1 and 2, 2 and 5 and 2 and 6, respectively (Table 2.1.5.1) and, following initial cloning into pBluescript II SK+, were subcloned as EcoRI/BglII fragments into a EcoRI/BamHI-digested pIJ5972. Final plasmid constructs were confirmed using DNA sequencing, then introduced into *S. coelicolor* M145 by conjugation (Sections 2.2.4.3 and 2.2.4.5).

Strains were grown in at least biological triplicate to exponential phase (OD450 0.8-1.0) before luciferase assays were performed, and normalised to optical density at 450 nm (Fig 3.2.3.2). Consistent with 5'end mapping data (Fig 3.2.2.1) most transcription appeared to depend on *rpoBp1* with a 30 % increase in relative luminescence seen when the p2 and p3 promoters were present (comparing *rpoBp1U*, *rpoBp1-2U* and *rpoBp1-3U*). Interestingly, the activity of *rpoBp1* (lacking the 5' UTR) was ~1.8-fold higher than *rpoBp1U*, indicating that the 5'UTR has a negative effect on transcription.

Overall, these experiments indicate that at least three promoters express the *rpoBC* operon, and that *rpoBp1* is the major promoter, consistent with earlier work (Barry,

Squires and Squires, 1979; Barry, G, 1985; Küster, Piepersberg and Distler, 1998), which also suggested that the *rpoB* 5'UTR plays a negative role in *rpoBC* expression.

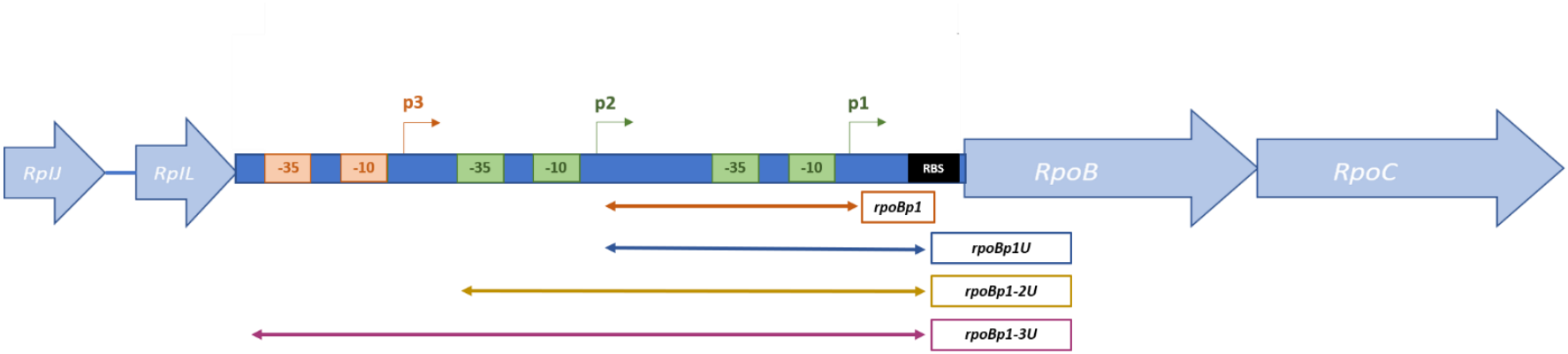


Figure 3.2.3.1: Truncations of the *rpoB* promoter region used for pIJ5972 fusions. The figure illustrates the promoter fragments p1, p1U, p1-2U, and p1-3U that were used in transcriptional fusions.

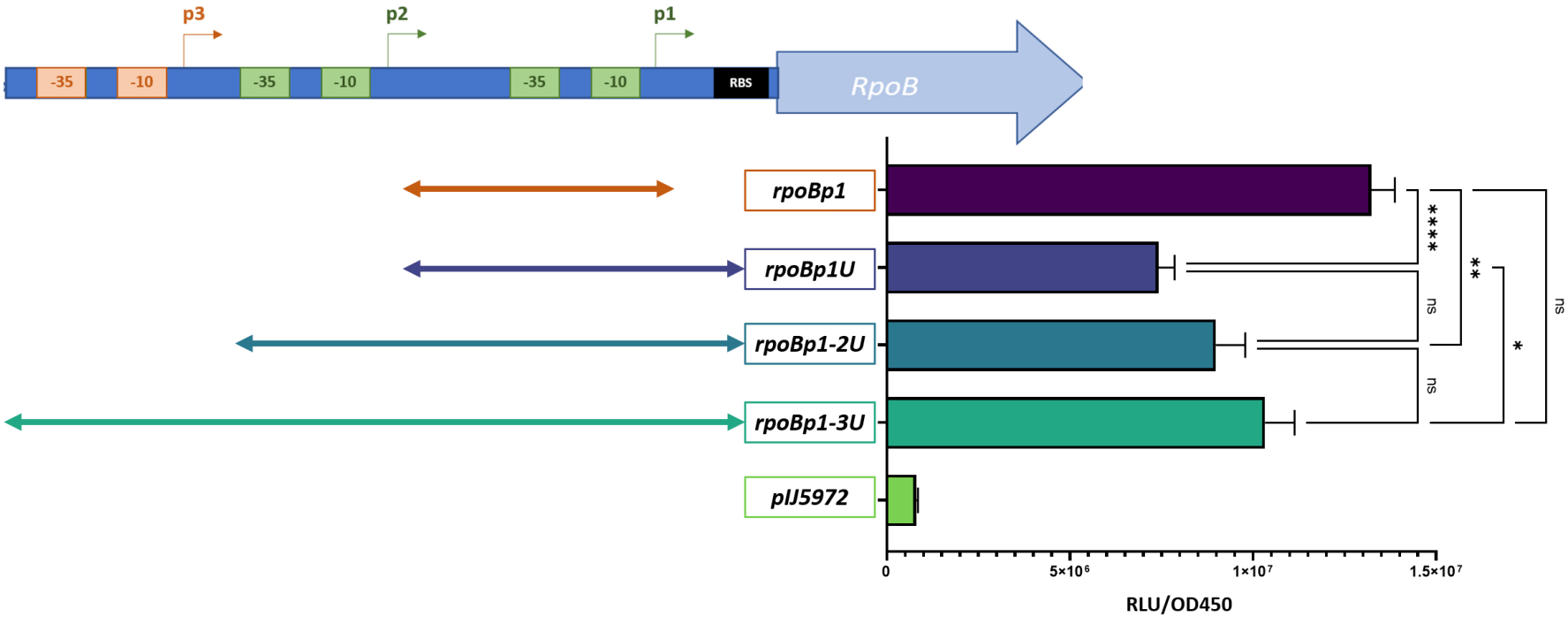


Figure 3.2.3.2: Analysis of *rpoBp*-luxAB transcriptional fusions during exponential growth. Strains were grown to OD450 0.8-1.0 in NMMP medium and luciferase activities normalised to optical density. pIJ5972 represents the luciferase activity of the vector-only control. Data are from at least triplicate biological replicates and significance was determined using a Brown-Forsythe and Welch ANOVA test. Error bars represent the S.E.M. (significance values: P < 0.05 = *, P < 0.01 = **, P < 0.0001 = ****).

3.2.4 Analysis of basal *rpoB* expression using pSNypet transcriptional fusions

The luciferase based-reporter system was a useful tool for determining basal promoter activity. However, luciferase was previously reported to be proteolytically stable in *E. coli* (Allen *et al.*, 2007), and initial experiments suggested that this was also the case in *S. coelicolor* (Appendix 7.3), limiting its use in monitoring rapid changes in gene expression. Therefore, an RNA-based dynamic reporter plasmid pSNypet was used; pSNypet is derived from the integrative plasmid pRT802 and includes a copy of the *ypet* fluorescence reporter that lacks translation initial signals (ribosome binding site and start codon, designated *ypet**) and is flanked by upstream tt_{sbiB}, a strong terminator in *Streptomyces* (created by Sven Reisloehner and Dr Sophie Nicod) (based on Horbal *et al.*, 2018; Huff *et al.*, 2010). The RNA products are unstable due to the absence of translation and this allows the quantification of promoter expression under dynamic conditions, including down-regulation.

S. coelicolor genomic DNA and primer pairs 7 and 8, 7 and 10, 10 and 11, and 11 and 12 were used for the amplification of *rpoBp1, rpoBp1U, rpoBp1-2U* and *rpoBp1-3U* (Table 2.1.5.1). Engineered upstream (XhoI) and downstream (HindIII) restriction sites allowed cloning into pSNypet via initial blunt-end cloning into pBluescript II SK (Sections 2.2.2.3 and 2.2.2.1). Plasmids were then conjugated into M145 using the ETR strain. Strains were grown in triplicate to OD_{450} 0.8-1.0 in NMMP medium before the RNA was extracted (Section 2.2.3.3) and then taken forward for DNase treatment and cDNA synthesis prior to qPCR analysis (Section 2.2.5.1).

Basal expression data (Fig 3.2.4.1) paralleled those obtained using luciferase (Fig 3.4.4.1), with a gradual increase in expression as the length of the promoter region increased; 3.1-, 7.2- and 8.3-fold increases were observed for the *rpoBp1U, rpoBp1-2U* and *rpoBp1-3U* compared to the empty vector control, as expected by the presence of additional promoters. In the absence of the 5'UTR, *rpoBp1* activity was increased 17-fold confirming that this region is negatively acting; however the difference was far greater than that observed using luciferase, where only a 1.8-fold increase was seen (Fig 3.4.4.1).

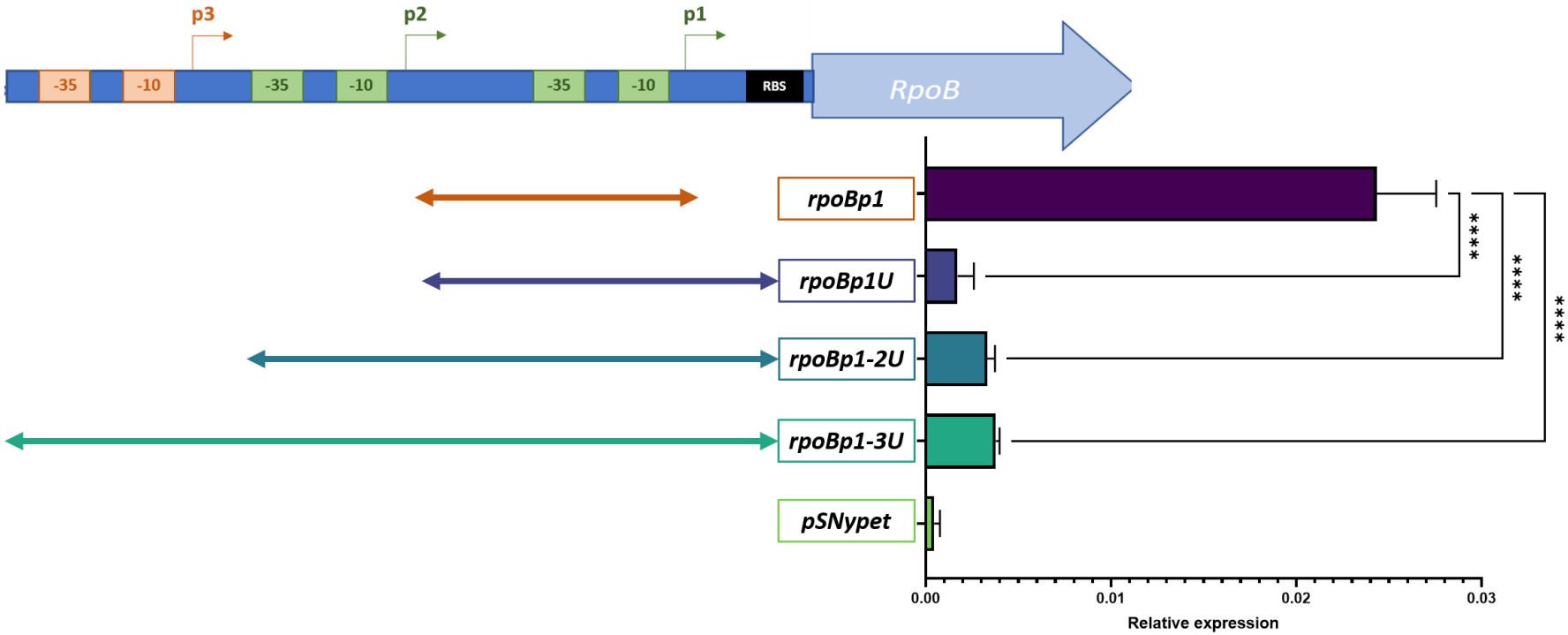


Figure 3.2.4.1: The effect of differing 5' and 3' *rpoBp* truncations on *ypet* reporter expression. Strains were grown to an OD450 of ~0.8-1.0 with *ypet** expression normalised using 16s stable rRNA expression and detected using primer pairs 65 and 66, and 67 and 68, respectively (Table 2.1.5.3). pSNypet represents the expression of the empty vector control. Data is from at least biological triplicate replicates and error bars represent the standard deviation. Significance values were calculated using a one-way ANOVA (P < 0.0001 = ****).

3.3 The effect of stress on *rpoBC* expression

3.3.1 The *rpoBC* operon is downregulated during oxidative stress

Previous microarray experiments by Kallifidas and colleagues (2010) detailed that several genes involved in RNA synthesis, including *rpoA*, *rpoB* and *rpoC*, were largely downregulated in response to diamide-induced oxidative stress. However, the mechanism that underlies this is not understood. To confirm the *rpoBC* operon is downregulated, *rpoB* expression was determined before and after a diamide-induced stress using qPCR.

M145 was pre-germinated and grown in biological duplicate in NMMP supplemented with amino acids to an OD450 of ~0.8-1.0, before oxidative stress was induced using diamide and samples collected (Sections 2.1.7.2 and 2.2.1.3). RNA extraction, DNase I treatment and cDNA synthesis were carried out as described in section 2.2.5.1 and *rpoB* transcript levels determined using qPCR primers 71 and 72 (Table 2.1.5.3). Expression was normalised to stable 16s rRNA, which was determined using primers 67 and 68. Actual transcript levels were quantified using a standard curve of known copies of *S. coelicolor* genomic DNA.

A significant ~2.5 fold decrease in *rpoB* expression was observed after 5 min diamide addition, which further decreased to ~3.3 fold less than basal expression (time 0) after 10 min. Expression remained low after 30 min, but was restored after 60 min. This data is in agreement with the previous microarray work, where expression of *rpoBC* was downregulated at least greater than 2-fold but returned to pre-treated levels ~40-60 min after diamide induced oxidative stress (Kallifidas *et al.*, 2010).

However, the mechanism that underlies this rapid downregulation is not understood, and was previously shown due to be independent of the σ^R-regulon. It is also thought the down-regulation of *rpoBC* and other growth-related genes is not facilitated by (p)ppGpp, as the levels of these alarmones does not significantly increase upon the induction of oxidative stress using diamide (Paget and Hesketh, unpublished results; Kallifidas *et al.*, 2010).

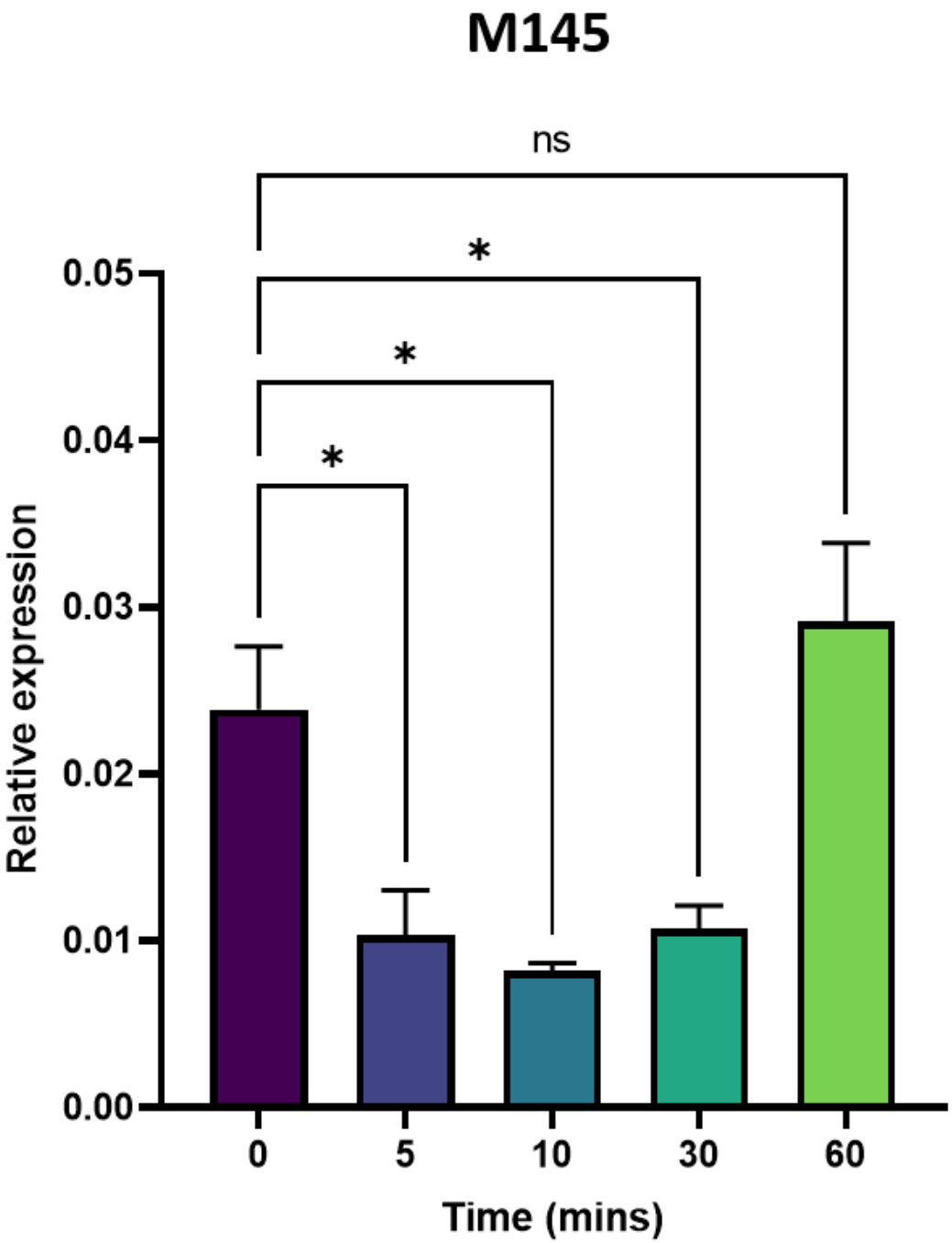


Figure 3.3.1.1: The expression of *rpoB* following diamide-induced oxidative stress. M145 was grown in biological duplicate to OD_{450} ~0.8-1.0 then treated with diamide (0.5 mM) Transcript levels were determined using qRT-PCR, using a standard curve of known quantities of *S. coelicolor* genomic DNA and normalisation to stable 16s rRNA. Error bars present the S.E.M for each data set. Significance values: P < 0.05 = *, P < 0.0001 = ****.

3.3.2 Construction of a stably integrated pSX400:*rpoBp1-rpoBC* fusion

In order to investigate regulation of the major *rpoBp1* promoter in isolation from the other promoters, and in its native chromosomal context, the non-replicating pSX400 plasmid, containing the lone *rpoBp1* promoter, was integrated into the genome of *S. coelicolor* by single-crossover homologous recombination, placing the full *rpoBC* operon under the control of the p1 promoter alone. pSX400 includes a degradation tag based on the *S. coelicolor* transfer-messenger RNA sequence (tmRNA) that, once translated, signals for the degradation of any linked truncated protein by C-terminal specific proteases (Keiler, Waller and Sauer, 1996). The amino acid sequence used for targeted degradation in pSX400 is RDSSQQAFALAA (L. Humphrey and M. Paget, pers. comm).

Initially PCR using *S. coelicolor* genomic DNA and primer sets 15 and 16 (Table 2.1.5.1) was used to amplify the ~1.2 kb N-terminal region of *rpoB* including ~160bp of upstream promoter DNA (including the *rpoBp1* promoter only; see Section 2.2.2.3). This was cloned into EcoRV-digested pBluescript II SK+, sequenced and further subcloned as a BamHI/DraIII digested fragment into pSX400 vector. The resulting plasmid, cloning produced the pSX400*::rpoBp1* plasmid, which includes 1 kb C-terminally truncated *rpoB* fused directly to the degradation tag, which should ensure that any truncated β protein produced is degraded, and therefore unable to interfere with expression (see Fig 3.3.2.1). pSX400*::rpoBp1* was used to transform *E. coli* ETZ before conjugation into *S. coelicolor* M145 (Sections 2.2.4.3 and 2.2.4.5).

Single-crossover recombinants were confirmed using colony PCR (see Section 2.2.2.5) A product size of ~1.3kb indicated that the plasmid had stably integrated into the genome of *S. coelicolor*, with the forward primer amplifying a sequence within pSX400 and the reverse primer annealing to part of *rpoB* (see Figs 3.3.2.1, 3.3.2.2) gDNA and empty plasmid were included as negative controls and were confirmed to not be amplified (data not shown). PCR products were purified (Section 2.2.2.3) and sequenced to confirm the expected structure, with one successful exconjugant taken forward and designated S301. Note that growth of S301 and the M145 parent strain were indistinguishable on standard lab media (see Section 4.4.1.2).

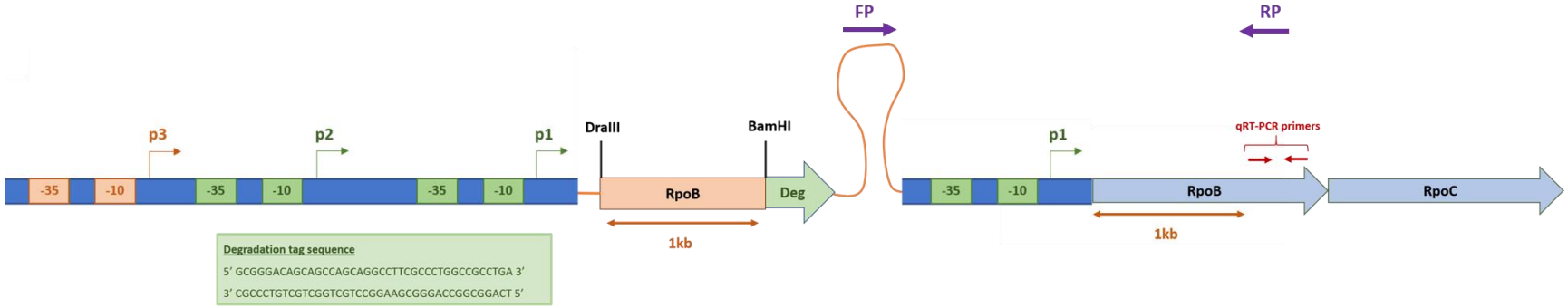


Figure 3.3.2.1: The expected *S. coelicolor* chromosome structure upon integration of pSX400::*rpoBp1*. The truncated *rpoB* (orange) is flanked by the DraIII and BamHI sites used for cloning, with the rest of the plasmid, integrated in the chromosome (represented by an orange loop). The 1kb fragment represents the sequence used for homologous recombination. The primers used for qRT-PCR are indicated in red. The primers used to confirm the integration of the plasmid into the *S. coelicolor* genome are shown in purple; FP = forward primer binding to pSX400, RP = reverse primer binding within the *rpoB* gene, outside of the 1kb of coding sequence used for recombination. Not to scale.

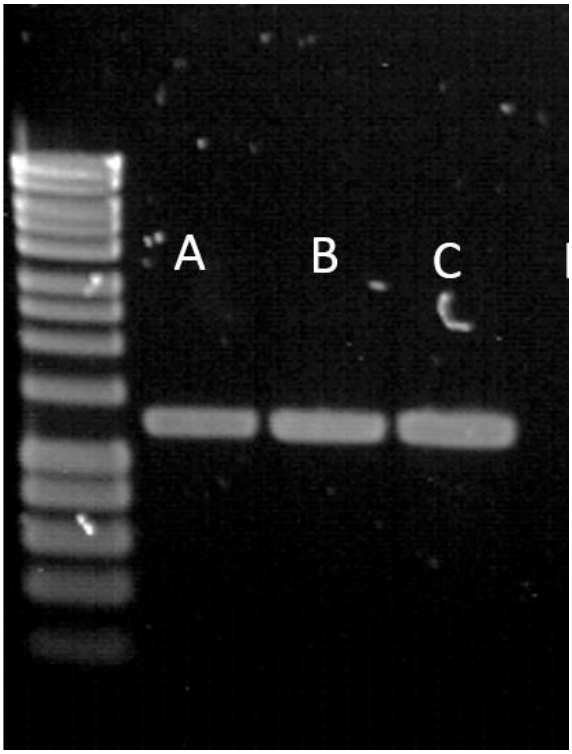


Figure 3.3.2.2: PCR products confirming integration of pSX400*::rpoBp1* into the *S. coelicolor* genome. PCR products were separated on a standard 0.8% agarose/ TAE gel for 50 min at 150V. A, B and C represent differing exconjugants restreaked from initial conjugation of M145 with the pSX400*:rpoBp1* plasmid. Forward and reverse primers (17 and 18 respectively, Table 2.2.5.1) bind to the pSX400 plasmid and the *rpoB* gene respectively, with a ~1.3kb band only observed upon integration of the plasmid into the *S. coelicolor* genome.

3.3.3 The *rpoBp1* promoter is downregulated in response to oxidative stress

To investigate the response of *rpoBp1* to oxidative stress in its native context, S301 was grown in biological triplicate in NMMP medium supplemented with amino acids (Section 2.1.7.2), and diamide-induced oxidative stress applied as described in Section 2.2.1.3. RNA was extracted from samples taken before and 5, 10, 20, 30, 45 and 60 min after the addition of diamide, before conversion to cDNA (Section 2.2.5.1). *RpoB* expression was determined using qRT-PCR with transcript levels normalised to stable 16s rRNA (see Section 3.3.1); note the *rpoB* primers used for qRT-PCR amplify a region downstream from the truncated *rpoB* sequence used for recombination (>1kb after start codon), ensuring that only the intact *rpoB* ORF was analysed (Fig 3.3.2.1).

A rapid and significant ~3.8 fold decrease in expression was confirmed within 5 min of diamide treatment (Fig 3.3.3.1), with levels of expression remaining low for ~30 min, then recovering after 1 h. These data are similar to those obtained for M145 (Fig 3.3.1.1) and to previously reported microarray data (Kallifidas *et al*, 2010) suggesting that the activity of *rpoBp1* determines the observed downregulation of the *rpoBC* operon in response to oxidative stress. However, the roles of *rpoBp2* and *rpoBp3* during oxidative stress are yet to be determined.

Interestingly, quantitative analysis revealed that the basal level (normalised to stable 16s rRNA) of the *rpoB* transcript in S301 (*rpoBp1* promoter alone) is around 2.7 fold lower than M145 (*rpoBp1-3*) consistent with the presence of the additional promoters (Figs 3.3.1.1 and 3.3.3.1), although larger than expected since *rpoBp1* is the main promoter for *rpoB*.

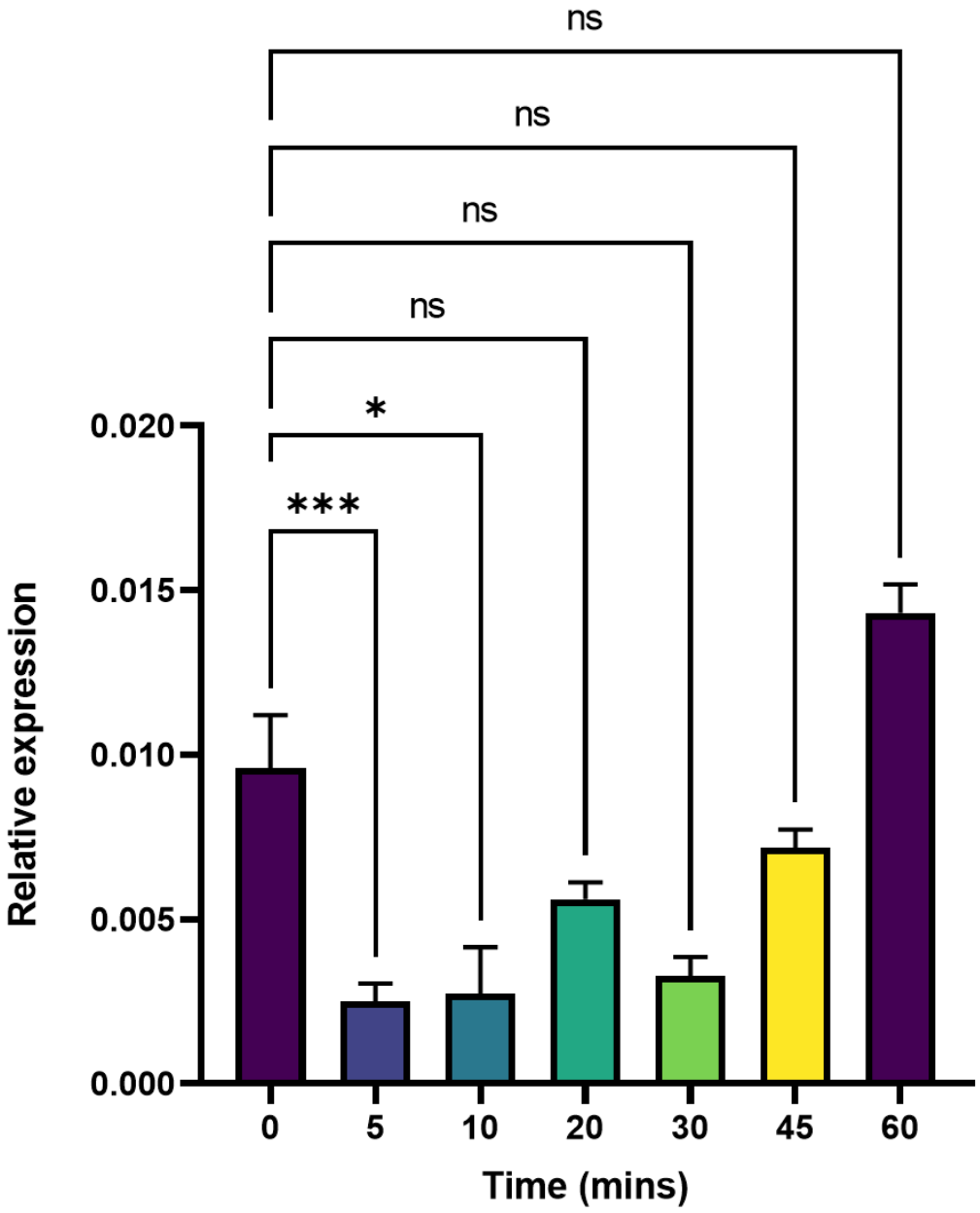


Figure 3.3.3.1: The expression of *rpoBp1* following diamide treatment in S301. S301 was grown in biological triplicate before treatment of cultures with 0.5 mM diamide, and *rpoB* expression was determined using qRT-PCR. Reads were quantified using a standard curve of known copies and expression relative to stable 16s rRNA. Error bars represent the S.E.M for each data set; For determination of significance a one-way ANOVA was used; $P < 0.05$ = *, $P < 0.0005$ = ***.

3.3.4 The *rpoBC* operon is downregulated during nutrient limitation in a partially ppGpp-dependent manner.

The stringent response in *S. coelicolor* is dependent on RelA-catalysed (p)ppGpp production, and involves global changes in transcription, including the rapid downregulation of growth- related genes, including stable rRNA (Hesketh, Sun and Bibb, 2001; Hesketh *et al.*, 2007a). However, perhaps surprisingly, the *rpoBC* operon was not found to be significantly down-regulated (Hesketh *et al.*, 2007a). One possible explanation is that the artificial experimental approach involved the induction of truncated *relA* using a thiostrepton inducible promoter, which might delay the response. In addition, the multiple *rpoBC* promoters might mask promoter-specific responses. Therefore, the control of *rpoBC* during stringent conditions was revisited here.

Early work by Strauch and colleagues (1991), using an S1 nuclease mapping approach, determined that all four *rrnD* promoters were down-regulated of within 20 min of nutrient downshift in *S. coelicolor.* Initially, experiments were performed to determine the stringent response in M145, investigated using a qRT-PCR approach. Primers were designed to amplify a product located within the *rrnD* 5' unstable region upstream of 16S rRNA, which should be rapidly degraded during rRNA maturation and therefore reflect transcriptional changes (Li, Pandit and Deutscher, 1999; Romero *et al.*, 2014). To confirm that any response observed was (p)ppGpp-dependent, experiments were performed using both M145 and M571 (Δ*relA*); M571 does not produce detectable ppGpp (Fernández-Martínez, Gomez-Escribano and Bibb, 2015).

These strains were grown to mid-log phase in NMMP liquid medium, then subjected to amino acid depletion, with RNA extracted before (time 0), and 5, 10, 15, 20, 30 and 60 min proceeding the stress, before being processed for qRT-PCR analysis (see Sections 2.2.1.4 and 2.2.5.1) .

In M145 (RelA+) expression of *rrnD* steadily decreased within 5 min of amino acid depletion, while expression remained relatively stable in M571 (RelA-) (Fig 3.3.4.1), confirming a RelA-dependent stringent response.

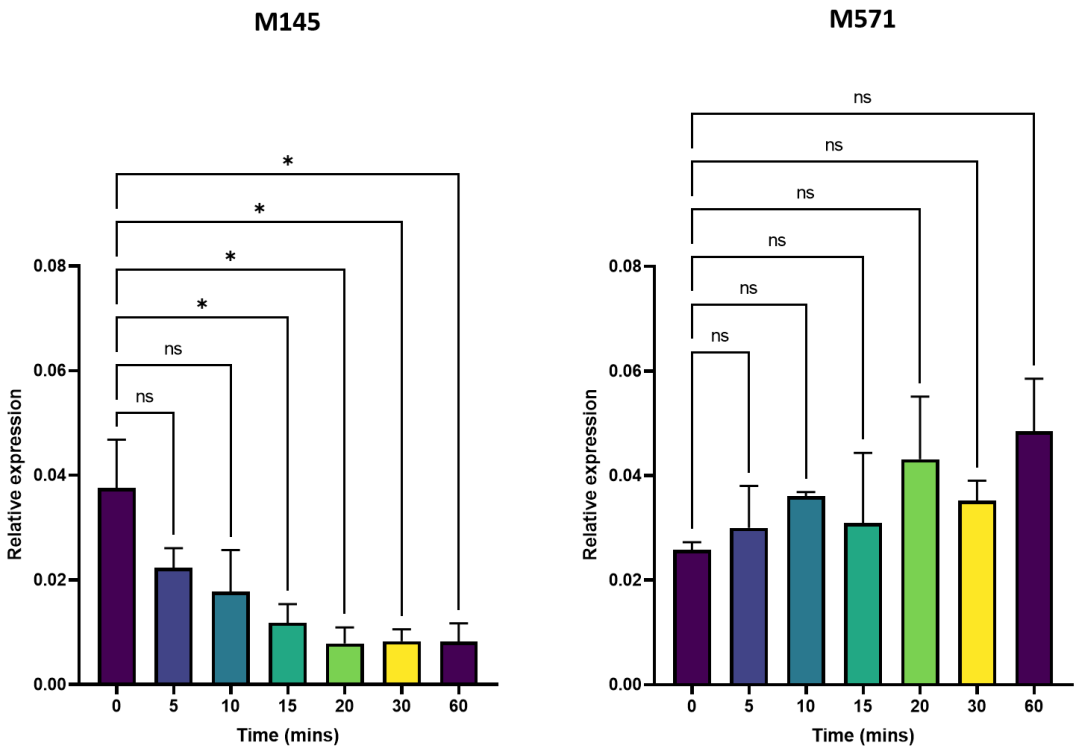


Figure 3.3.4.1: Expression of *rrnD* is stringently controlled in M145 (RelA+) but not in M571 (RelA-). The expression of *rrnD* was determined using qRT-PCR and normalised to stable 16S rRNA. cDNA was diluted 1/1000 fold for qPCR testing, using primer sets 69 and 70, and 67 and 68 for the amplification of rrnD unstable and 16S stable sRNA regions respectively, with level of cDNA determined using a standard curve of known gDNA copies (Table 2.1.5.3). The error bars shown represent the S.E.M of the results in biological triplicate, and changes in expression were determined relative to stable 16S rRNA at time zero. Statistical significance was calculated using an ordinary one-way ANOVA, where equal distribution between data and standard deviations is assumed. (P< 0.05 = *)

After confirming the stringent response, the same cDNA was taken forward for quantification of *rpoB* expression in M145 and M571 using qRT-PCR with expression was normalised again to stable 16S rRNA. Similar to *rrnD*, amino acid deprivation led to a steady decrease in *rpoB* expression in M145, with the lowest level reached after 30 min, where a two-fold drop in relative expression is seen (Fig 3.3.4.2). However, this decrease was not observed in M571 suggesting that this response is ppGpp-dependent. Surprisingly, unlike *rrnD*, *rpoB* expression appeared to be restored to near basal (time 0) expression levels after 60 min. This may suggest increased sensitivity of *rpoBp* to intracellular conditions, allowing restoration of transcriptional capacity as the strain recovers from the stress, for example allowing for the synthesis of its own amino acids.

Interestingly, although *rpoB* expression was not downregulated in M571, it increased ~3 fold compared to time zero 60 min following nutrient downshift, however the reason for this remains unclear (Fig 3.3.4.2).

Overall, these experiments confirm the induction of a stringent response in M145 and provide evidence that the downregulation of the *rpoBC* operon during this stress is carried out in a ppGpp-dependent manner.

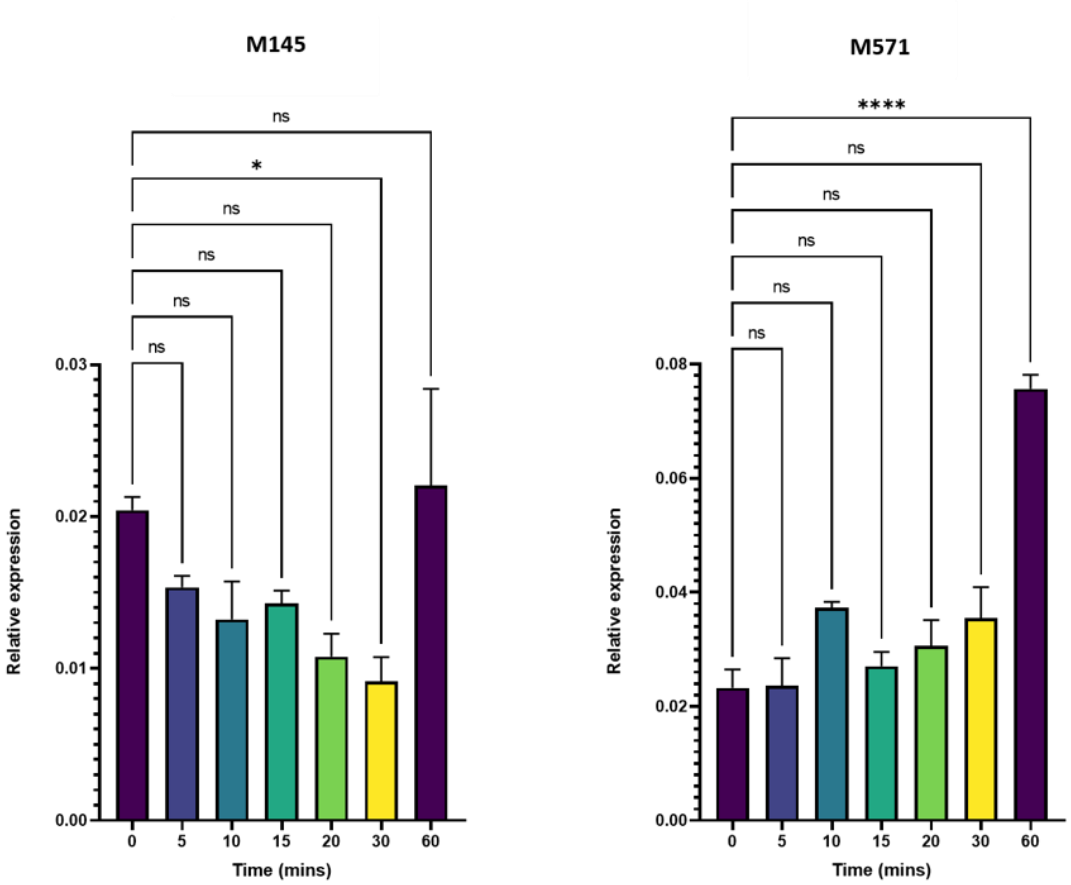


Figure 3.3.4.2: Expression of *rpoB* is stringently controlled in M145 (RelA+) but not in M571 (RelA-). The expression of *rpoB* was determined using qRT-PCR and normalised using the stable 16s rRNA, using primers 71 and 72, and 67 and 68, respectively (Table 2.1.5.3). The error bars shown represent the S.E.M and significant changes in expression were determined relative to expression of *rpoB* at time zero. Statistical significance was calculated using an ordinary one-way ANOVA, where equal distribution between data and standard deviations is assumed. ($P < 0.05$ = *, $P < 0.0001$ = ****).

3.3.5 The *rpoBp1* promoter is stringently controlled

As discussed above, the main *rpoBp1* promoter is downregulated during oxidative stress. To investigate whether this promoter was also stringently controlled, the M145 strain containing the previously constructed pIJ5972*::rpoBp1U* plasmid was taken forward for nutrient downshift, RNA extraction and qRT-PCR (Section 3.3.4); *luxA* transcript levels were quantified 0, 5, 10, 15, 20, 30 and 60 min after nutrient downshift. The qRT-PCR approach, which was used due to the stability of bacterial luciferase (see Appendix section 7.3), revealed a rapid decrease in *luxA* expression, reaching its lowest point at 30 min where transcript levels were ~11 fold less than those observed before stress (Fig 3.3.5.1). While the trend presented by the *rpoBp1* region alone is similar to that seen for the M145 strain containing the whole promoter region, the drop in transcript of *rpoBp1U* is both larger and more significant (Figs 3.3.4.2 and 3.3.5.1). This suggests that *rpoBp2* and *rpoBp3* contribute to continued transcription during nutrient downshift, ensuring sustained low level expression of the essential RNAP subunits.

Unlike the native *rpoB* transcripts (controlled by *rpoBp1-3*) *luxA* transcript levels were not fully restored 60 min after the induction of stress, which further suggests that this increase might be due to increased activity of *rpoBp2* and/or *rpoBp3*.

To investigate whether the stringent control of *rpoBp1* depended on *relA*, the pIJ5972:: *rpoBp1U* fusion was studied in M571, although in this case the time course was only performed for 30 min and with two biological replicates. No significant changes in *luxA* expression was observed (Fig 3.3.5.2) indicating that the stringent control of *rpoB,* and *rpoBp1* specifically, is ppGpp-dependent.

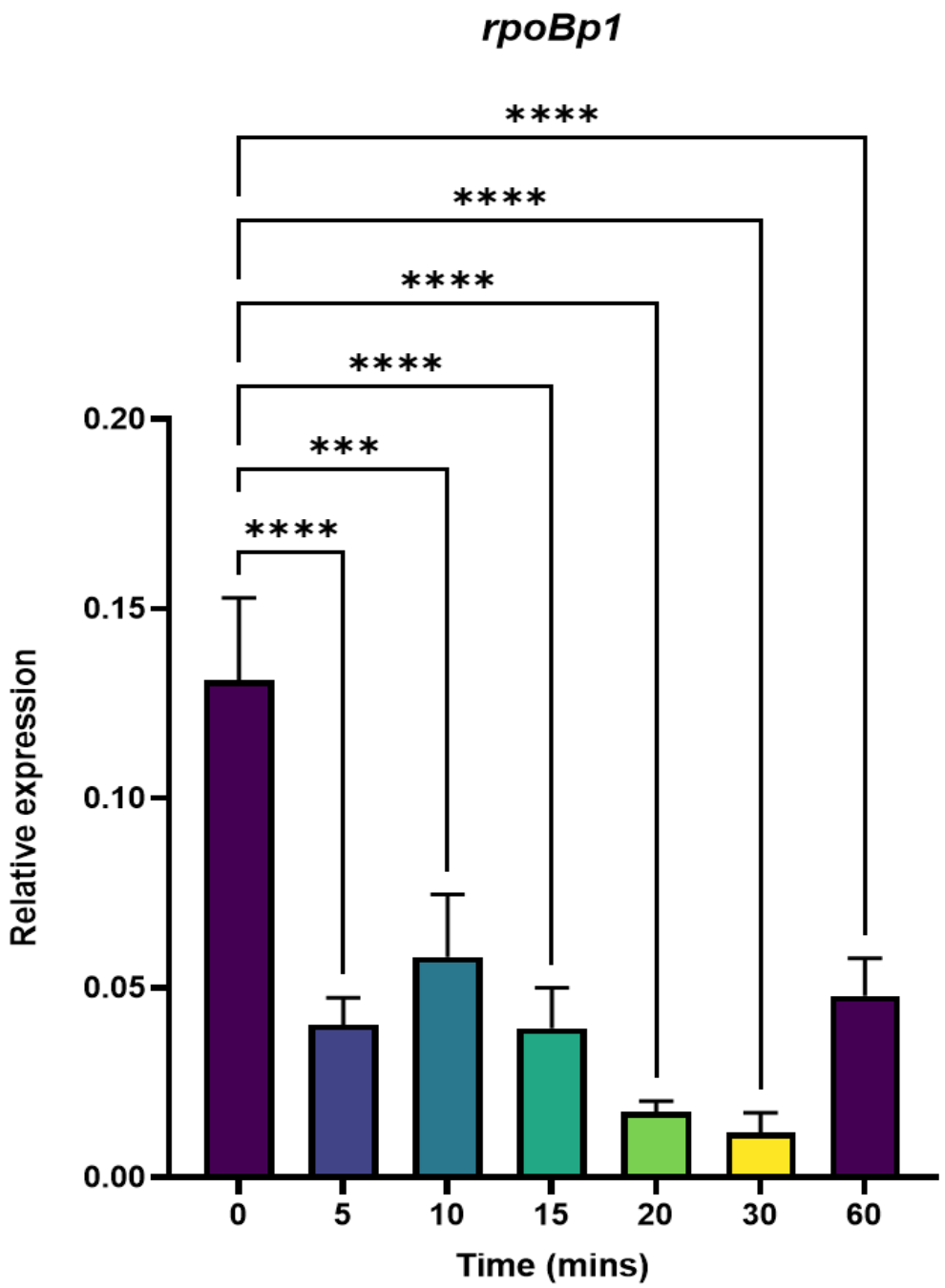


Figure 3.3.5.1: The *rpoBp1* promoter is stringently controlled. Relative expression of the *luxA* transcripts of the pIJ5972::*rpoBp1U* plasmid during stringent conditions in an M145 background. *luxA* expression was determined by normalisation to 16s stable transcript level using primer pairs 73 and 74, and 67 and 68, respectively (Table 2.1.5.3). Copy number was determined using a standard curve prepared using pIJ5972 vector-only plasmid and *S. coelicolor* genomic DNA. Error bars represent the S.E.M for each time point, grown and tested in both biological and technical triplicate. Data was confirmed to be normally distributed using a D'Agostino and Pearson test and Shapiro-Wilk test. An ordinary one-way ANOVA, was used to determine significance. ($P < 0.0005$ = ***., $P < 0.0001$ = ****).

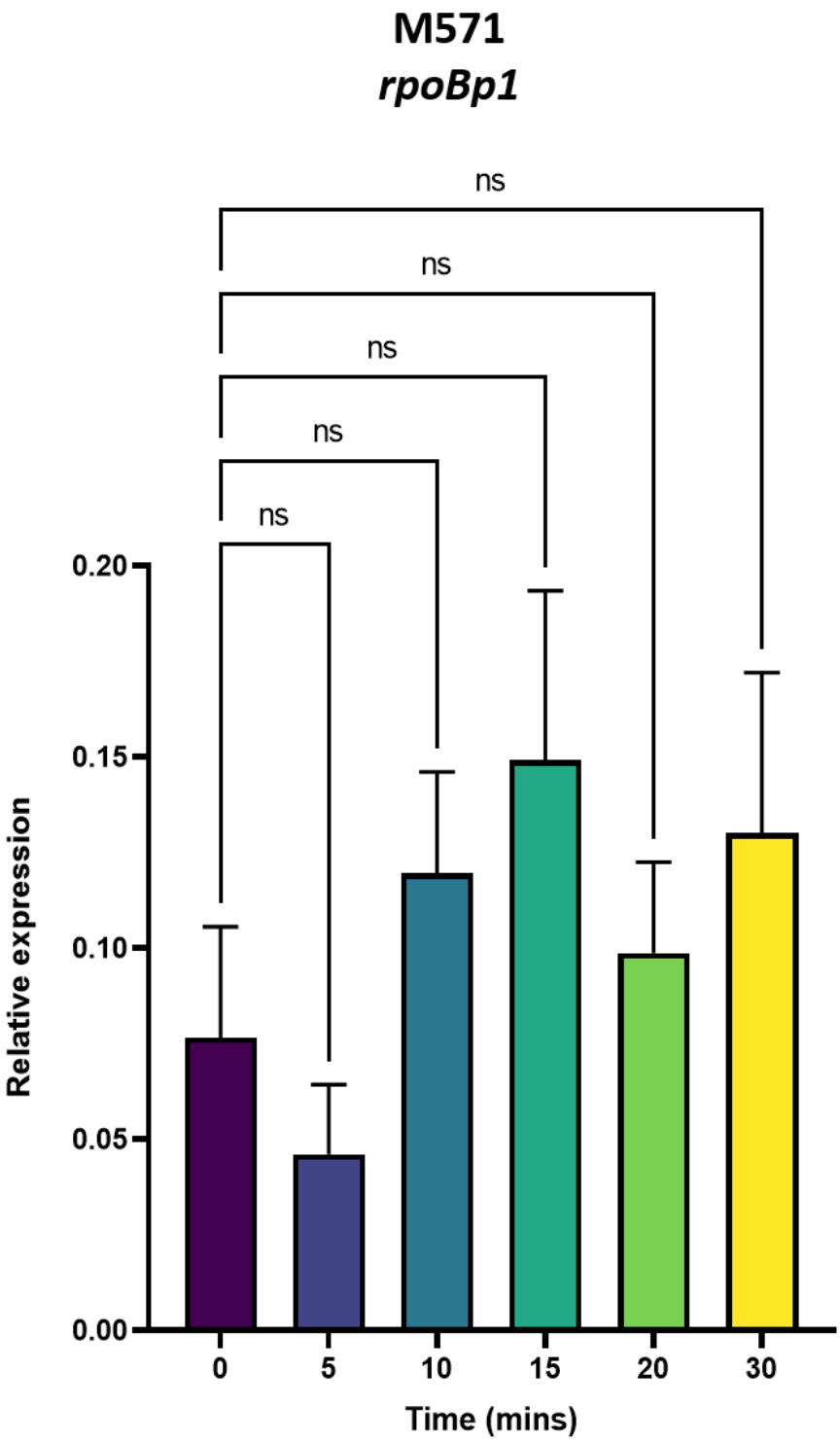


Figure 3.3.5.2: The *rpoBp1* promoter is not downregulated during stringent conditions in M571 (RelA-). Relative expression of the *luxA* transcripts of the pIJ5972::*rpoBp1U* plasmid during stringent conditions in an M145 background. *luxA* expression was determined by normalisation to 16s stable transcript level using primer pairs 73 and 74, and 67 and 68, respectively (Table 2.1.5.3). Transcript number was determined using a standard curve prepared using pIJ5972 vector-only plasmid and *S. coelicolor* genomic DNA. Error bars represent the S.E.M for each time point, grown in biological duplicate and cDNA samples tested in technical triplicate. Significance was determined using a one-way ANOVA as data and SD were both normally distributed.

3.4 Linker-scanning mutagenesis of the *rpoBp1* 5'UTR reveals several potential cis-acting regulatory elements.

3.4.1 Identification of potential regulatory elements in the 5'UTR

Initial experiments characterising the basal expression of differing *rpoBp* constructs provided evidence for a negatively acting regulatory element or elements present in the 5'UTR (Sections 3.2.3 and 3.2.4). This was in agreement with a previous study that identified an inverted repeat in this 5'UTR, hypothesised to be acting as an attenuator of *rpoBC* expression in *S. griseus* (*RpoBa*) (Küster, Piepersberg and Distler, 1998). Sequence alignments revealed an inverted repeat sequence in the 5'UTR of the *S. coelicolor* promoter region, located in a similar position, but with only 57% identity to *rpoBa* in *S. griseus* (Fig 3.4.4.1)*;* the *S. coelicolor* inverted repeat sequence was positioned from +71 to +103 bp relative to the *rpoBp1* +1 transcription initiation point. To investigate this inverted repeat (also designated *rpoBa*) and identify any additional potential secondary structures, the entire 5'UTR of *S. coelicolor rpoB* (+1 to +161 nt) was analysed using an RNA folding algorithm (Mathews, 2014). The most probable secondary structure (Fig 3.4.1.2) was based on a transcript that included an additional 5' U nucleotide (5'UUUU), since most transcripts have this addition (see Section 4.2), although note that the RBS (spanning residues +162 to +168) was not included. A large stem loop, extending beyond but centred around *rpoBa,* was identified. A second structure was also seen to span the last 30 nt of the 5'UTR, (before the RBS), however it's probability of structure is much lower than observed for *rpoBa*, and hence is not largely discussed in the present study.

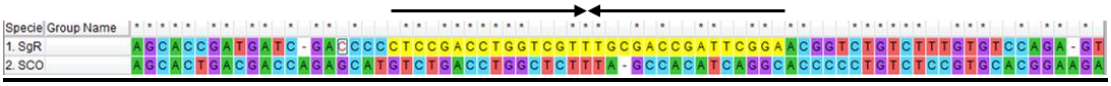


Figure 3.4.1.1: Alignment of the *rpoBa* inverted repeat sequence in *S. griseus* (SgR) and *S. coelicolor* (SCO). The Clustal W2 sequence alignment tool was used with the MEGA-X programme. There is 57% identity between the two inverted repeat sequences. The inverted repeat sequence is highlighted in yellow for SgR, and displayed by black arrows.

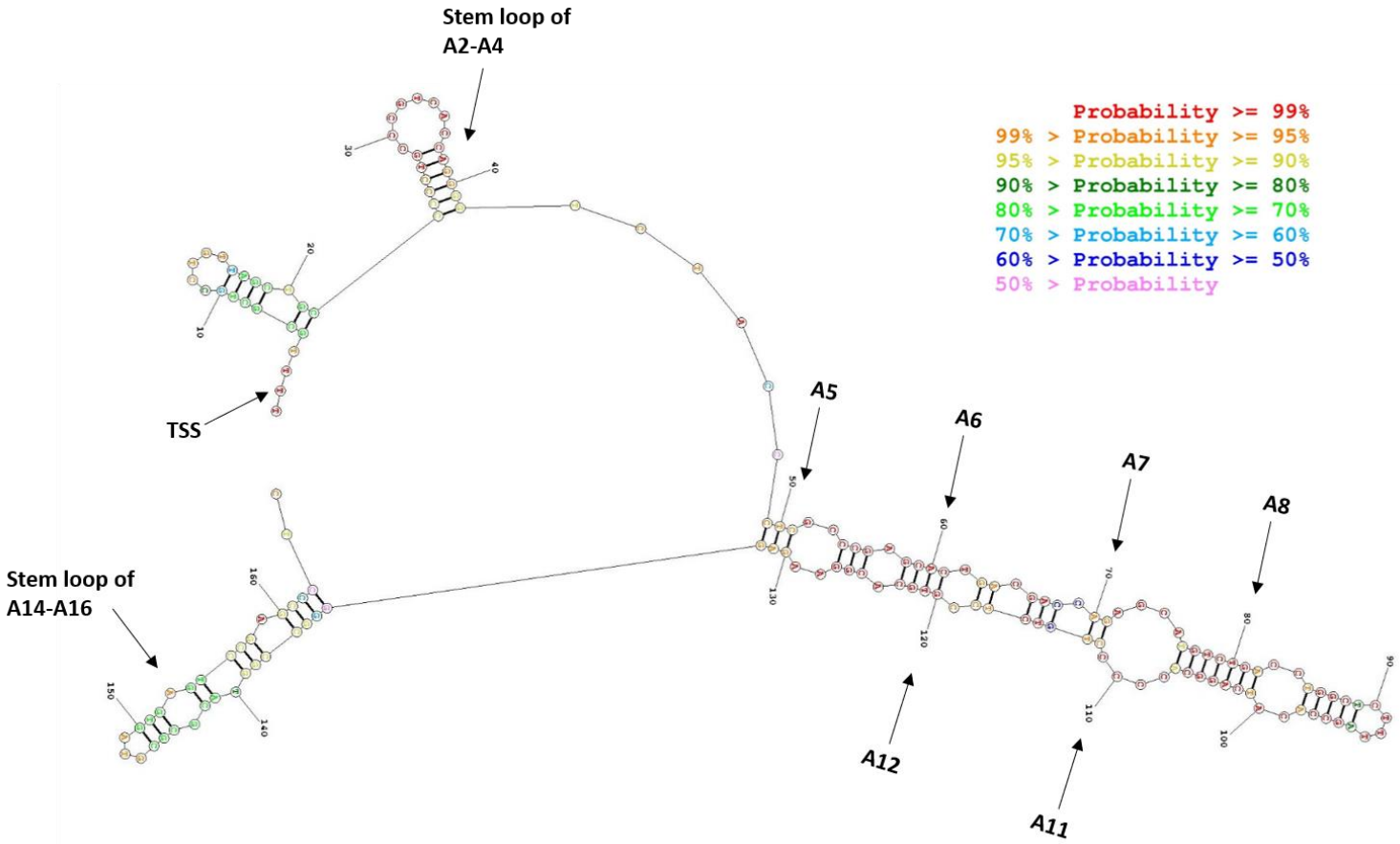


Figure 3.4.1.2: Hypothesised RNA secondary structure for the 5'UTR of *rpoB*. The RNA sequence spanning from the TSS down to but not including the RBS for the *rpoBC* operon (+161) was analysed using secondary structure prediction (Mathews, 2014); 4 U nucleotides were included at the 5' end of the transcript representing the sequence of the majority of the *rpoB* transcripts (Fig 4.2.1.3). Default fold options were used; maximum of 10% energy difference and temperature of 310.15K. Every 10bp from the +1 nucleotide is marked, and several of the attenuator mutations were included for reference (A2-A4, A5-A8, A11-A12 and A14-A16). Probabilities for pairing between bases are provided as colours; red = >99%, orange = >95%, yellow = >90%, dark green = >80%, light green = >70%, light blue = >60%, dark blue = >50%,, pink = <50%.

3.4.2 Construction of linker mutations in *rpoB* 5'UTR

To identify transcriptional regulatory elements, 10 bp scanning mutagenesis was performed on the 161 bp 5'UTR region, with expression quantified using transcriptional fusions to *luxAB* in pIJ5972. Sixteen consecutive mutations (A1-A16) were constructed by replacing native sequence with 10 bp linker DNA (5'CAGCTAGCCA) that contains an NheI restriction site to allow for quick screening of transformants (Fig 3.4.2.1). Mutagenesis was carried out using inverse PCR (Section 2.2.2.4), using primer pairs 33-64 (Table 2.1.5.2), and pBS-*rpoBp1U* as template (see Section 3.2.3). Note that pBS-*rpoBp1U* does not include the *rpoB* RBS, since these transcriptional fusions rely on the *luxA* RBS for translation. Mutated promoter fragments were subcloned as a BglII/EcoRI fragments into BamHI/EcoRI -digested pIJ5972, then conjugated into *S. coelicolor* M145 (Sections 2.2.2.1 and 2.2.4.5).

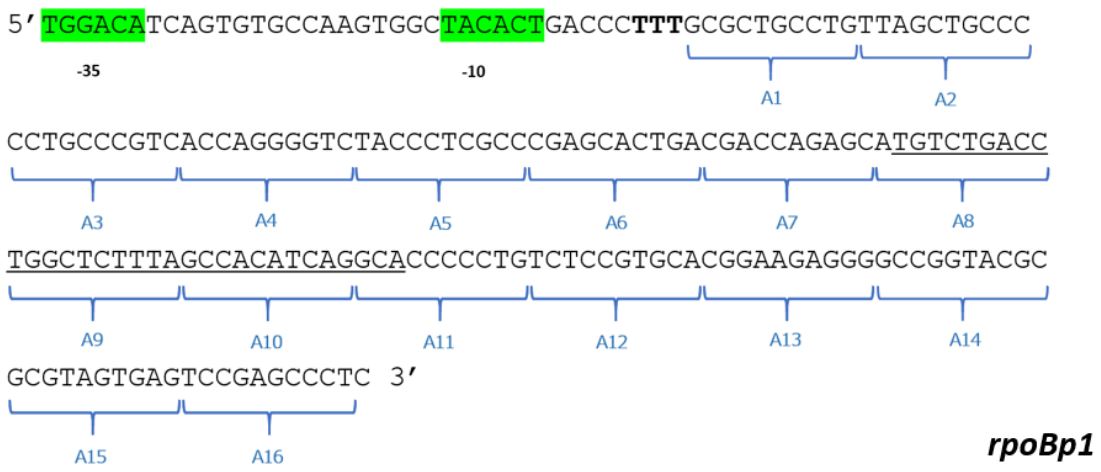


Figure 3.4.2.1: Mutagenesis of the *rpoB* 5'UTR. The 16 consecutive mutations were carried out in the pIJ5972*::rpoBp1* constructs, by replacement of native sequence with 5' CAGCTAGCCA, which contains an NheI restriction site. The underlined region spanning mutations A8 to the beginning of A11 represents the inverted repeat sequence of *rpoBa*.

3.4.3 Basal expression of *rpoBp1* 5'UTR mutant promoters

The *S. coelicolor* M145 strains containing each pIJ5972-*rpoBp1U* mutant, along with WT control, were grown to mid-log phase in NMMP medium and luciferase assays were performed, and results normalised to OD450 (Section 2.2.9) (Fig 3.4.3.1).

As the inclusion of the 5'UTR causes a drop in expression (Fig 3.2.3.2), it was anticipated that several mutations would increase/restore expression by eliminating the effect of potential attenuators. Mutations A1, A3, A4, A10, A11, A14, A15 and A16, all significantly increased expression of *luxAB*, which suggests that there are several cis-acting negative regulatory elements within the 5'UTR. For simplicity, these are considered as four groups based on adjacent mutations.

A1: the A1 mutation of the first 10 transcribed bases from *rpoBp1*, displayed the largest increase in luciferase activity. This mutation disrupts a minor predicted stem loop structure (Fig. 3.4.1.2), which might account for this. However, considering the hypothesis that the +4 G nucleotide might control reiterative transcription at *rpoBp1* (see Section 4.2), it is important to note that the +4G has been changed to +4C in the A1 mutation and this in itself might have a significant impact on transcription; this possibility is considered in detail in Chapter 4.

A3-A4: the ~1.5 – 2-fold increase in expression seen for these mutations might be accounted for by the disruption of the putative stem-loop structure covering these residues. This stem-loop has a 6bp stem, a predicted folding free energy of -9.34 kcal/mol and is highly conserved (Fig 3.4.1.2.)

A10-A11: these mutations disrupt the *rpoBa*. However, other mutations that would be expected to disrupt the secondary structure had little effect; A8 was unaffected, and A9 showed a small increase in expression that did not meet the statistical significance test.

A14-A16: these mutations disrupt the potential secondary structure located immediately upstream of the *rpoB* RBS. However, since the *luxA* RBS is located ~46 nt downstream from the cloned promoter DNA, it seems unlikely that translation of *luxAB* would be affected, and that the observed influence is therefore likely to be at the transcription level. Inclusion of the RBS into the RNA folding algorithm was shown to prevent the formation of this secondary structure spanning the A14-A16

mutations, however, resulted in the formation of a less probable structure (see appendix Fig 7.7). Nonetheless, it is possible that in the native context such potential secondary structures could affect ribosome recruitment, which is highlighted as further work.

Aside from mutations increasing expression, A6 and A7 significantly decreased expression in comparison to the wildtype *rpoBp1U* promoter region (Fig 3.4.3.1). This contradicts the increased expression by A10 and A11 mutations, that also reside in the *rpoBa* sequence, and may suggest the formation of an anti-terminator around this region (Figs 3.4.1.2 and 3.4.3.1). However, upon the study of secondary structure within RNA spanning the TSS (+1) to the middle of *rpoBa* (signified by TTT, +90), no probable structural formations were observed (data not shown).

Overall, these experiments provide evidence for attenuation mechanisms within the 5'UTR of *rpoBC* as previously reported and suggest the action of multiple cis acting regulatory elements within the 5'UTR (Newman, Linn and Hayward, 1979; Küster, Piepersberg and Distler, 1998). The results displaying both the upregulation and downregulation of *LuxAB* by differing mutants, also suggests that more than one regulatory element is acting within this 161 bp region. Further site-directed mutagenesis, resulting in the weakening of secondary structure would strengthen the evidence for secondary structure based regulation, which has been highlighted as future work.

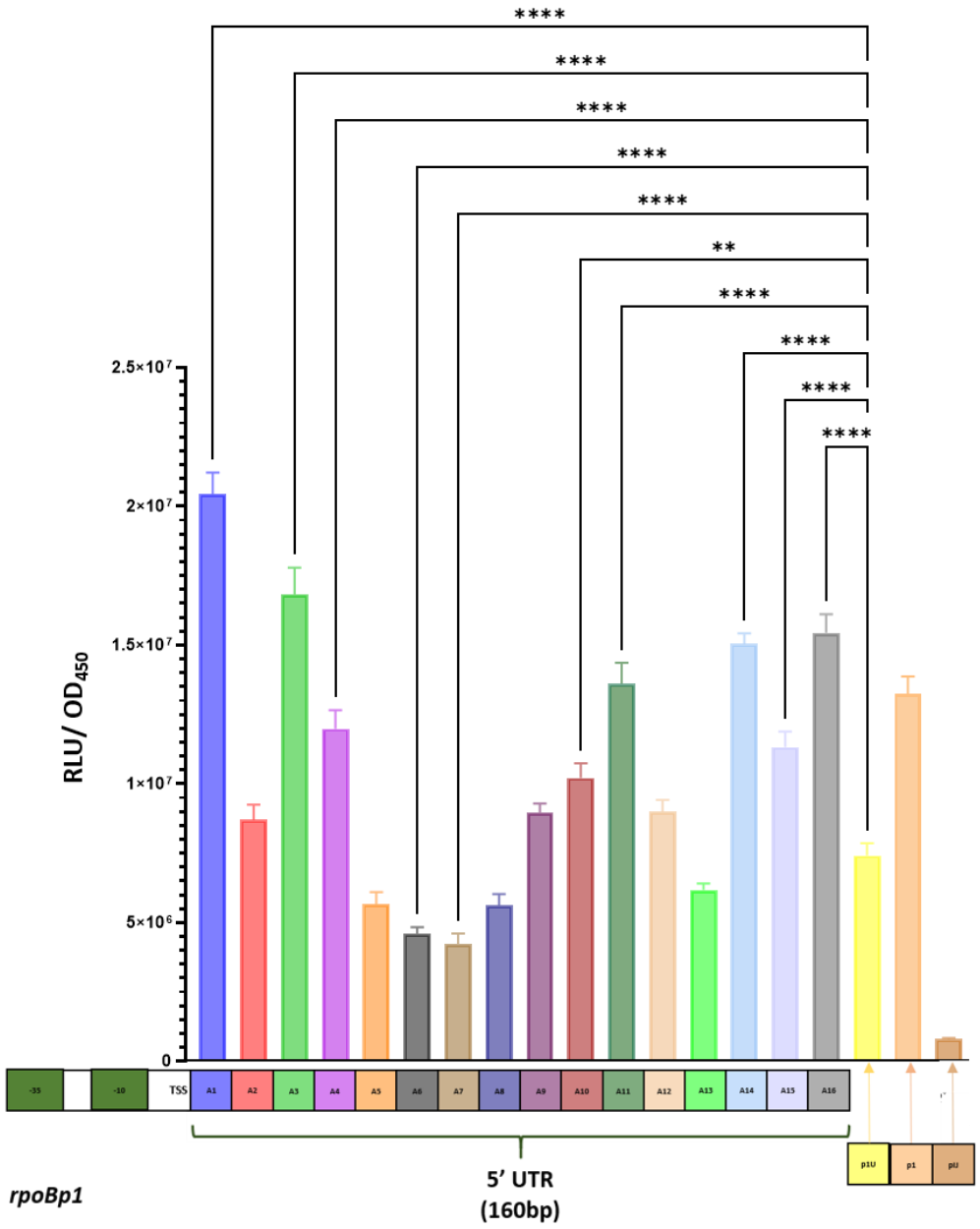


Figure 3.4.3.1: The luciferase activity of *rpoBp1* constructs containing differing 5'UTR mutations. A1 – A16 represent the 16 consecutive mutations introduced into the 5'UTR region. Luciferase activity of the *rpoBp1U* (p1U), *rpoBp1* (p1) and the empty vector pIJ5972 (pIJ) have been added for reference. All luciferase assays were carried out in strains grown at least in biological triplicate. All RLU readings were taken in technical triplicate and normalised using OD450. The error bars represent the S.E.M. Significance values were calculated for each mutation when compared to the *rpoBp1U* construct expression. Due to standard deviations being significantly different we performed a Brown-Forsythe and Welch ANOVA for statistical analysis; $P < 0.05$ = *, $P < 0.005$ = **, $P < 0.0001$ = ****.

3.5 Discussion

Relatively little is known about how bacteria control the production of the core RNAP subunits ($\beta\beta\alpha_2\omega$), and therefore the transcriptional capacity of the cell. Where studied, it is thought that RNAP assembly is limited by β and β' synthesis, since α was present in > 2-fold excess compared to the large subunits in *E. coli* (Engbaek, Gross and Burgess, 1976; Dykxhoorn, St. Pierre and Linn, 1996; Shepherd, Dennis and Bremer, 2001; Izard *et al.*, 2015). Thus, the control and expression of the ubiquitous *rpoBC* operon is crucial in determining overall transcriptional activity. In organisms such as *Streptomyces*, that undergo metabolic and physiological differentiation, when primary growth has ended, the availability of RNAP may play an important role in the remodelling of the transcriptome, including the expression of otherwise silent biosynthetic gene clusters. Studies have largely associated this control as post-transcriptional, as a result of feedback from the availability of the large subunits (Wei *et al.*, 2017). However, due to increased production of antibiotics, resultant from *rpoB* mutations (Hosaka *et al.*, 2009), as well as in response to stress (Chakraburtty and Bibb, 1997; Gomez-Escribano *et al.*, 2008), it was of keen interest to investigate transcriptional control of the *rpoBC* operon during varying conditions.

3.5.1 *rpoBC is* transcribed independently of *rplJL* from three promoters

In *E. coli*, transcription of *rpoBC* is determined by two strong promoters residing upstream of *rplK* and *rplJ*, with transcription decoupled from the ribosomal protein genes by an intergenic attenuator (Steward and Linn, 1991). This differs in the case of the Gram-positive *B. subtilis* where *rpoBC* is expression from a strong promoter just upstream of the *rpoB* gene (Boor, Duncan and Price, 1995), similar to that reported in *S. griseus* (Küster, Piepersberg and Distler, 1998). However, TSS-mapping (Jeong *et al.*, 2016a) revealed two further transcription start points in *S. coelicolor*. Using alignments and known consensus sequences for promoter regions, these correspond to three promoters, with considerable conservation across the *Streptomyces* genus: two putative σ^{HrdB}-dependent (*rpoBp1* and *rpoBp2*) and one putative σ^{ShbA}-dependent (*rpoBp3*) (Figs 3.2.2.1, 3.2.2.2 and 3.2.2.3). As might be expected, transcriptional fusions revealed an increase in expression as the number of promoters included in constructs increased (*rpoBp1U* < *rpoBp1-2U* < *rpoBp1-3U*) (Figs 3.2.3.2 and

3.2.4.1). The use of three independent promoters to control *rpoBC* expression provides a possible means to respond to different intracellular or extracellular conditions, integrating different signals to meet cellular demand. However, the control of *rpoBp2* and *rpoBp3* have yet to be analysed.

Multiple promoters also control *rpoBC* in *M. tuberculosis*, and are differentially expressed in response to antibiotic stress (Zhu *et al.*, 2018); upon rifampicin treatment the main promoter is rapidly downregulated, however expression of *rpoB* is continued via increased transcription from the second promoter. The possible control of *rpoBp3* by σ^{ShbA} is of particular interest and may provide a means to coordinate production of the principle sigma factor σ^{HrdB} with core RNAP. Interestingly, both the σ^{shbA} promoter for *HrdB* in *S. coelicolor* (Newell *et al.*, 2006) and the expression of the *M. tuberculosis sigJ* factor (Hu and Coates, 2001) are unaffected by rifampicin treatment, providing further evidence for differential promoter expression, however this needs to be confirmed Transcriptional read-through from upstream ribosomal protein genes, similar to that observed in *E. coli*, also cannot be ruled out, and requires further experiments to disprove this.

3.5.2 *rpoBp1* is the main promoter for *rpoBC* expression during exponential growth

Transcriptional fusion experiments with truncated promoter sequences revealed that *rpoBp1* is the main promoter under standard laboratory conditions (Fig 3.2.3.2). This promoter corresponds to the 5' end point with the highest number of sequencing reads for *rpoBC* transcription in *S. coelicolor* (Jeong *et al.*, 2016a), and is homologous to that identified in *S. griseus* (Küster, Piepersberg and Distler, 1998), with high sequence conservation observed in all *Streptomyces* organisms analysed (Fig 3.2.2.1). Consistent with a key role for *rpoBp1*, *rpoBp2* is relatively poorly conserved, present in only 11 out of the 16 *Streptomyces* genomes aligned (Fig 3.2.2.2), and although *rpoBp3* is very well conserved, its activity appears to be weak (Figs 3.2.3.2 and 3.2.4.1).

3.5.3 The *rpoBp1* promoter is stringently controlled

Analysis of *rpoBp1*-luxAB transcriptional fusions using qRT-PCR revealed that this key promoter is subject to negative stringent control in a RelA-dependent manner (Fig 3.3.5.1). The failure to identify the *rpoBC* operon as being subject to negative stringent control in previous work (Hesketh *et al.,* 2001) might be due to the additional *rpoB* promoters masking or compensating the effects on overall transcription.

The ppGpp effector of the stringent response is unable to bind to RNAP in *Streptomyces* and is hypothesised to alter expression indirectly via decreased GTP levels as observed for other Gram-positive bacteria (reviewed by Sivapragasam and Grove, 2019; Corrigan *et al.*, 2016). In *B. subtilis* both the consumption of GTP for the production of ppGpp, and the inhibition of GTP biosynthesis enzymes results in reduced intracellular GTP which in turn facilitates the rapid downregulation of genes where GTP is the iNTP (Krásný *et al.*, 2008; Kriel *et al.*, 2013, 2014). Recent work in *B. subtilis* has also shown that mutations in both *rpoB* and *rpoC,* are capable of suppressing the phenotype of ppGpp-null mutants (Osaka *et al.*, 2020). These mutations, located near the active site of the RNAP, are thought to cause a iNTP preference activating and inhibiting promoters with +1A and +1G as the TSS respectively, mimicking transcriptional changes induced by the stringent response (Lane and Darst, 2010). The depletion of GTP also results in the downregulation of rRNA operons in *S. aureus* where promoters initiate with GTP, suggesting iNTP stringent control is common amongst Gram-positive organisms (Geiger *et al.*, 2012). A simultaneous increase and decrease in ppGpp and GTP respectively, has already been observed in both *S. coelicolor*, *S. clavuligerus* and *S. griseus* (Gomez-Escribano *et al.*, 2008; Hesketh *et al.*, 2007; Ochi, 1987; Strauch *et al.*, 1991), hence, the identity of the iNTP may also prove crucial for determining downregulation in *Streptomyces* during stress. Promoters of both *rrnD* and *hrdB*, which contain GTP as the iNTP (data not shown) (Jeong *et al.*, 2016), are rapidly downregulated during the stringent response in *S. coelicolor* providing evidence for the importance of GTP in gene regulation. However, this is not the case for the *rpoBp1* promoter, where transcription unusually initiates at +1T, with GTP present downstream at +4, suggesting an alternative mechanism of downregulation exists here (Fig 3.2.2.1). It is also unlikely

that ppGpp solely affects GTP levels as a means to elicit a stringent response (Corrigan *et al.*, 2016).

3.5.4 The *rpoBp1* promoter is down-regulated during disulphide stress

This work confirmed earlier studies that showed that *rpoBC* was downregulated during oxidative stress (Kallifidas *et al.*, 2010), and extended this by showing that *rpoBp1* governs this response. In a very similar manner to the earlier study, expression was restored by 60 min (Fig 3.3.3.1), which is likely to reflect the recovery of reducing redox conditions following the complete consumption of diamide (Kallifidas *et al.*, 2010). However, the mechanism by which *rpoBC* and other growth-related genes are down-regulated during oxidative stress is unknown, although is independent of the key disulphide stress regulator σ^R (Kallifidas *et al.*, 2010). Similarly, in *B. subtilis*, the majority of genes that are subject to negative stringent control are also downregulated by diamide (Ole Leichert, Scharf and Hecker, 2003) suggesting that diamide might induce a stringent response either through the production of ppGpp or by reducing GTP directly. The fact that (p)ppGpp levels were reported to remain unchanged during oxidative stress (Paget and Hesketh, unpublished results; Kallifidas *et al.*, 2010), suggest that a direct effect on GTP levels might underlie the response. One possible mechanism is through inosine monophosphate dehydrogenase (GuaB), the enzyme that catalyses the conversion of inosine 5'-phosphate (IMP) to xanthosine 5'-phosphase (XMP), and is the first committed step for the production of guanine nucleotides. GuaB has an active site cysteine thiol, which is readily modified by thiol oxidising agents (Hochgräfe *et al.*, 2007; Park and Roe, 2008; Imber, Pietrzyk-Brzezinska and Antelmann, 2019). Indeed, the *guaB* gene is upregulated by σ^R during disulphide stress which might reflect the need to replenish active levels of the essential enzyme during this stress. Interestingly, GuaB has also been identified as a target of ppGpp in *Streptomyces griseus* (Ochi, 1987; Sivapragasam & Grove, 2019), as well as within the *Bacillus* genus (Osaka *et al.*, 2020), suggesting this mechanism may be well spread within Gram-positive organisms. Similar to the stringent response (Hesketh et al., 2007), disulphide stress also results in the downregulation of the ATP synthetase operon as well as genes associated with protein synthesis (30S and 50S proteins) (Kallifidas *et al.*, 2010). Consequently, both responses to nutrient and

disulphide stress may work in *Streptomyces* through alterations to GTP levels; the role of the conserved +4G residue is investigated in the next chapter.

3.5.5 Negatively-acting regulatory elements are present in the 5'UTR of *rpoBC*

Initial promoter assays revealed a large drop in expression upon the inclusion of the 5'UTR (Figs 3.4.4.1 and 3.2.4.1) suggesting the presence of an attenuator or transcriptional terminator within this region. Attenuation of *rpoBC* is particularly important for the decoupling transcription from that of *rplJL* in *E. coli*, with 80% of transcripts terminated just upstream of *rpoB* (Newman, Linn and Hayward, 1979; Steward and Linn, 1991). However, in several Gram-positive bacteria, most transcription initiates from independent *rpoBC* promoters, although attenuation of the resulting transcription remains a possibility.

The inverted repeat that was proposed to act as an attenuator upstream of *rpoBC* in *S. griseus* (Küster, Piepersberg and Distler, 1998) is not as well conserved as other regions of the UTR and shared only 57% homology to the same region in *S. coelicolor* (Fig 3.4.4.1). Mutagenesis of this stem loop structure however, didn't result in restoration of expression levels, with 2/3 of the mutations that lie within this structure (A8, A9 and A10), having little effect on *rpoBp1* promoter expression (Fig 3.4.3.1). This suggests that either the stem loop has minimal effects on expression, or that *in vivo*, this structure is not truly formed. Alternatively, it is possible that our mutations were not sufficient to fully disrupt the secondary structure, or other anti-termination mechanisms reside in this 5'UTR. In *B. subtilis* for example, a trp RNA-binding (TRAP) attenuation protein is capable of binding within the 5'UTR of the *trpE* transcripts causing the formation of a hairpin that sequesters access to the RBS (Merino, Babitzke and Yanofsky, 1995). Elongation factors, NusA and NusG (see Section 1.2.5.1) are able to facilitate pausing of the RNAP in this 5'UTR, increasing the likelihood of TRAP binding and terminating transcription (Mondal, Yakhnin and Babitzke, 2017). Similar attenuation has been hypothesised for the *trp* operon in the *Streptomyces* genus and may be speculated to be a more common form of post-transcriptional control in this genus. The increase in expression observed for the A14 and A16 mutations may be an example of a related mechanism of attenuation,

affecting ribosomal recruitment. This is also supported by the apparent secondary structure formed around the A14-A16 mutations (Fig 3.4.1.2).

Riboswitches, often located in the 5'UTR of mRNA, are also known to modulate gene expression, by direct binding of metabolites to these cis-elements, which further alters the structure of RNA and determines the regulatory response (Garst *et al*., 2011; Breaker, 2012). In *B. subtilis* several operons are regulated by attenuators at regions similar to *E. coli* counterparts however rely more heavily on riboswitches and/or RNA binding proteins for control (Johnson *et al.,* 2020). Few examples of riboswitches have been characterised in the *Streptomyces* genus (Tezuka and Ohnishi, 2014; Busche *et al.,* 2016), however, computational analysis detailed several forms of riboswitches may be present (Seliverstov *et al.,* 2005). More recently a cyclic di-AMP-responsive ydaO riboswitch motif was identified in the 5'UTR of the *rpfA* gene, and may implicate a more widespread regulatory role of the messenger cyclic di-AMP, already known to control development in *Streptomyces* (Hengst *et al.*, 2010; St-Onge *et al.*, 2015; Sivapragasam and Grove, 2019). The fact that cyclic di-AMP is synthesised from two GTP molecules further relates production to intracellular GTP concentration and potentially the purine based stringent response thought to be apparent in *Streptomyces* (discussed in Section 1.2.4.2). It hence is entirely plausible that a similar mechanism of gene regulation is acting in the 5'UTR of *rpoB* to control its expression, supported by the rapid drop in expression observed during the stringent response previously (Fig 3.3.5.1).

Overall, the variable changes in expression observed for mutations spanning the *rpoB* 5'UTR provides initial evidence for control by multiple cis/trans acting elements, however, this would need to be confirmed with further experiments.

Summary

The mechanisms that control RNAP concentration within the cell are not well understood in most bacteria, including organisms belonging to the *Streptomyces* genus. Within this chapter we have successfully characterised several elements of control of the *rpoBC* operon, encoding the large β and β' subunits of this enzyme, that are likely rate-limiting in regard to formation of the core enzyme (Dykxhoorn, St.

Pierre and Linn, 1996), including 3 promoters, and several areas of regulation within the 5'UTR.

Mutations within the RNAP, more specifically *rpoB* have already been associated with increased expression of antibiotic production, and hence this further understanding of RNAP control may assist in activation of silent biosynthetic gene clusters for the production of novel antibiotics (Hosaka *et al.*, 2009).

Chapter 4: Reiterative transcription at the *rpoBp1* promoter

4.1 Overview

Reiterative transcription (RT) is a mechanism of gene regulation that results from the "slippage" of RNAP, causing either the loss or additional nucleotides to be encoded (reviewed by Turnbough (2011), Liu *et al.*, 2022). These slippage events occur on homopolymeric tracts in the template DNA, consisting of three or more of the same nucleotide, and cause the RNAP to slip when transcribing, due to unstable RNA:DNA hybrids within the active site of the enzyme (see Fig 1.2.5.1.1) (Cheng, Dylla and Turnbough, 2001; Han and Turnbough, 2014; Molodtsov *et al.*, 2014). This mechanism itself is sensitive to the surrounding concentration of nucleotides and can result in either transcription termination, by release of the nascent RNA, or return to canonical transcription producing full length transcripts (Han and Turnbough, 1998; Molodtsov *et al.*, 2014). However, only a few examples of RT have been reported in varying organisms, most involving nucleotide biosynthetic genes, and none of which belong to the *Streptomyces* genus.

Within this chapter, TSS mapping data was used to confirm RT at the main promoter for the *rpoBC* operon (*rpoBp1*). Experiments also focused on understanding the role RT might play in controlling expression from *rpoBp1 in vivo*, during both exponential growth and in response to stress. Mutations disrupting the slippage prone TSS for *rpoBp1* were introduced using CRISPR-cas, with *rpoBp* transcriptional-fusions used to assay expression changes. *In vitro* transcription assays were also used to investigate slippage events and nucleotide sensitivity of the *rpoBp1* TSS.

4.2 Reiterative transcription occurs at *rpoBp1 in vivo*

Previously, RNA-sequencing data from both *S. coelicolor* and *S. venezuelae,* confirmed the largest proportion of transcripts initiated at the *rpoBp1* promoter (see Section 3.2.1), at an unusual +1 UTP TSS (Fig 3.2.2.1). The +1 UTP is the first nucleotide in a highly conserved homopolymeric T-tract (Fig 4.2.1.1); tracts of at least three of the same nucleotide is sufficient for RT to occur during transcription initiation (Xiong and Reznikoff, 1993). To identify and study the extent of RT at the *rpoBp1* promoter, transcripts initiating at this promoter were extracted from TSS-mapping data from *S. coelicolor* (Jeong *et al.*, 2016b) (NCBI bioproject PRJNA285265 and

Illumina MiSeq SRA run SRR2043956, and Section 2.2.3.4) (Ettwiller *et al.*, 2016) and *S. venezuelae* (Bush *et al*., pers comms)(as used in Section 3.2.1); to prevent bias, a 30 bp sequence containing the second GTP after the TSS (T-tract) (5' GCTGCCTGTTAGCTGCCCCCTGCCCGTCAC) and downstream DNA was used extract identical sequences from the raw sequencing data using the Galaxy platform (see Fig 4.2.1.1) (Afgan *et al.*, 2018). Data was organised according to the number of U's at the 5' end of the transcript, indicative of slippage, and plotted as a percentage of total transcripts, which can be seen in Figure 4.2.1.1 for *S. venezuelae* and Figures 4.2.1.2 and 4.2.1.3 for *S. coelicolor*. Transcripts were also observed to initiate on the CTP and GTP directly preceding and after the T-tract respectively (5' **C**TTT**G**), however, they did not represent a significant proportion.

RT at the *rpoBp1* promoter is a common occurrence for both *S. coelicolor* and *S. venezuelae,* with the highest abundance of transcripts containing four U's at the 5' end, representing at least one slippage event. A similar trend for slippage is observed for both organisms, however transcripts initiating at the CTP preceding the T-tract is only observed for *S. venezuelae*. These CTP initiated transcripts also displayed evidence for slippage on this T-tract, however due to the small number of transcripts, this was not included in the quantification of RT. Initiation with +1 GTP was also observed for both *Streptomyces* species, which may suggest some selectivity for start site for *rpoBp1,* however the importance of this, and the slippage occurring at the main TSS for *rpoBC* is yet to be elucidated. Due to the small percentage of transcripts initiating on GTP, this phenomenon was not investigated further.

We considered that stress or growth phase that influences NTP levels may affect the rate of RT, and thereby transcription at *rpoBp1*. During stringent conditions, GTP levels have been observed to drop rapidly after the induction of stress, however UTP levels remain unchanged (Strauch *et al.*, 1991). Such changes in [NTP] might promote RT at *rpoBp1* due to GTP deficiency preventing addition to nascent RNA and further commitment to the transcription. TSS data for *S. venezuelae* was collected over a time course into sporulation (10, 14, 18 and 24 h growth), where changes in [NTP] might be expected; however, there was no obvious change in slippage rates from progression from early exponential phase (10 h) to late stationary phase (24h) (individual data not shown). However, it is important to note that TSS analysis can only detect transcription products that arise from promoter escape – therefore slippage that leads

immediately to termination would not be detected; indeed termination has previously been shown to be the most common outcome for UTP-dependent RT (Turnbough, 2008, 2011).

One particularly noticeable result was the increased number of transcripts containing one 5'U nucleotide at transcripts for *S. venezuelae* (~25%), which is much higher than observed for both sets of the *S. coelicolor* data (~2.5-5%), although the basis for this is unknown.

It would be favourable to further confirm the extent of RT at *rpoBp1 in vivo* using such methods as S1 nuclease mapping or primer extension, however, previous attempts using these approached and probes or primers specific to *rpoB* respectively, were unsuccessful (data not shown). This is hypothesised to be a result of the identified secondary structure in the 5' UTR of *rpoBC*, studied in Section 3.4, affecting probe/primer binding.

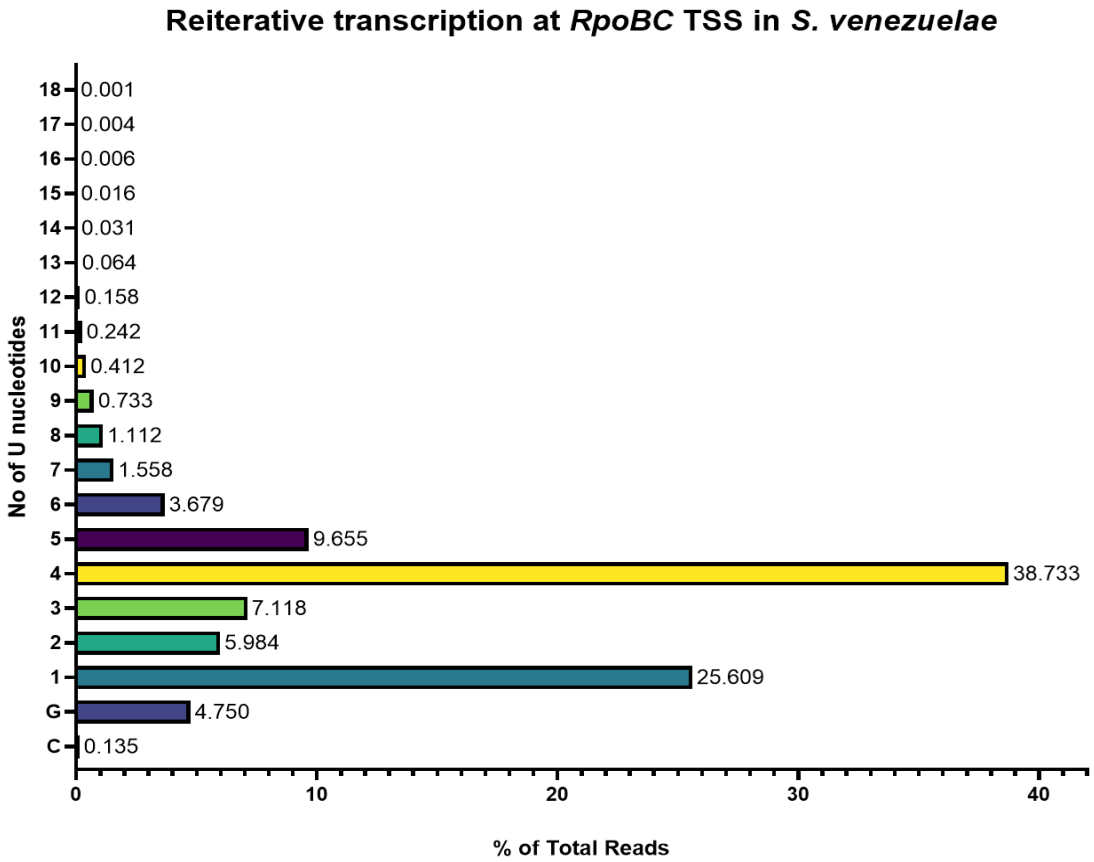


Figure 4.2.1.1: Reiterative transcription at the *rpoBp1* TSS in *S. venezuelae*. Transcripts from TSS-mapping data (Ettwiller *et al*., 2016; Matt Bush, pers comm) were extracted using a 30bp search substring (see Section 4.2). A total of 336690 sequences were extracted from raw fastq data using the galaxy platform (Afgan *et al*., 2018) using an intext (grep) search. Transcripts are organised according to the number of U's at the 5' end (TSS) of the transcript, as a percentage of the total. All 4 time points from the data set (10,14, 18 and 24h after growth) were used to analyse reiterative transcription. 69 bp of the *rpoBp1* region is shown including the -35 and -10 elements (bold green), the T-tract (bold red) and the sequence used to search transcripts (underlined).

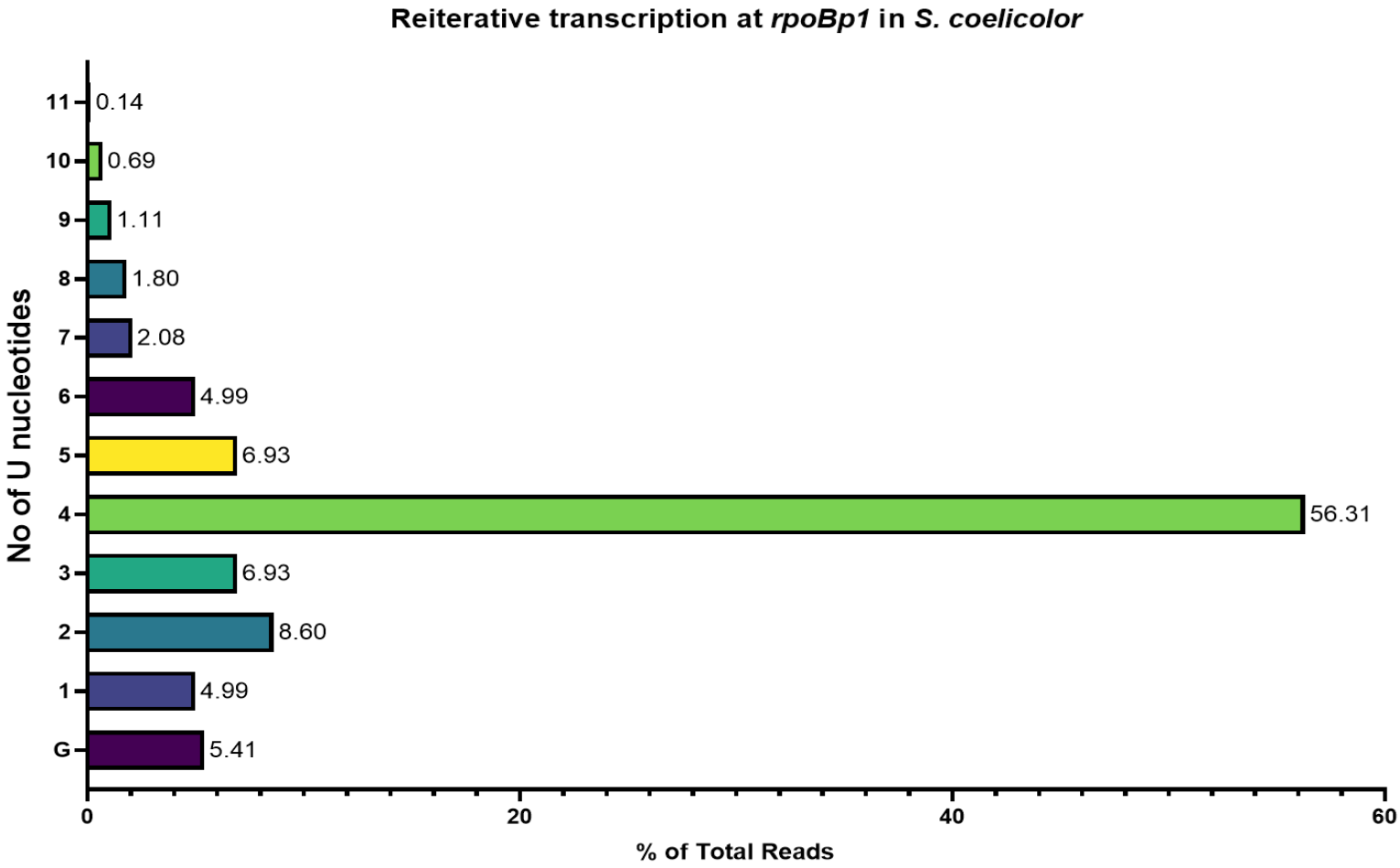


Figure 4.2.1.2: Reiterative transcription at the *rpoBp1* TSS in *S. coelicolor*. A total of 721 transcripts from *S. coelicolor* (Jeong *et al.*, 2016b) (NCBI bioproject PRJNA285265 and illumina MiSeq SRA run SRR2043956) were extracted using a 30bp search substring homologous to *rpoBp* (see Section 4.2). Transcripts are organised according to the number of U's at the 5' end (TSS) of the transcript, as a percentage of total transcripts analysed. Transcripts initiated with the GTP (that is directly after the T-tract) are also included in the analysis.

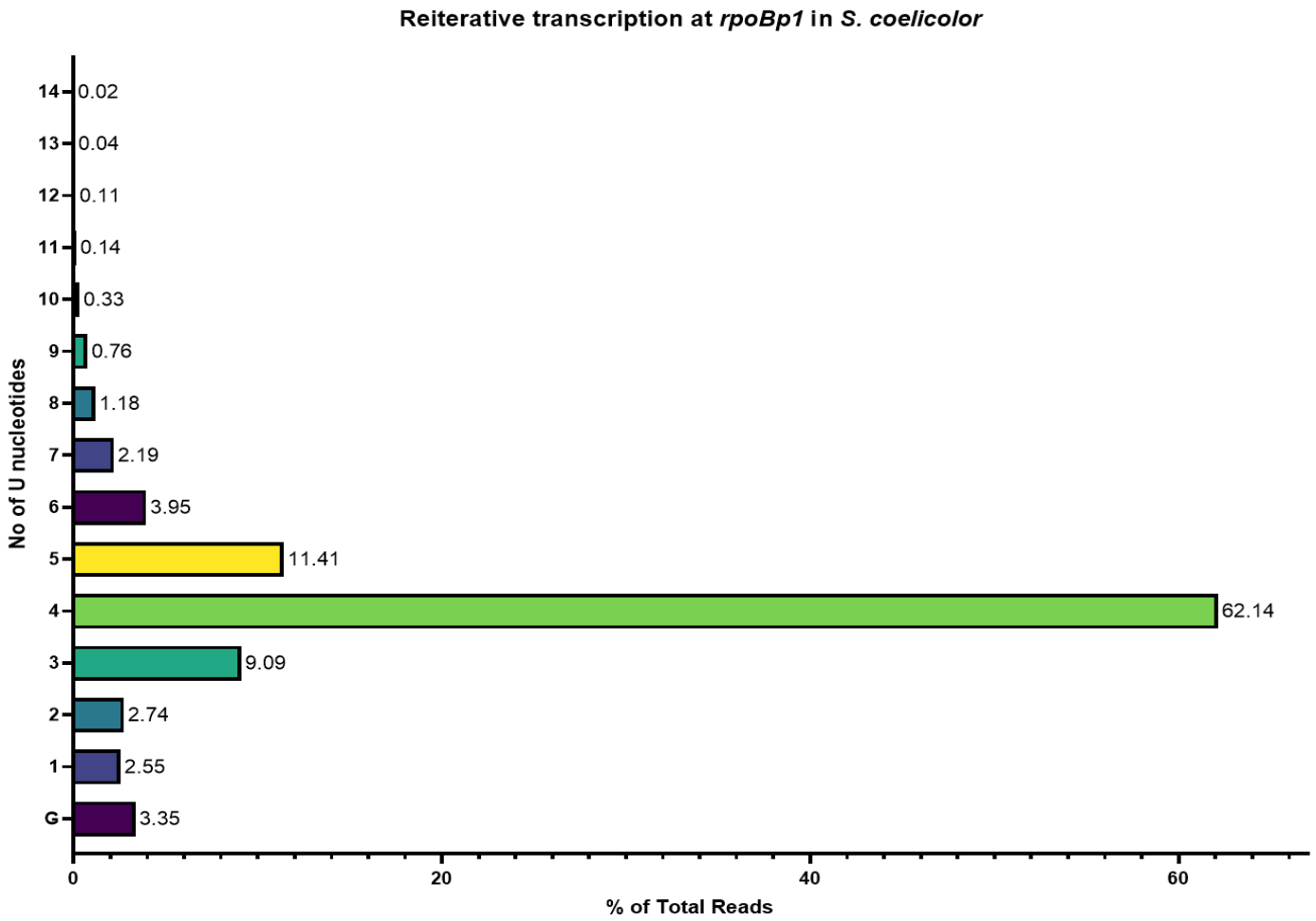


Figure 4.2.1.3 Reiterative transcription at the *rpoBp1* TSS in *S. coelicolor*. Transcripts (5520) from the raw TSS-mapping data (Ettwiller *et al*., 2016), taken during exponential phase (see Section 2.2.3.4), were extracted using the galaxy platform (Afgan *et al.*, 2018) and an intext (grep) search with 30bp of *rpoBp* sequence. Sequences are organised according to the number of U's at the 5' end (TSS) of the transcript, as a percentage of total transcripts analysed. Transcripts initiating at GTP (that is directly after the T-tract) are also included in the analysis.

4.3 Reiterative transcription occurs at *rpoBp1 in vitro*

4.3.1 Development of an *in vitro* transcription method

To confirm RT at the *rpoBp1* TSS, *in vitro* transcription assays were performed using RNAP containing the principle sigma factor σ^{HrdB}; problems with σ^{HrdB} were addressed by developing a new purification strategy.

4.3.1.1 Purification of the σ^{HrdB} sigma factor

It was decided to overproduce σ^{HrdB} as a fusion to the ubiquitin-like protein SUMO, which has been shown to improve solubility (Butt *et al.*, 2005). The sequence encoding σ^{HrdB} was amplified using *S. coelicolor* genomic DNA and primers 75 and 76 (Table 2.1.5.4) placing BglII and HindIII restriction sites after the start codon and after the stop codon, respectively. The amplified fragment was then sub-cloned into pBluescript II SK+, confirmed using sequencing before digested as a BglII/HindIII fragment and subcloned into BamHI/HindIII-digested pET28B-SUMO. In this fusion hrdB is fused in frame to an N-terminal SUMO that additional includes a 6XHis-tag. The constructed pET28B-SUMO::σ^{HrdB} plasmid was transformed into the *E. coli* expression strain BL21 (pLysS) before the protein was expressed, cells lysed and purified using nickel affinity chromatography (Sections 2.2.6.1, 2.2.6.2, 2.2.6.3; Fig 4.3.1.1.1.A).

Elutions from the Ni-NTA column were taken forward for SUMO digestion (Section 2.2.6.6) overnight, followed by subtractive purification using Ni-NTA to capture the 6XHis-SUMO. The flow through, containing σ^{HrdB} with the expected size, was collected as ~1 ml fractions before the addition of 10 ml elution buffer to the column to elute any bound 6XHis-SUMO tag and Ulp protease (Fig 4.3.1.1.1B). To ensure removal of all Ulp protease and any contaminating SUMO, all flow through fractions from the second Ni-NTA purification were taken forward for size exclusion chromatography (HiLoad 16/600 superdex 75 pg), with proteins washed and eluted into 2x storage buffer (Section 2.2.6.5). Fifty 2 ml fractions were collected. All fractions with notable peaks at 280nm were analysed using SDS-page and Coomassie staining (Fig 4.3.1.1.2 A and B). The symmetrical main peak at approximately 50 ml elution volume indicates a monodispersed protein solution; fractions 5-7 appeared as

pure σ^{HrdB} by PAGE and were taken forward for concentrating (Section 2.2.6.8) to a final volume of 500 μL. For further storage glycerol was added to the σ^{HrdB} protein to a final concentration of 50%, flash-frozen in aliquots and stored at -80°C. The σ^{HrdB} protein was stored at a final concentration of 0.8 mg/ml and was included in subsequent *in vitro* transcription reactions.

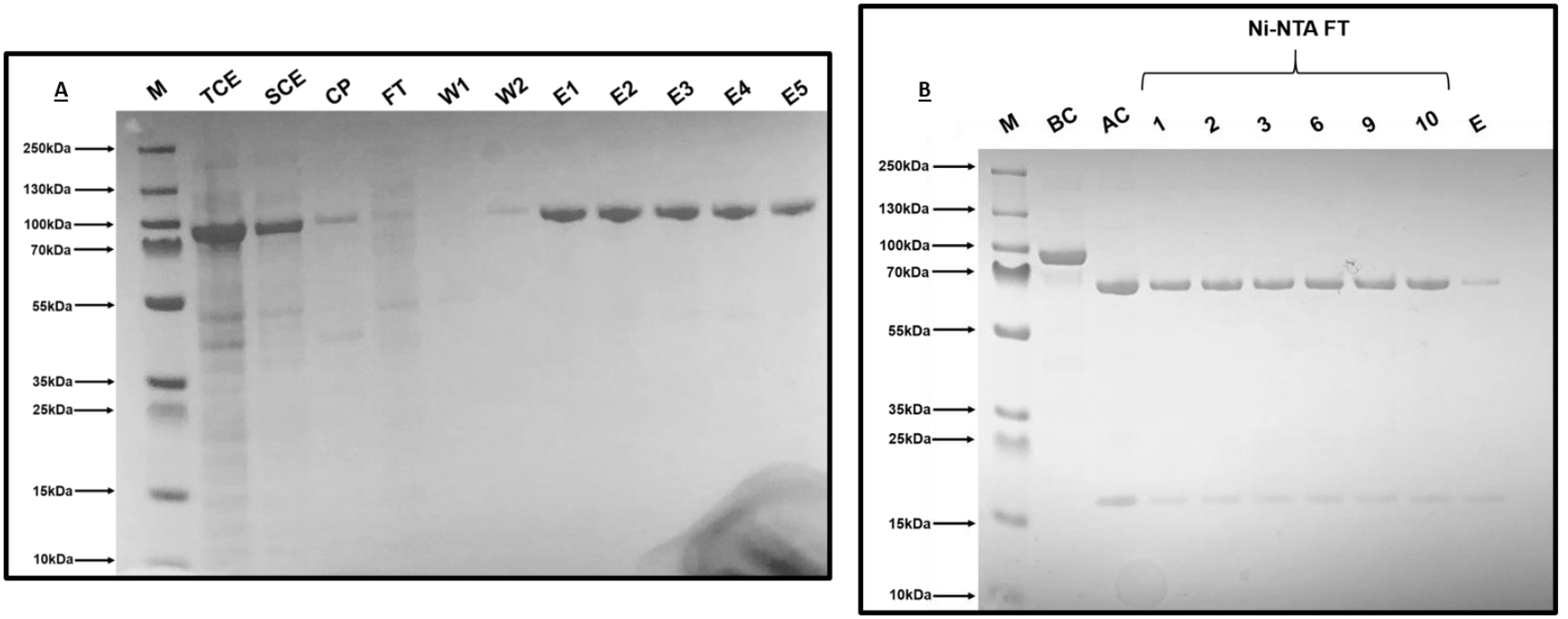


Fig 4.3.1.1.1 The purification of σ^{HrdB} using affinity chromatography. Protein samples (5 µL) were assessed SDS-PAGE separation using a 4-12% gradient gel and separated at 120 V for 1 h 20 min and visualised by Coomassie staining. A) contains fractions from the first Ni-NTA purification. From left to right – Protein marker (M) Total Cell Extract (TCE), Soluble Cell Extract (SCE), Cell Pellet (CP), Flow Through (FT), Washes (W1 and W2, 10mM and 25mM respectively) and elution's (E1-E5, 250mM imidazole). B) containing protein samples collected from the second Ni-NTA purification; samples taken before (BC) and after cleavage (AC) using Ulp protease, flow through fractions containing the σ^{HrdB} protein, and final elution (E, 250mM imidazole).

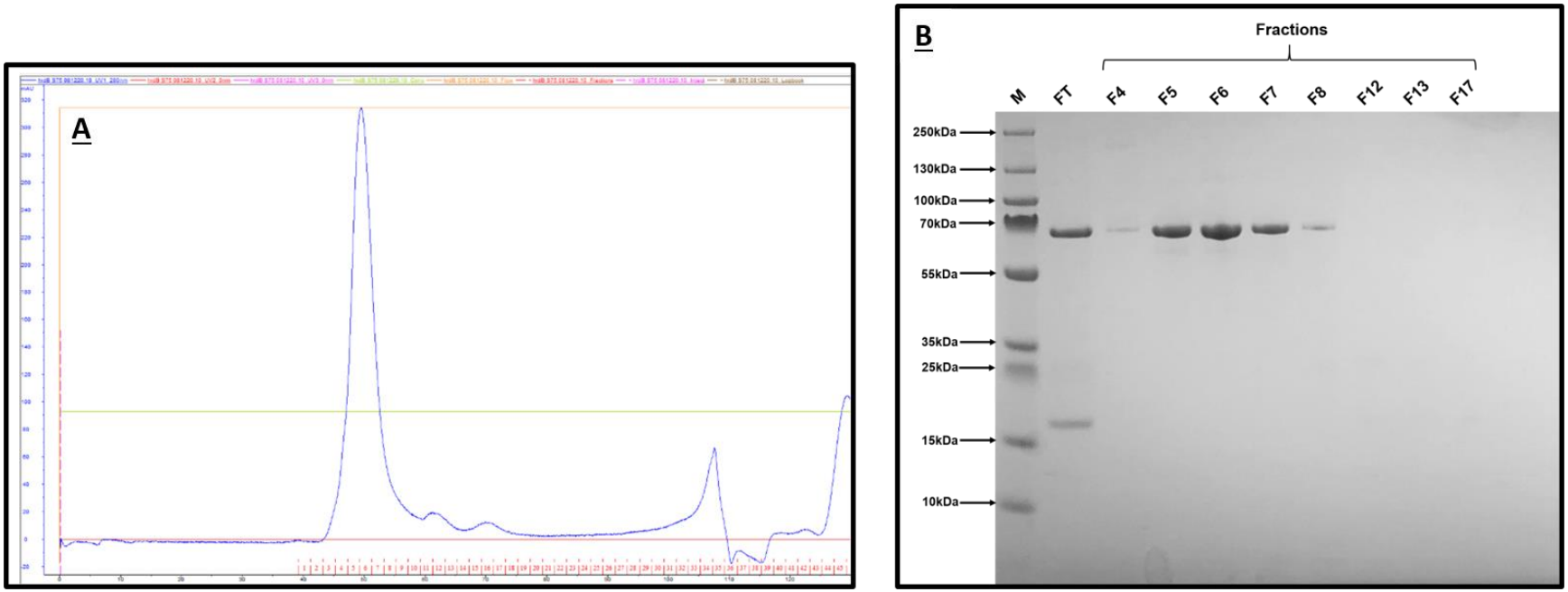


Figure 4.3.1.1.2: The purification of σ^{HrdB} using gel filtration. A) The chromatogram produced from the S75 superdex gel filtration column after the addition of 45 ml of 2x storage buffer. Separate 1 ml fractions were collected. B) SDS-PAGE analysis using a 4-12% NuPAGE gel of fractions from gel filtration (Section 2.2.6.8); Samples (5 µL) were separated at 120 V for 1 h 20 min and visualised by Coomassie staining. From left to right, Protein Marker (M), Flow through (FT) after second Ni-NTA purification, and fractions from collected from size exclusion chromatography (F4-F17).

4.3.1.2 Optimisation of the *in vitro* transcription method

All components used for *in vitro* transcription assays (buffers and concentrations of, NTPs, protein and DNA template) were optimised using the *rplJ* template previously used (Tabib-Salazar *et al.*, 2013), separated on 8% denaturing urea-acrylamide gels and visualised by phosphoimaging (Section 2.2.8.1). Due to the known role of both CarD and RbpA during transcription initiation (Newell *et al.*, 2006; Stallings *et al.*, 2009; Jensen *et al.*, 2019), these proteins were included as standard in assays. Initiation at *rpoBp1* was confirmed by using templates, *rpoBp1*(1) and *rpoBp1*(2) that differ in the extent of downstream DNA (Figs 4.3.1.2.1 and 4.3.1.2.2) (see Section 4.3.2) producing run-off transcripts of ~44 nt and ~74 nt, respectively (see Appendix Figs 7.6.1 and 7.6.2).

Previously purified RNAP holoenzyme from J1981, RbpA (Aline Tabib-Salazar, pers comms) and CarD (Laurence Humphrey, pers comms), were used in all *in vitro* transcription reactions. To confirm additional activity from our purified σ^{HrdB}, we performed assays without this protein. RbpA, CarD and RNAP were also removed from reactions to confirm effects on transcription from the *rpoBp1*(1) template (Fig 4.3.1.2.2). Interestingly, both reactions lacking CarD and RbpA increase transcription compared to standard conditions. CarD has previously been observed to negatively some promoters (M. Paget, pers comm), and CarD and RbpA can have been shown to have negative effects at some promoters in mycobacteria (Jensen *et al.*, 2019). σ^{HrdB} increased transcription from *rpoBp1*, with a slight increase in band intensity, confirming the purified protein was active.

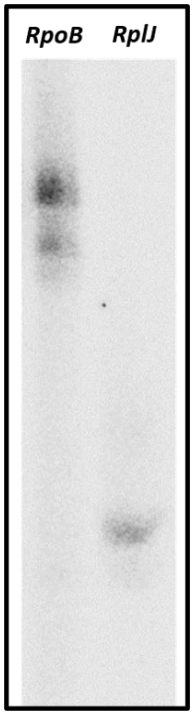


Figure 4.3.1.2.1: *In vitro* transcription analysis at *rplJ* and *rpoBp1* promoter regions. An 8% denaturing urea gel was run for 1h 20 min at 600 V, 50 w at 55°C, and visualised by phosphoimaging. The *rpoBp1(2)* template and *rplJ* templates are around 76 and 44 nt long in transcript length respectively. The *rplJ* and *rpoBp1*(2) templates were amplified using primers 80 and 81 and 78 and 79, respectively and purified by gel extraction (Section 2.2.8.2).

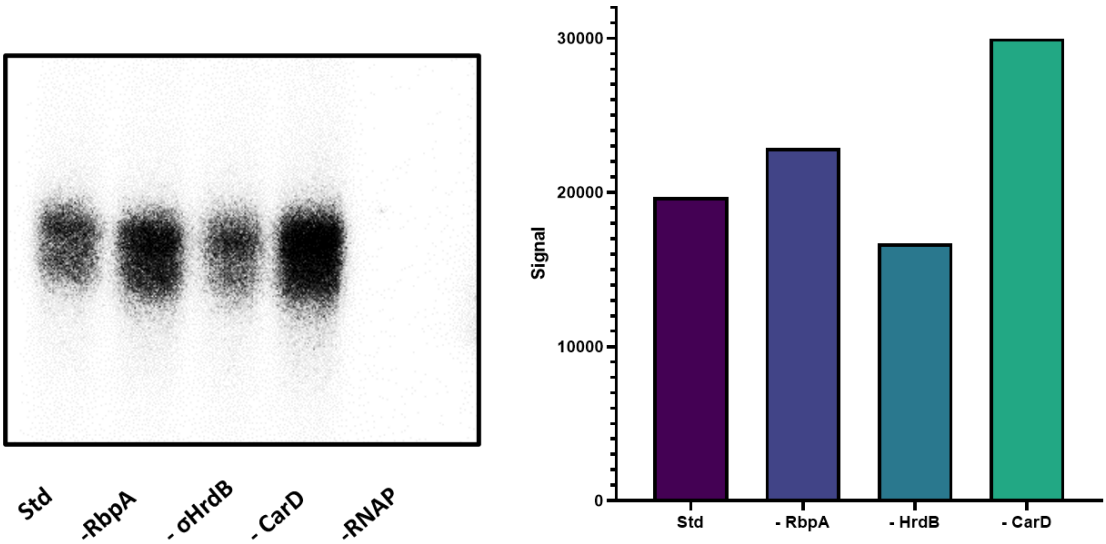


Figure 4.3.1.2.2: The effects of RbpA, CarD, RNAP and σHrdB on in vitro transcription at the *rpoBp1* promoter. An 8% denaturing urea gel was run for 1h 20 min at 600V, 50w at 55°C and visualised by phosphoimaging. The *rpoBp* (1) template was used for *in vitro* transcription assays with band intensity quantified using Image studio lite, and plotted in GraphPad prism. From left to right: Standard conditions (Std), -RbpA, -σ^{HrdB}, -CarD and -RNAP. Intensity was not reported for -RNAP reaction due to a lack of band. The *rpoBp1*(1) template was amplified using primers 77 and 78, and purified by gel extraction.

4.3.2 Reiterative transcription at the *rpoBp1* promoter is UTP-dependent

RNA-sequencing detailed extensive RT occurring at the *rpoBp1* TSS *in vivo* (Section 4.2), however, with standard *in vitro* transcription conditions, no RT was observed (Fig 4.3.1.2.1); a ladder of transcripts or a large smear, depending on denaturing gel percentage, can provide evidence for RT (Qi *et al.*, 1996). RT is usually influenced by nucleotide levels, often those associated with the homopolymeric tract, or the following nucleotide that is required to commit to canonical transcription (Han and Turnbough, 1998). Using an NTP mix containing only cold UTP (100 μM) and additional α-UTP radiolabel (see Section 2.1.8.4.2), *in vitro* transcription was carried out using the *rpoBp1*(1) and *rpoBp1*(2) templates.

UTP alone should facilitate RT from the *rpoBp1* TSS (5' TTT), producing a poly-U product, which was confirmed by a ladder of transcripts produced from *rpoBp1*(1) and *rpoBp1*(2) templates (Fig 4.3.2.1), with each band increasing in length by 1 nucleotide. A separate reaction including all NTPs (and UTP*) in equimolar amounts confirmed full run-off transcription for both *rpoBp1* templates used.

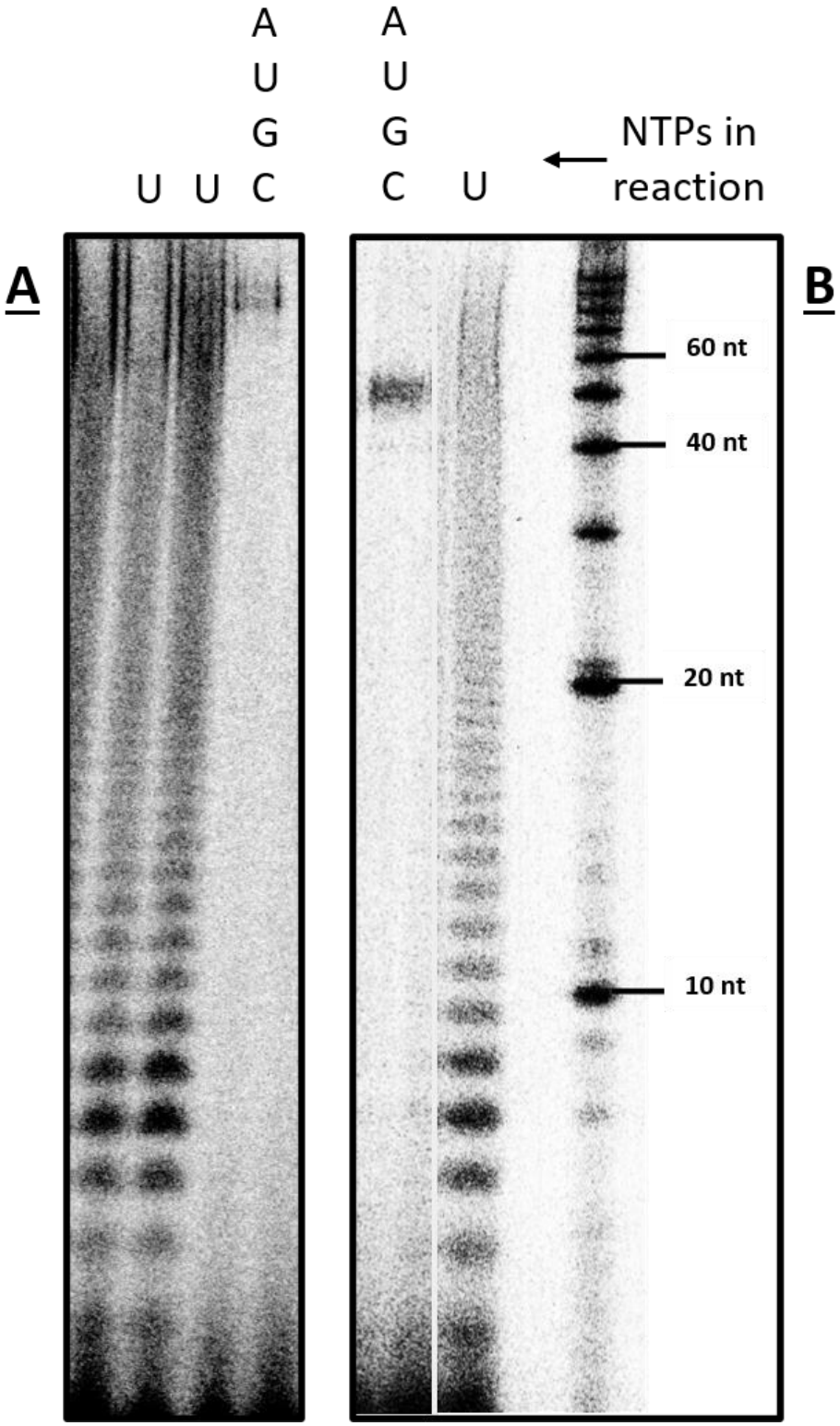


Figure 4.3.2.1: *In vitro* transcription assays showing reiterative transcription at *rpoBp1*. Samples (8 μL) were loaded onto a 24% denaturing urea (7M) gel and separated at 600v, 50w for 2h 45 min at 55°C and visualised by phosphoimaging. A and B represent the transcripts produced from the *rpoBp*(2) (~74 nt, in technical replicate) and *rpoBp*(1) (~44 nt) templates respectively. The same RNA markers (FisherSci, AM7778) were used for gel A, but image was aligned to the ladder in B for reference. A,U,G,C and U signify the nucleotides included in the *in vitro* transcription assays.

4.3.3 Removal of the homopolymeric T-tract results in the abolition of reiterative transcription

In order to confirm that RT on *rpoBp1* templates occurs at the specific identified TSS, the homopolymeric T-tract, *in vitro* transcription assays were carried out as previously (Section 4.3.2) using a DNA template containing a 5'TTTG > 5'**ACT**G mutation; the template was obtained by PCR from the genome of *S. coelicolor* (M145_*rpoB*_ACT, see Section 4.4.3).

While RT was observed for the WT *rpoBp1* promoter template no transcription products were detected in reactions containing the *rpoBp1*-ACT DNA template and UTP nucleotides alone, indicating that RT occurs at the T-tract in WT *rpoBp1* (Fig 4.3.3.1).

.

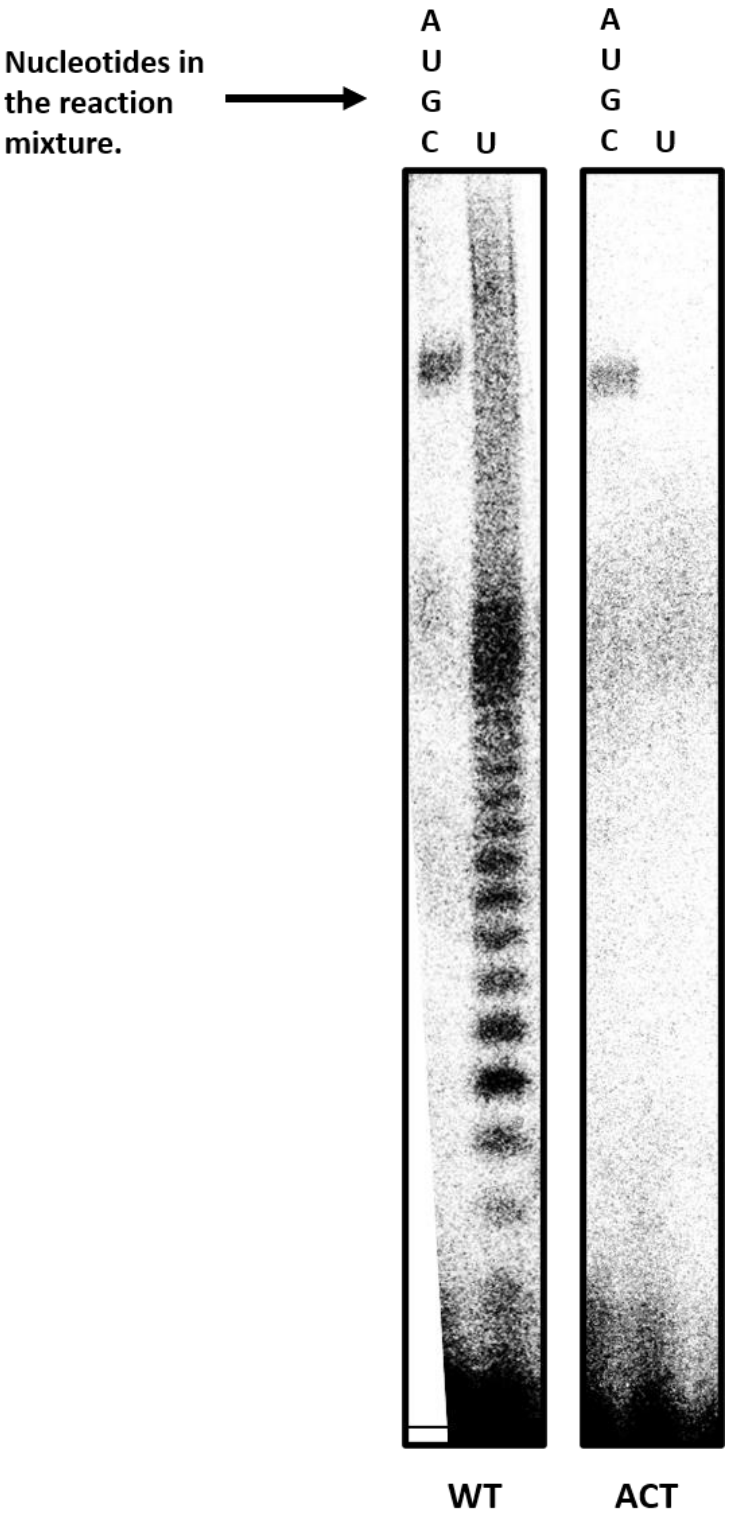


Figure 4.3.3.1: *In vitro* transcription assays from mutant *rpoBp1* TSS. Samples (8 μL) were loaded onto a 24% denaturing urea gel and separated at 600 V, 50 W at 55°C for 2 h 45 min and visualised by phosphoimaging. WT and ACT represent the reactions with the wildtype and ACT mutant DNA templates respectively. A,U,G,C and U signify the nucleotides included in the *in vitro* transcription assays.

4.3.4 Reiterative transcription at *rpoBp1* TSS is sensitive to nucleotide after the homopolymeric tract.

RT is sensitive to intracellular nucleotide concentrations, usually those encoded in the homopolymeric tract, or the following nucleotides encoded in the DNA template that commit RNAP to canonical transcription (Turnbough, 2011). In the case of the *rpoBp1* promoter, the T-tract is followed by a GTP nucleotide (Fig 4.3.4.1), raising the possibility that the promoter and therefore *rpoBC* expression might be modulated by [GTP].

Initial experiments were performed to confirm that the absence of the nucleotide following the T-tract (GTP) would cause RT. Assays were performed using the *rpoBp1*(1) template with an NTP mix (α-UTP) lacking only the GTP nucleotide (Section 2.1.8.4.2). A visual ladder of transcripts, indicative of RT, was observed raising the possibility that *rpoBp1* might be sensitive to GTP levels (Fig 4.3.4.2, lane 2).

Another template, containing a mutation of this (+4) GTP to a CTP nucleotide, TTT**G** > TTT**C**, was used to confirm sensitivity to the nucleotide directly after this T-tract (Fig 4.3.4.1). *In vitro* transcription assays were carried out as above, with the DNA template amplified from plasmid pIJ5972::*rpoBp1U_TTTC* containing this mutation (Section 4.4.12) using primers 77 and 78. Upon the removal of GTP, slippage was no longer observed on the TTTC mutant template, however, RT was apparent in the reaction lacking CTP (Fig 4.3.4.2, lanes 10 and 11). This provides further evidence for sensitivity of RT to levels of the nucleotide encoded directly after the *rpoBp1* T-tract.

RT was not detected in assays using wildtype *rpoBp1* lacking CTP. Since CTP is the second nucleotide encoded after the TSS (+5), this suggests that the nucleotide directly after the T-tract is the most crucial for promoting RT (Figs 4.3.4.1 and 4.3.4.2, lane 3). The removal of GTP from TTTC mutant DNA template containing reactions is also provided for reference, again with no apparent RT.

Overall, these experiments suggest that the intracellular nucleotide concentration of the NTP that is encoded directly after the homopolymeric tract might be an important determinant of RT; however, in order to test this hypothesis further, a range of NTP concentrations should be tested.

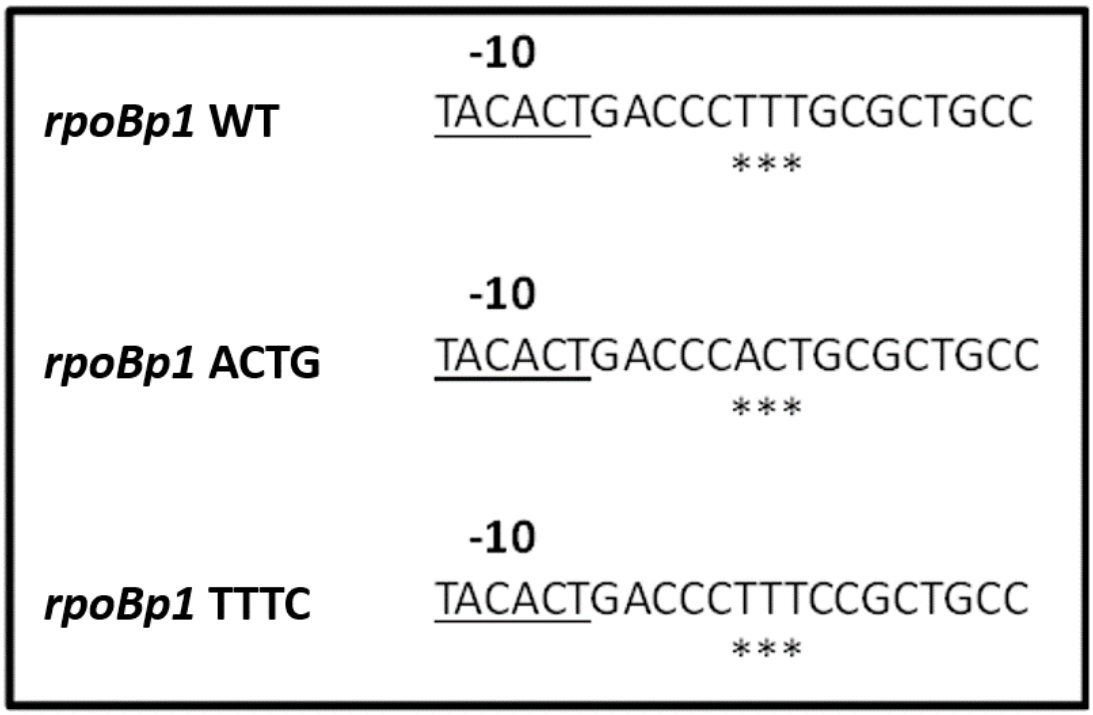


Figure 4.3.4.1: The *rpoBp1* based DNA templates used for *in vitro* transcription assays. The -10 promoter element and TSS of the three differing DNA templates used for the assays are presented. WT represents the wildtype *rpoBp1* promoter containing the homopolymeric T-tract. All mutations were carried out in the *rpoBp1*(1) DNA template. The positioning of the TSS is indicated by ***.

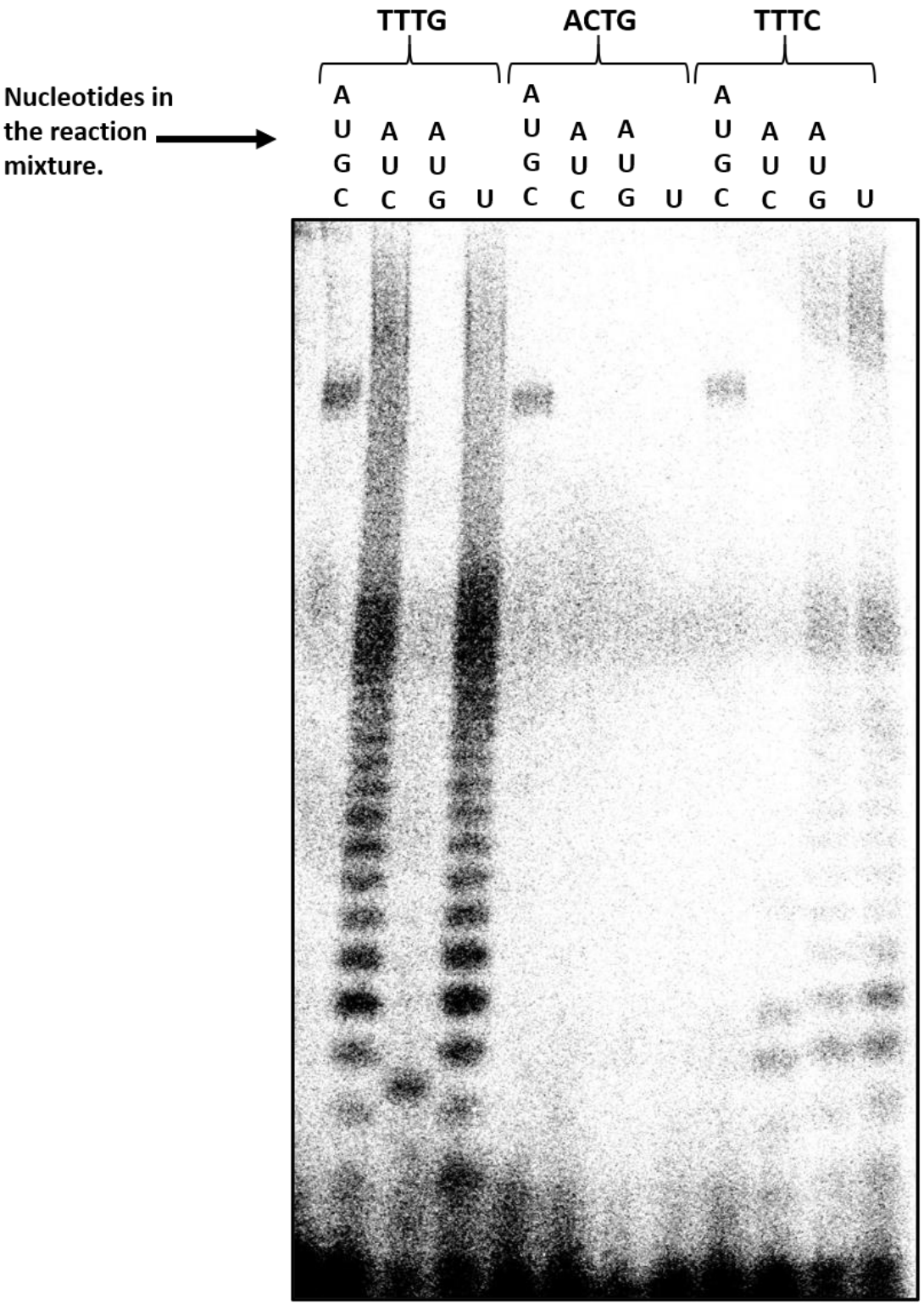


Figure 4.3.4.2: *In vitro* transcription assays of mutant *rpoBp1* templates. Samples (8 μL) were run on a 24% denaturing urea (7M) gel at 600 V, 50 W for 2h 45 min at 55°C, and visualised by phosphoimaging. Lanes 1-4 contain the wildtype *rpoBp1* TSS (TTTG), lanes 5-8 contain the ACTG mutant DNA template, and lanes 9-12 contain the TTTC mutant template. *rpoBp1*(1) was the basis for all templates. The nucleotides included into reaction mixtures are specified above wells.

4.3.5 Depletion of GTP results in increased slippage and a downregulation in transcription at *rpoBp1 in vitro*.

The presence of GTP as the iNTP, has previously been shown to be crucial in *B. subtilis* in controlling expression levels by sensing intracellular GTP concentrations, especially important for downregulation of genes during the stringent response (Krásny and Gourse, 2004; Krásný *et al.*, 2008). The same is not known for *Streptomyces*, yet GTP levels decrease during stringent conditions (Strauch *et al.*, 1991), thus the same mechanism may be apparent in this genus, facilitating the downregulation of genes in response to stress. GTP concentration too, may be of particular importance at the *rpoBp1* promoter, where removal of GTP from *in vitro* transcription assays, caused extensive RT (Section 4.3.4).

To determine the sensitivity of *rpoBp1* to GTP levels, *in vitro* transcription reactions were carried out using concentrations of GTP ranging from 0-1000 μM; remaining nucleotides were maintained at 100 μM and both α-UTP and α-CTP radionucleotides were used to quantify both RT and run-off transcription respectively. All multi-round *in vitro* transcription assays were carried out as previously using the wildtype *rpoBp1*(1) templates (Fig 4.3.4.1).

As might be expected, a gradual decrease in run-off transcript intensity was observed as GTP concentrations were reduced, becoming undetectable at ~3.1 μM (Fig 4.3.5.1B). A ladder of transcripts, indicative of RT, appears in reactions with 6.25 μM GTP or less when assayed with α-UTP (Fig 4.3.5.1A). The extent of slippage seems to drop slightly for the 3.1 μM reaction using the α-UTP, suggesting an anomaly within results due to the RT observed at 6.25 μM.

When transcripts were labelled with α-UTP, high [GTP] reduced the level of detected product. However, this might be due to a shift in the +1 TSS to the +4GTP, bypassing the T-tract, which would reduce the level of product labelling in this GC rich transcript. This is supported by the higher intensity bands at 1000 μM and 100 μM GTP when labelling with α-CTP and a shift in product size that was repeatedly observed, with a shorter transcript produced at 100-1000 μM The bands of greater length could also indicate RT that has returned to canonical transcription, however sequencing would be required to confirm this.

Overall, these experiments confirm a decrease in full length transcription upon the reduction of GTP levels, with a simultaneous increase in RT at *rpoBp1 in vitro*, suggesting sensitivity of this promoter to intracellular GTP levels.

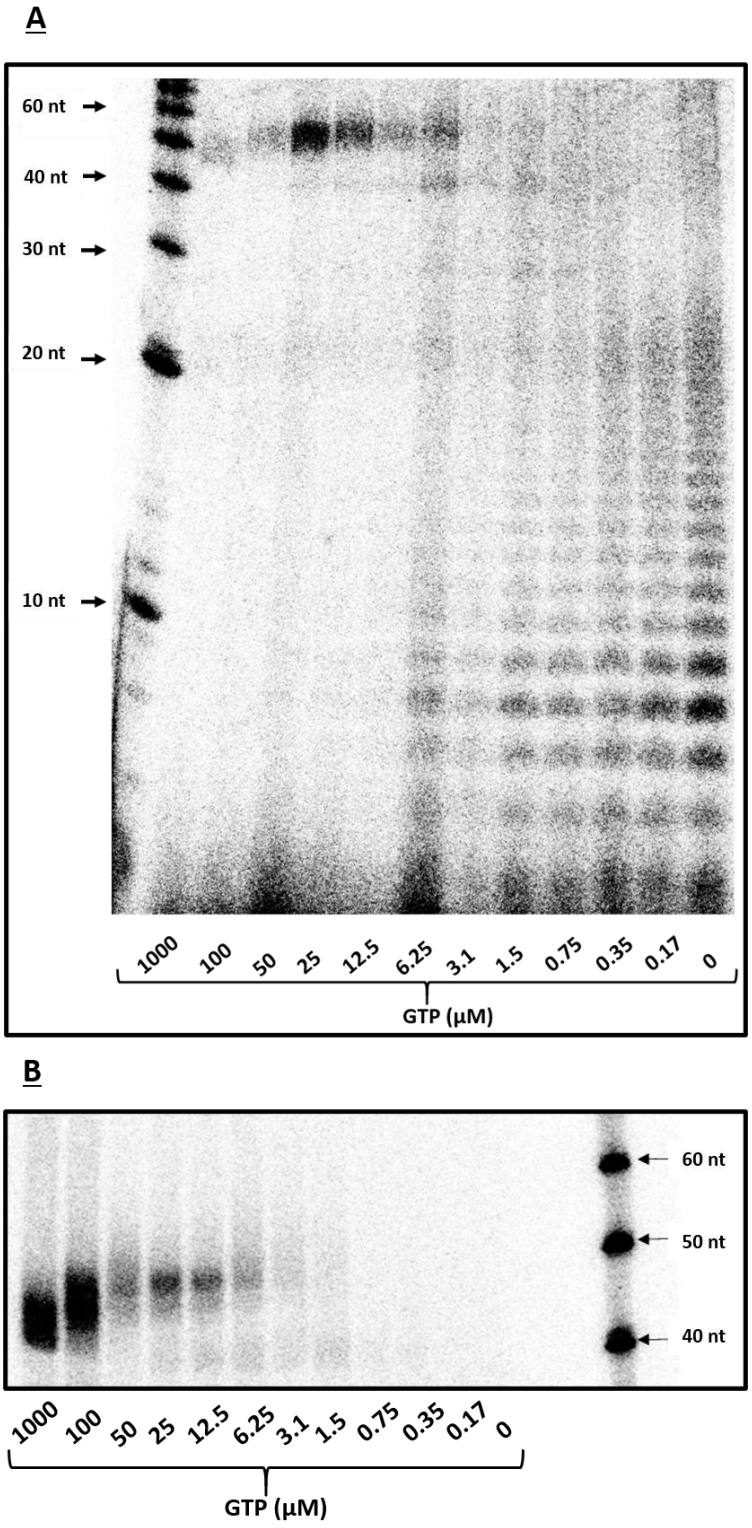


Figure 4.3.5.1: *In vitro* transcription assays showing the effects of GTP depletion on the wildtype *rpoBp1* promoter. GTP depletion were carried out in the presence of (A) α-UTP or (B) α-CTP. Samples from assays were run on 24% and 8% denaturing urea gels at 600 V 50 W and 55 °C for 2 h 45 min and 1 h 20 min in images A and B, respectively. Both gels were visualised using phosphoimaging. RNA markers (FisherSci, AM7778) are indicated on both images. The concentration of GTP in each reaction is specified below each lane.

4.3.6: Transcription from *rpoBp1* is downregulated in depleted GTP levels when T-tract is removed.

Previous experiments confirmed that *rpoBp1* RT at the TSS is sensitive to GTP levels *in vitro* (Fig 4.3.5.1). To determine whether RT at the *rpoBpl* plays a regulatory role by allowing the sensing of [GTP], depletion *in vitro* transcription assays using performed using the 5'ACTG mutant DNA template (Fig 4.3.4.1), where slippage does not occur (Fig 4.3.4.2). In contrast to our hypothesis, a similar response to [GTP] was observed; transcripts were detected to ~3.1 μM GTP but were absent at 1.5 μM GTP (Fig 4.3.6.1), suggesting the homopolymeric TSS plays little role in response to GTP levels *in vitro*.

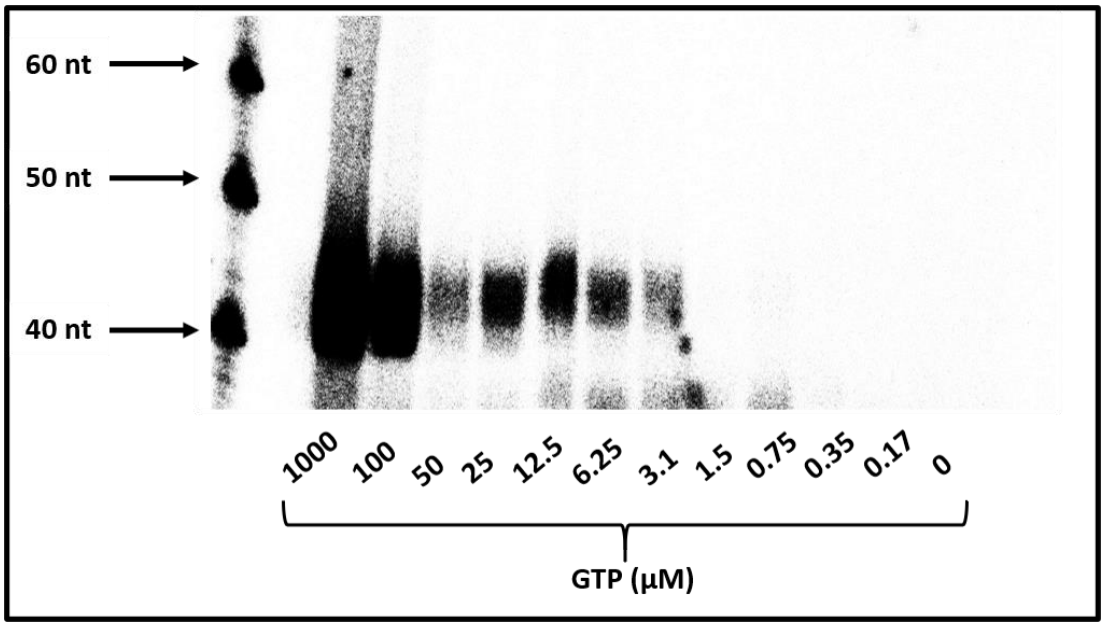


Figure 4.3.6.1: *In vitro* transcription assays showing the effects of GTP depletion on the ACTG mutant *rpoBp1* promoter. The lengths of the RNA markers (FisherSci, AM7778) are indicated. α-CTP labelled samples were run on an 8% denaturing urea gel for 1h 20 min at 600 V 50 W and 55 °C and visualised using phosphoimaging. GTP concentrations (μM) are provided underneath each lane.

4.3.7: Transcription from *rpoBp1* is downregulated in depleted GTP levels when +4 GTP is replaced with CTP

Previous experiments suggested the importance of the +4 nucleotide in determining the extent of RT at *rpoBp1* (Section 4.3.5). However, similar sensitivity to [GTP] was seen for both the wildtype (5'TTTG) and mutant 5'ACTG promoters, suggesting that +4 GTP by itself might be crucial in coordinating transcription initiation with GTP levels, as observed in *B. subtilis* (Figs 4.3.5.1 and 4.3.6.1) (Krásný *et al.*, 2008).

To investigate further, *in vitro* transcription assays were performed using the 5'TTTC template (Fig 4.3.4.1). Surprisingly, the 5'TTTC template displayed no difference in sensitivity to [GTP] compared to WT, suggesting that +4 GTP is not by itself a sensor of GTP levels (Fig 4.3.7.1); as is seen in WT, run-off transcription is largely diminished in 3.1 μM levels of GTP, with complete loss of transcription products at 1.5 μM. The sequence of the band ~10 nt, seen to increase as GTP is depleted is yet to be confirmed, but is likely to be 5' UUUCC, since the first GTP to occur in this template is at +6 (Fig 4.3.4.1), with deviation from the predicted size possibly due to anomalous mobility (Qi *et al.*, 1996).

Overall, these results show that the +4 GTP of the *rpoBp1* region can influence RT but is unlikely to be a special sensor that down-regulates transcription in response to depleted GTP levels *in vitro*. However, these artificial *in vitro* experiments lack key elements of the regulatory region including the UTR, and so further work is required to confirm this.

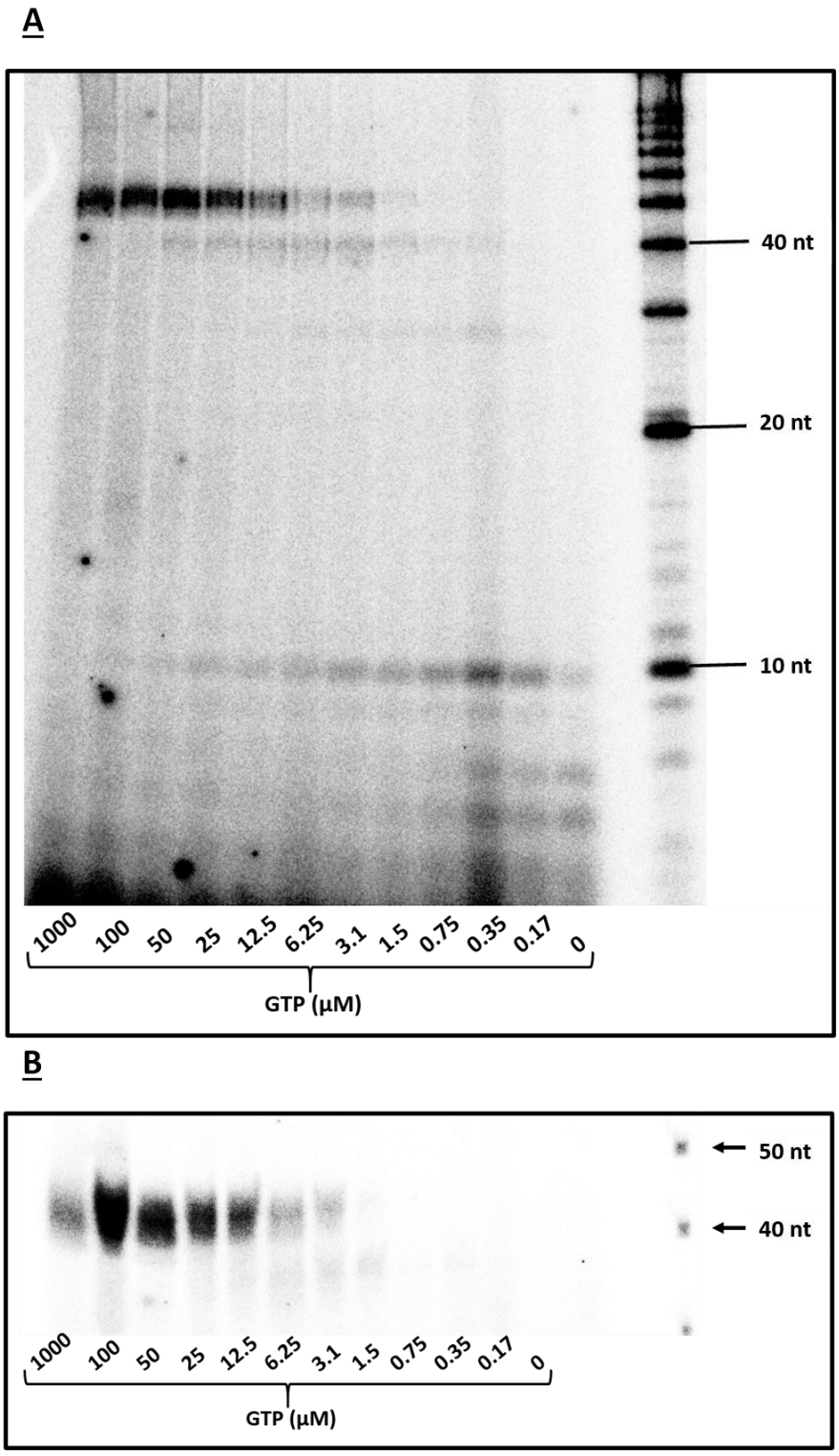


Figure 4.3.7.1: *In vitro* transcription assays showing the effects of GTP depletion on the TTTC mutant *rpoBp1* promoter. A and B represent the assays carried out with the α-UTP and α-CTP radionucleotides respectively. Samples from the assays were run on 24% and 8% denaturing urea gels at 600 V 50 W and 55 °C for 2h 45 min and 1h 20 min in images A and B, respectively, with both gels visualised using phosphoimaging. The lengths of the RNA markers (FisherSci, AM7778) are indicated on both images. The concentration of GTP in each reaction is specified below each lane.

4.3.8 Transcription of *rpoBp1* is downregulated upon GTP depletion in the presence of high UTP concentrations.

While depletion of the nucleotide that occurs after a polymeric tract can cause slippage, an increase in RT can also occur at high concentrations of the nucleotide that is encoded by the polymeric tract (Turnbough, 1992). Since increased slippage at *rpoBp1* might influence the sensitivity of the system to [GTP], the GTP depletion assays used as previously (Section 4.3.5) were repeated at "high"- (1000 μM) or "standard" (100 μM) UTP concentrations (Section 2.1.8.4.2). However, the different concentrations of UTP did not appear to influence the sensitivity to [GTP] (Fig 4.3.8.1).

Notably, at high [GTP] and high [UTP] the proportion of full length transcripts that are thought to initiate with +1T was higher than at standard [UTP] (Fig 4.3.8.1) where it was proposed that transcription initiates at +4GTP (Section 4.3.6.1). It is possible that high [UTP] competes effectively with GTP to ensure initiation on the T-tract, with a lack of observable increase in transcript size with decreasing [GTP] providing evidence for this (Fig 4.3.8.1A). Therefore it is possible that the promoter is particularly sensitive to the ratio of UTP to GTP, although variable start site selection is not unusual (Yu *et al.*, 2017).

For reference the same GTP depletion assays were performed in the presence of low UTP (20 μM UTP) as above, however, bands were only observed at GTP concentrations of 100 μM and 50 μM (data not shown). This provides further evidence that the homopolymeric T-tract is the most abundant TSS for *rpoBp1,* and a certain level of UTP is required for successful transcription initiation. Transcription may have only been observed for the higher concentrations of GTP (100 μM and 50 μM), as sufficient levels are available for initiation at the GTP present after this homopolymeric T-tract (+4G), providing further evidence for start site selectivity at *rpoBp1*.

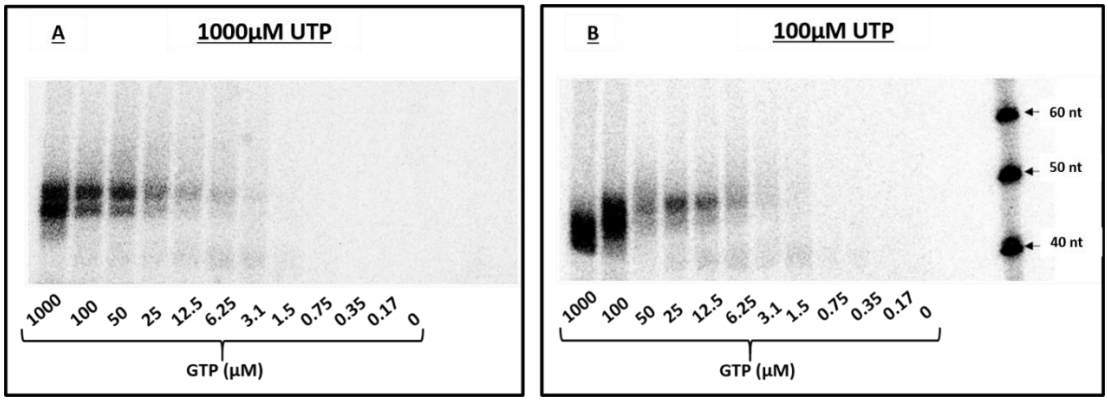


Figure 4.3.8.1: *In vitro* transcription assay showing the effects of GTP depletion on *rpoBp1* in the presence of high UTP nucleotides. All samples (8 µL) were run on 8% denaturing urea gels at 600 V 50 W and 55 °C for 1 h 20 min and imaged using phosphoimaging. The lengths of the RNA markers (FisherSci, AM7778) are indicated on image B. A and B represent the *in vitro* transcription assays carried out with high (1000 µM) and standard (100 µM) UTP concentrations respectively, using α-CTP radionucleotide and wildtype *rpoBp1*(1) DNA template. GTP concentrations used for depletion are stated under each lane.

4.4 *rpoBp1*-dependent transcription of *rpoB* increases in an TTT>ACT mutant

4.4.1 Construction and phenotypic characterisation of a stably integrated mutant pSX400:*rpoBp1-rpoBC* fusion into the genome of *S. coelicolor*

To determine the effects RT has on the expression of *rpoBC*, the pSX400*::rpoBp1_ACT* plasmid, containing the *rpoBp1* promoter (including ~170 bp upstream DNA) and a mutated TSS (TTT > ACT), was constructed and integrated by single crossover homologous recombination into the *S. coelicolor* genome as seen in Section 3.3.2, producing strain S302. Primers 15 and 16 were used to amplify 1.2kb of *rpoB* promoter from the pIJ12738*::rpoB_ACT* plasmid, introducing the ACT mutation into the construct. Integration of plasmids into the genome of *S. coelicolor* was confirmed using PCR as previously (Section 3.3.2).

Due to the integration of the non-replicating pSX400 vector, which simultaneously removes promoters *rpoBp2* and *rpoBp3*, the expression of *rpoBC* in S302 and S301 (created previously, Section 3.3.2) may differ to the M145 parent strain, and present phenotypic differences. A single colony from each of the three strains was streaked onto MS, R2, minimal medium and Bennett's agar for study of phenotype and antibiotic production (Section 2.1.7.2). MS plates containing thiostrepton (10 µg/ml)

to select for pSX400 were used to further confirm integration and test for general growth and sporulation (Fig 4.4.1.1).

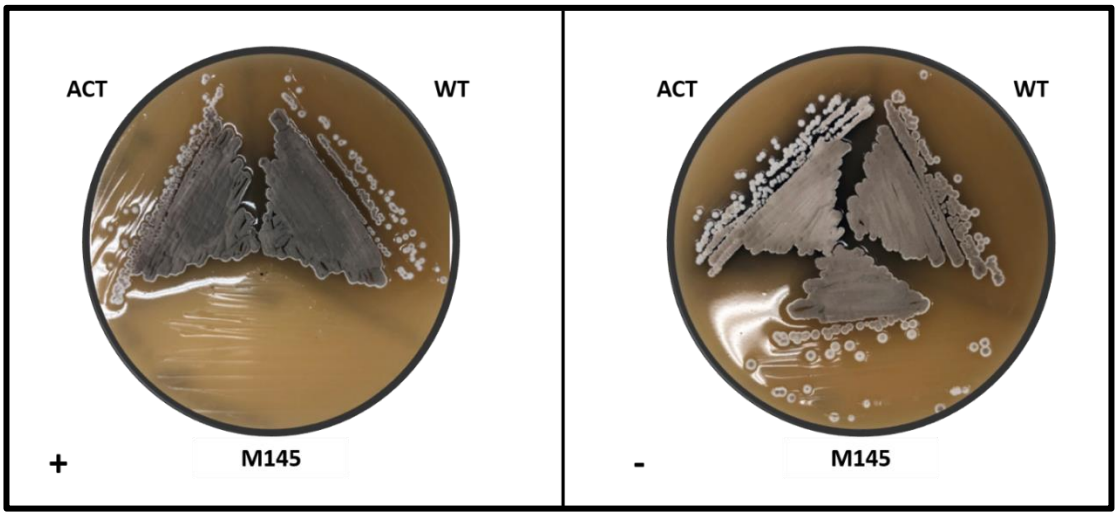


Figure 4.4.1.1: Sensitivity of M145 and S301 and S302 to thiostrepton. M145, S301 (WT) and S302 (ACT) were streaked onto MS agar both with and without 10 µg/ml thiostrepton and grown for 4 days at 30°C. (+) and (-) represent agar containing and lacking thiostrepton, respectively.

Phenotypic differences were not apparent on the minimal media between the parent M145 and the strains containing the integrated plasmids, however, like MS agar, a slight decrease in sporulation was observed on Bennetts agar for S301 and S302 (Fig 4.4.1.2). The S302 (ACT) strain appeared to produce more actinorhodin on MS agar containing thiostrepton when compared to the wildtype S301 (TTT) strain. However, the effect on antibiotic production is not consistent as the parent M145 was seen to produce slightly more of this antibiotic on R2 medium than the other two strains. However, overall, the fusion of the *rpoBC* operon to *rpoBp1* (TTT) or *rpoBp1*(ACT) by single crossover integration appeared to have little effect on the level of growth and pigmented antibiotics production compared to the M145 parent strain.

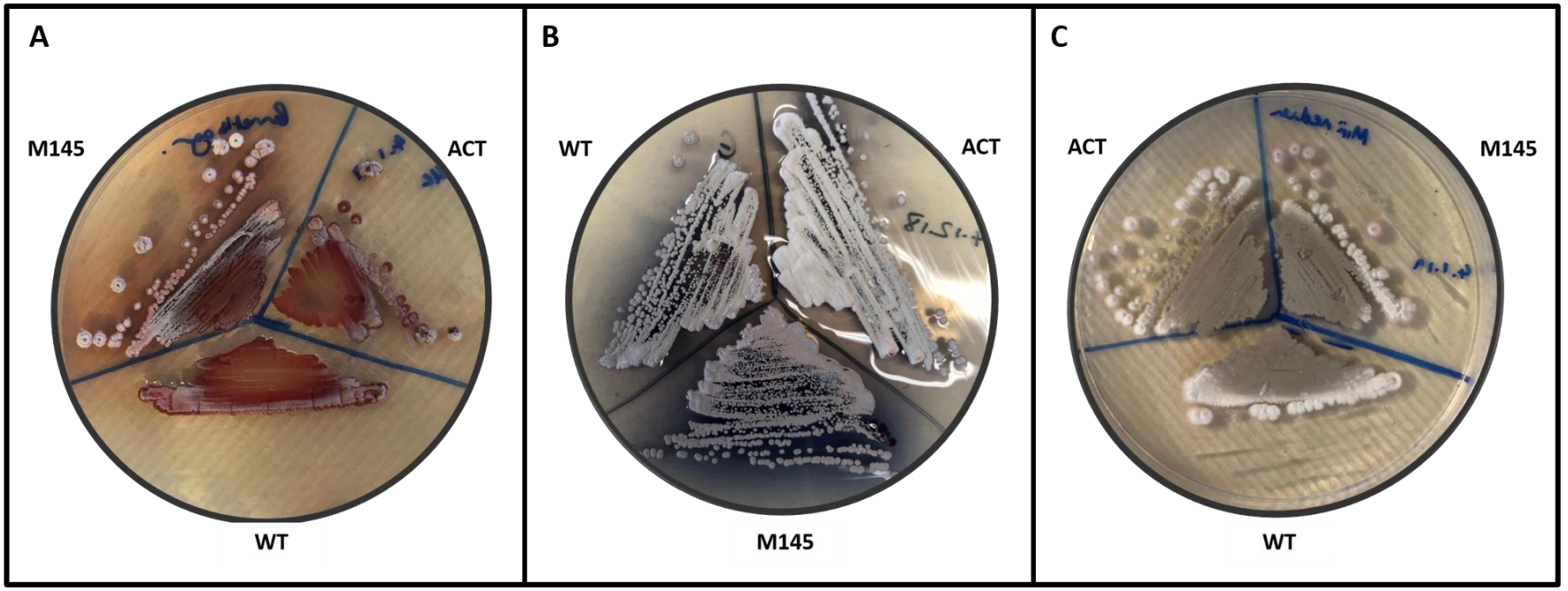


Figure 4.4.1.2: Growth of M145, S301 and S302 on R2, minimal medium and Bennetts agar. M145, S301 (WT) and S302 (ACT) were streaked onto the agar with no selection; A, B and C represent colonies on Bennetts, R2 and minimal medium agar, respectively. Strains were grown on at 30 °C for 4 days.

4.4.2 *rpoB* expression is increased in the mutant S302 (ACT) strain

To determine the effects of the ACT mutation on *rpoB* expression *in vivo*, M145 and both S301 (TSS) and S302 (ACT) were grown to exponential phase (OD450 of 0.8-1.0) in NMMP media (Sections 2.1.7.2 and 2.2.1.2), maintaining thiostrepton selection, and RNA purified (Section 2.2.3.3) and processed for cDNA synthesis and qRT-PCR (Section 2.2.5.1).

RpoB expression was the highest for M145, with both plasmid integrated strains having a reduced level of expression in comparison; a ~1.5 and ~9.5 fold decrease in expression in strains S302 and S301, respectively (Fig 4.4.2.1). This decrease is likely to be due to the absence of the upstream promoters, *rpoBp2* and *rpoBp3* in the constructed single-crossover fusion.

In comparison to S301 (TTT), the S302 (ACT) mutation exhibited a ~6.3 fold increase in *rpoB* transcript levels, providing evidence that RT negatively impacts transcription from *rpoBp1*. However, the replacement of +1UTP with +1ATP may contribute to increased expression, as ATP is a more favourable initiating nucleotide (Turnbough, 2008), with further mutagenesis required to dissect this.

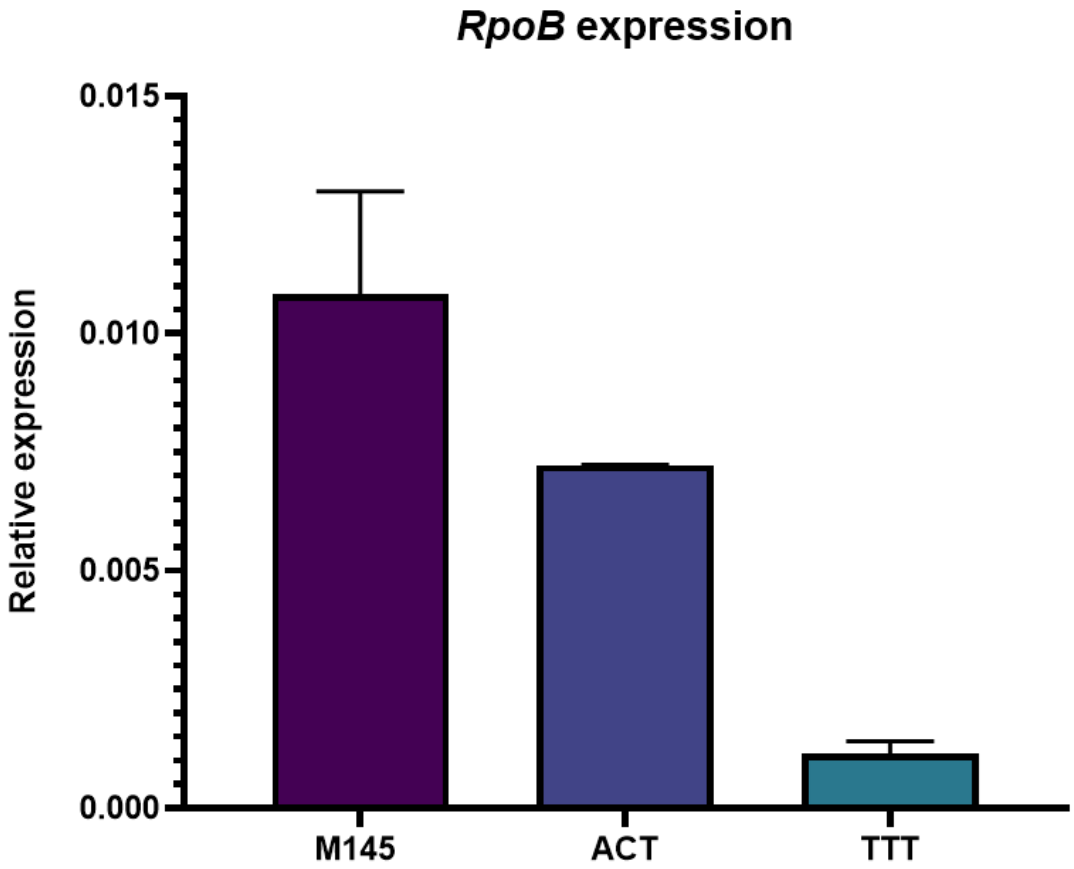


Figure 4.4.2.1: Basal *rpoB* expression in M145, S301 and S302. Strains were grown in biological triplicate to an OD450 of ~0.8-1.0, with *rpoB* expression determined using qRT-PCR and normalised to stable 16s rRNA, using primers 71 and 72, and 67 and 68, respectively (Table 2.1.5.3). Error bars present the S.E.M of each strain. TTT and ACT indicate the S301 and S302 strains, respectively.

4.4.3 Construction of a stable mutation of *rpoBp1* TSS using a CRISPR-cas based system

Previous construction of an *in vivo* mutation using the integrated *pSX400::rpoBp1_ACT* also resulted in additional plasmid DNA being incorporated into the *S. coelicolor* genome, including a duplication of the regulatory region, which have unknown effects on *rpoB* expression (Section 4.4.2). The *Streptococcus pyogenes* CRISPR/Cas system has become a well-known genome editing system widely used for the generation of specific mutations in a variety of organisms. A CRISPR-Cas based system has been optimised for use in *Streptomyces*, and was used to construct a stable chromosomal *rpoBp1* TTT>ACT mutation following recommended guidelines (Cobb and colleagues (2015).

For correct use of this system, both a spacer sequence, containing the 5' NGG 3' PAM site required for cas9-specific cutting, and a sequence for HR rescue of the lethal double stranded break (dsb), are required for site directed mutagenesis and cloning into the pCRISPOmyces plasmid (Cobb, Wang and Zhao, 2015). The spacer sequence (Sp1) was created by annealing oligonucleotides 24 and 25 (Section 2.2.2.8) before ligated into a BbsI-digested pCRISPomyces 2.0 plasmid, producing the pCRISPO-_Sp1 plasmid. The HR template was made using PCR and primers 26 and 28, and 27 and 29, and *S. coelicolor* genomic DNA amplifying 1kb of DNA either side of the *rpoBp1*, whilst simultaneously introducing an ACT mutation at the TSS (Section 2.2.2.3). Primers 26 and 27 were also used alone to produce a template with the wildtype T-tract still in place (Table 2.1.5.1). These were then PCR purified and cloned into the pCRISPO_sp1 plasmid using Gibson assembly (Sections 2.2.2.3 and 2.2.2.6) producing plasmids pCRISPO_Sp1_ACT and pCRISPO_Sp1_WT containing the mutated and wildtype T-tract TSS, respectively. Note that these HR templates were also cloned into the empty pCRISPOmyces plasmid alone to act as a control; pCRISPO_HR_WT and pCRISPO_HR_ACT plasmids contain the 2kb HR templates for wildtype and ACT mutated *rpoBp* regions, respectively, (with no presence of the spacer sequence), and were used to confirm a lack of background effects on *rpoB* expression.

To prevent further cutting by cas9 of the *rpoBp* region, the last GTP of the PAM sequence, (5' NGG 3') was also mutated in the HR templated provided. To reduce

likely effects of this secondary PAM mutation on *rpoB* expression, the alignments of *rpoBp* seen in Section 3.2.2 were used to identify a NGG site containing an unconserved GTP within differing *Streptomyces* strains. Potential spacer sequences were also BLAST against the *S. coelicolor* genome to reduce the likelihood of off-site effects. The final NGG site selected for the spacer sequence lies 47 bp away from the T-tract being mutated. The pCRISPO_Sp1_ACT and pCRISPO_Sp1_WT plasmids were further taken forward for mutagenesis of the spacer 1 PAM site within the HR template using PCR and primers 27 and 30 and 26 and 31 (Table 2.1.5.1) before purification and assembly as stated above, producing the pCRISPO_WT_Sp1* and pCRISPO_ACT_Sp1* plasmids, respectively.

All mutations were confirmed with sequencing and initial cloning of all constructs was carried out using Dh5α for transformations and miniprep (Sections 2.2.4.3 and 2.2.3.1). Once confirmed, all plasmids were transformed into *E. coli* ETZ before conjugation into M145 (Section 2.2.4.5), selecting for apramycin resistance. Control conjugations using plasmids containing the spacer sequence only (pCRISPO_Sp1), and templates without PAM mutations (pCRISPO_Sp1_ACT, pCRISPO_Sp1_WT) provided evidence for cas9 activity due to a lack of colonies, expected from the lethal dsb. M145 colonies containing the final test plasmid (pCRISPO_ACT_Sp1*) containing both the PAM and ACT mutation were taken forward for colony PCR using cloning primers 26 and 27, and sequenced using both these and primer no 32, internal to *rpoB,* to confirm a successful recombination event and induction of this ACT (and PAM) mutation. However, upon sequencing of strains only the ACT mutation was confirmed, with the PAM site used for cutting left unmutated. This may indicate this specific GTP mutation supposedly required to prevent further cutting with cas9 was lethal to the strain. This however is more favourable and indicates any change in *rpoB* expression is resultant from this stably integrated TTT>ACT mutation that was introduced. This strain was then further cultured at a high growth temperature of 37°C to facilitate the removal of the temperature sensitive plasmid, and taken forward for replica plating (Section 2.2.10). A strain conferring renewed sensitivity to apramycin was again sequenced, confirming this stable TTT>ACT mutation (M145_*rpoB*_ACT) and was taken forward for further determination of the effects on *rpoB* activity and phenotype.

The pCRISPO_WT_Sp1* exconjugants were initially constructed to confirm any effects this NG**G** (PAM) mutation may have on *rpoB* expression, however due to the lack of its presence in the final strain containing the ACT mutant, this plasmid/these exconjugants were not taken forward for testing of *rpoB* expression.

4.4.4 The mutagenesis of *rpoBp1* TSS has little effect on growth and phenotype

The removal of potential slippage events themselves may have an effect on not only the expression of *rpoB*, but the level of β-subunit and thus level of RNAP itself. An altered level of RNAP may affect growth rate, which was determined by comparing growth of mutant M145_*rpoB*_ACT strain to the parent M145 strain.

Both M145 and M145_*rpoB*_ACT were pre-germinated for ~3 h and grown in biological duplicate in YEME (10% sucrose), at a starting OD450 of 0.06 (Section 2.2.1.2). Initially strains were grown at 30°C for 13.5 h before readings were taken every 2 h. Figure 4.4.4.1 displays the growth of both strains over the course of several hours. All replicates of both M145 and ACT mutant strains were seen to grow at a similar rate, providing evidence that the introduced mutation has little to no effect on growth rate of *S. coelicolor*.

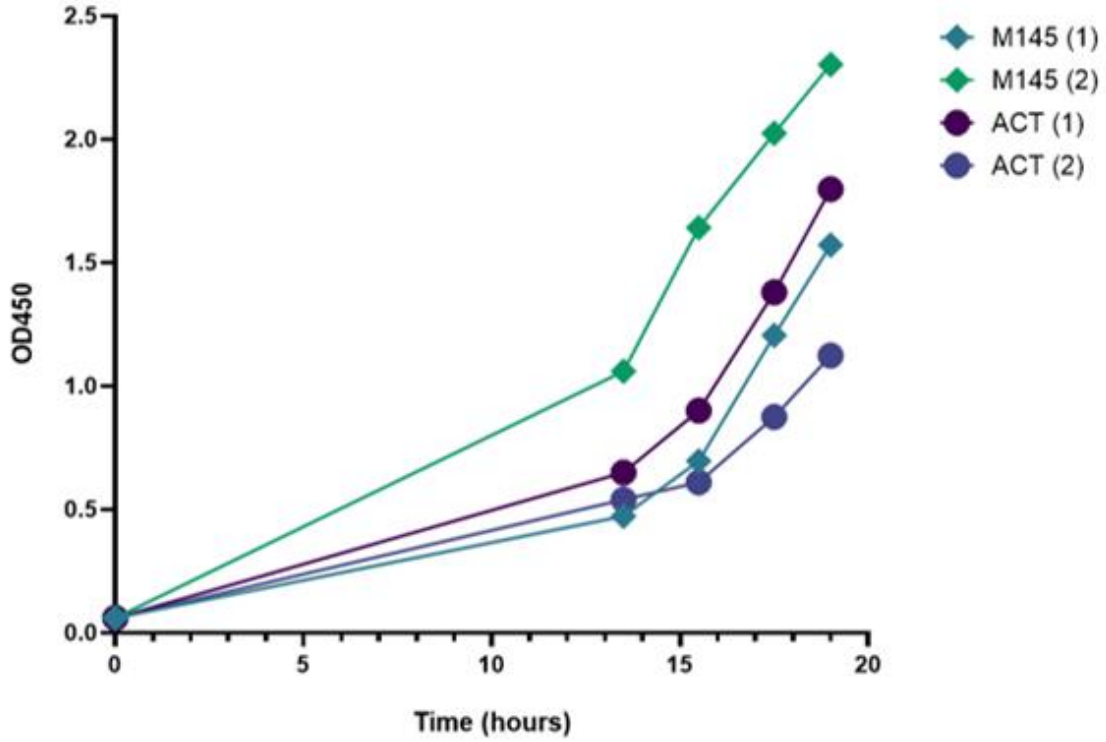


Figure 4.4.4.1: Mutagenesis of the TSS for *rpoBp1* has no effect on overall growth. Both M145 (squares) and M145_*rpoB*_ACT (circles) were grown over a course of several hours in YEME (10%) from a starting OD450 of 0.06.

The effect of the TTT-ACT mutation on the growth on solid MS, SMMS, R5 and nutrient agar and was also determined (Figs 4.4.4.2 and 4.4.4.3). Colonies exhibited similar growth after 2 days on MS media, however the mutant produced more blue pigmented actinorhodin (Fig 4.4.4.2).

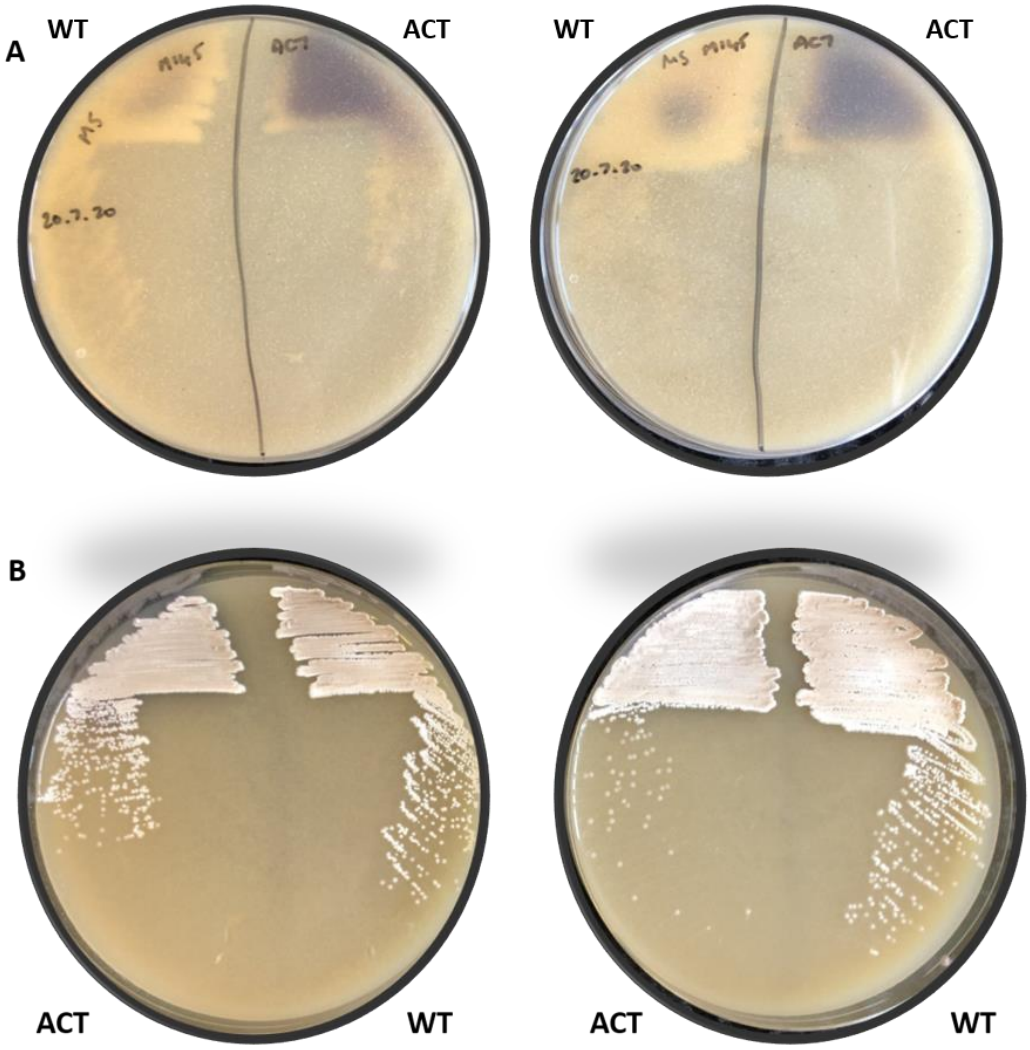


Figure 4.4.4.2: Growth of M145 and mutant M145_*rpoBp*_ACT strains on MS agar. Blue pigment signifies production of actinorhodin antibiotic. ACT and WT (TTT) signifies the M145_*rpoB*_ACT and M145 strains, respectively. Both colonies were grown for 2 days at 30°C before images were taken of both colony surface (B) and underside of plates (A) of growth in duplicate.

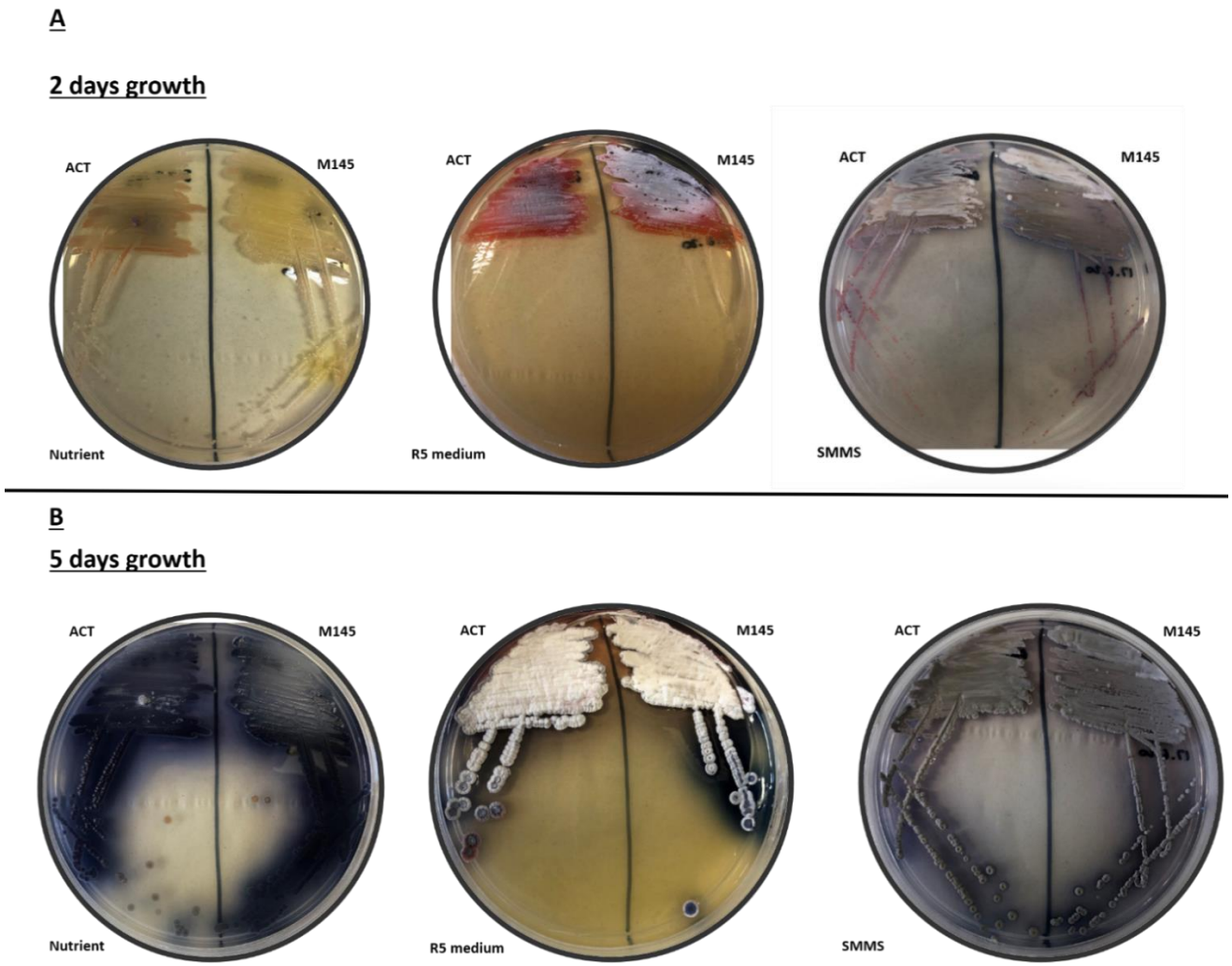


Figure 4.4.4.3: M145_*rpoBp*_ACT and M145 on nutrient, R5 and SMMS agar. Strains were all grown for either 2 days (A) or 5 days (B) at 30°C before images were taken. ACT and M145 specifies the M145_*rpoB*_ACT and M145 strains, respectively. Images of colonies on nutrient, R5 and SMMS agar are shown from left to right.

It was considered that any change in the production of RNAP might influence sensitivity to the antibiotic rifampicin, which targets the β-subunit of the enzyme (Campbell *et al.*, 2001). Therefore, a spore dilution series was tested for growth on nutrient and SMMS agar containing various levels of rifampicin (0 - 20 μg/ml). However, little difference in sensitivity to rifampicin was observed between the two strains on both SMMS and nutrient agar (Figs 4.4.4.4 and 4.4.4.5).

Overall, the TTT>ACT mutation appears to cause little or no change in overall colony morphology and growth.

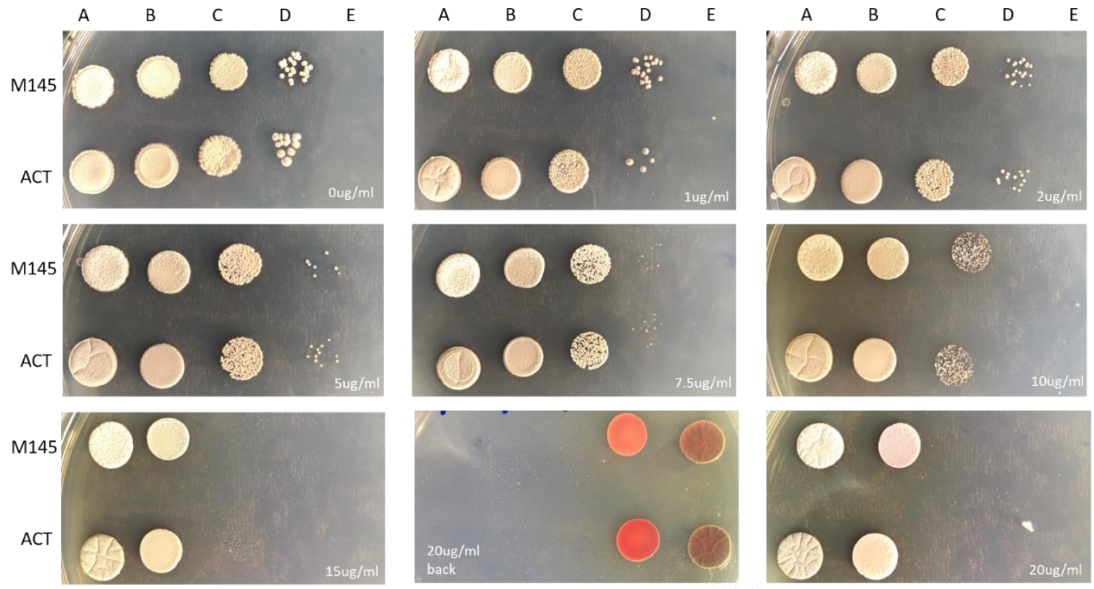


Figure 4.4.4.4: M145 and M145_*RpoBp*_ACT show similar sensitivity to rifampicin on SMMS agar. Both strains were standardised to the same OD450, serially diluted and plated onto SMMS plates containing differing rifampicin concentrations of 0, 1, 2, 5, 7.5, 10, 15, and 20 ug/ml (concentration is noted on bottom right of each plate image). A, B, C, D and E represent neat standardised stock and 10^2, 10^4, 10^6, 10^8 serial dilutions of both strains. Note all images are of the colony surface unless stated; 2 images were provided for the 20 μg/ml plate with one representing the back of the plate to display RED antibiotic production. Note this back plate is reversed compared to the A-E dilutions stated above.

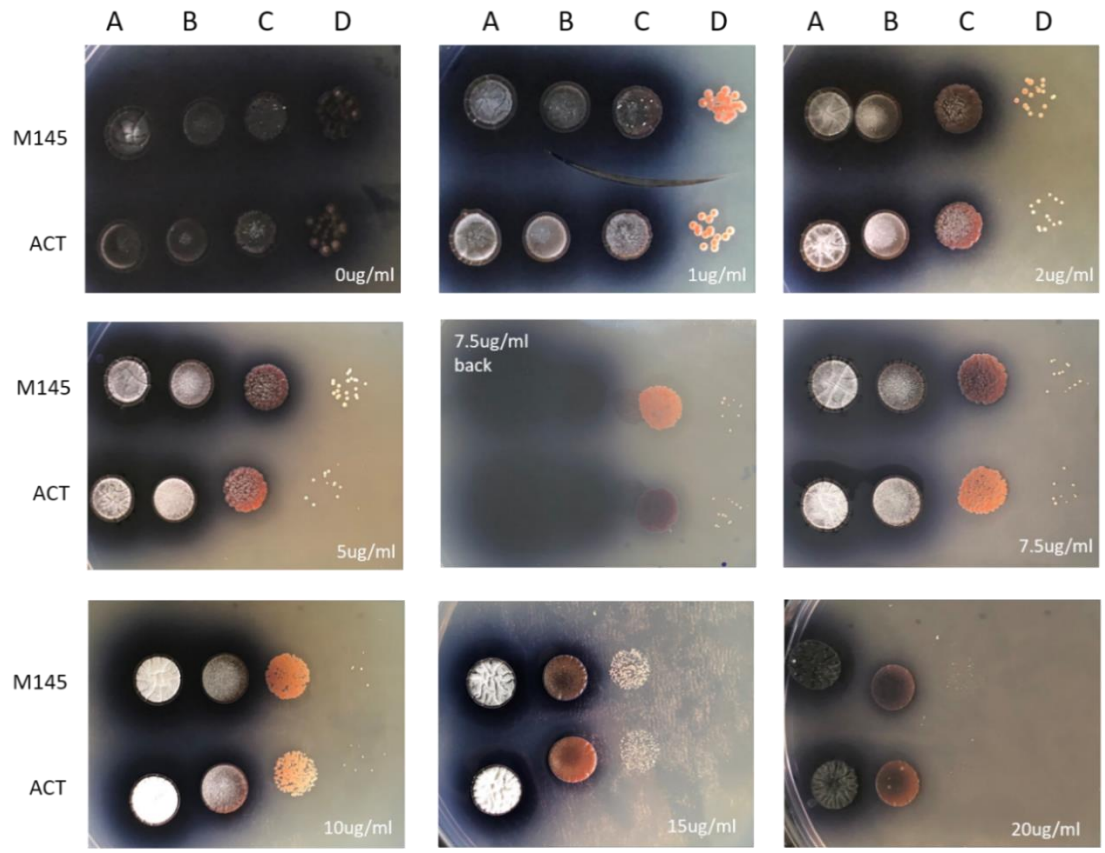


Figure 4.4.4.5: M145 and M145_*RpoBp*_ACT show similar sensitivity to rifampicin on nutrient agar. Both strains were standardised to the same OD450, serially diluted and were plated onto nutrient agar plates containing differing rifampicin concentrations of 0, 1, 2, 5, 7.5, 10, 15, and 20 ug/ml (concentration is noted on bottom right of each plate image). A, B, C and D represent neat standardised stock and 10^2, 10^4, 10^6 serial dilutions respectively. Note all images are of the colony surface unless stated; 2 images were provided for the 7.5ug/ml plate with one representing the back of the plate to display RED antibiotic production. Note this back plate is flipped with M145 and ACT strains swapped in respect to those stated at the side of the images.

4.4.5 A stable *rpoBp1* TTT>ACT mutation causes increased *rpoB* expression *in vivo*

To determine effects of the TTT>ACT mutation on basal *rpoB* expression, transcript levels in M145_*rpoB*_ACT, were compared with two strains of interest, the parent M145 and M571 (RelAΔ), with the latter used in subsequent stringent response investigations. All three strains were inoculated in biological triplicate into NMMP medium at a starting OD450 of 0.06 using mycelial stocks (Sections 2.1.7.2 and 2.2.1.2) and grown to early exponential phase (OD450 ~0.8-1.0), before 8 ml of each culture was extracted for RNA purification and subsequent cDNA synthesis (Sections 2.2.3.3 and 2.2.5.1).

In a similar way to the pSX400 single crossover integration experiments, *rpoB* expression was significantly increased 1.7-2-fold in the ACT mutant when compared to both M145 parent strain and the Δ*relA* strain, M571 (Fig 4.4.5.1); again this suggests that slippage at the *rpoBp1* TSS has a detrimental effect on *rpoB* expression, while noting that the changes might be due to a preferred +1ATP (Turnbough, 2008).

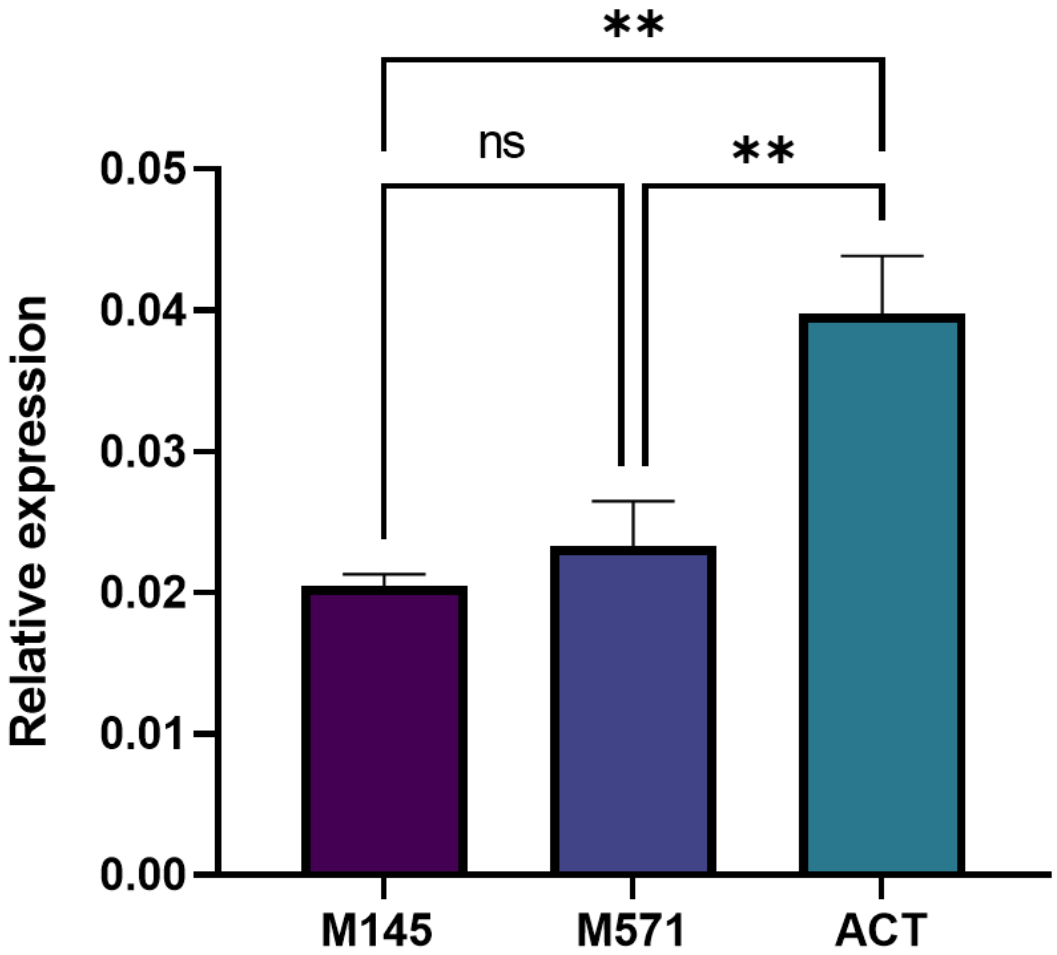


Figure 4.4.5.1: Basal *rpoB* expression of M145, M571 and mutant M145_*rpoB*_ACT strain. Strains were grown to an OD450 of ~0.8-1.0 in NMMP medium, with *rpoB* expression determined and normalised to stable 16s rRNA using primer pairs 71 and 72 and 67 and 68, respectively (Table 2.1.5.3). cDNA copy number was determined using a standard curve of known gDNA copies and error bars presented represent the standard error of the mean (S.E.M) of data in at least biological triplicate. Statistical significance was determined using an ordinary one-way ANOVA where equal distribution between data and standard deviations is assumed ($P < 0.01$ = **).

4.4.6 *rpoB* is downregulated in response to stringent conditions in the M145_*rpoB*_ACT mutant

Previously *rpoB* was confirmed to be under the stringent control in M145 (Section 3.3.4). To investigate a potential role of RT in the stringent control of *rpoBC*, M145_*rpoB*_ACT was subjected to nutrient downshift and the effects on *rpoB* expression quantified.

The M145_*rpoB*_ACT strain was taken forward for the induction of stringent response and RNA extraction as seen in Section 2.2.1.4. cDNA synthesis was carried out as seen in Chapter 3.3.4 and for reference the M145 and M571 results from this chapter, were included.

As was seen in M145, following nutrient down-shift, a steady decrease in *rpoB* expression over 30 min was seen in M145_*rpoB*_ACT, although individual values did not reach the threshold set for statistical significance (Fig 4.4.6.1). It should be noted that all individual biological replicates for both M145 and the mutant ACT strain showed a gradual decrease in expression to around 30 min (data not shown). No downregulation of *rpoB* occurred in M571, confirming changes in expression are due to the functional RelA protein present in the M145 and ACT mutant strains. Overall, the data suggest that RT is not required for the stringent control of the *rpoBC* operon.

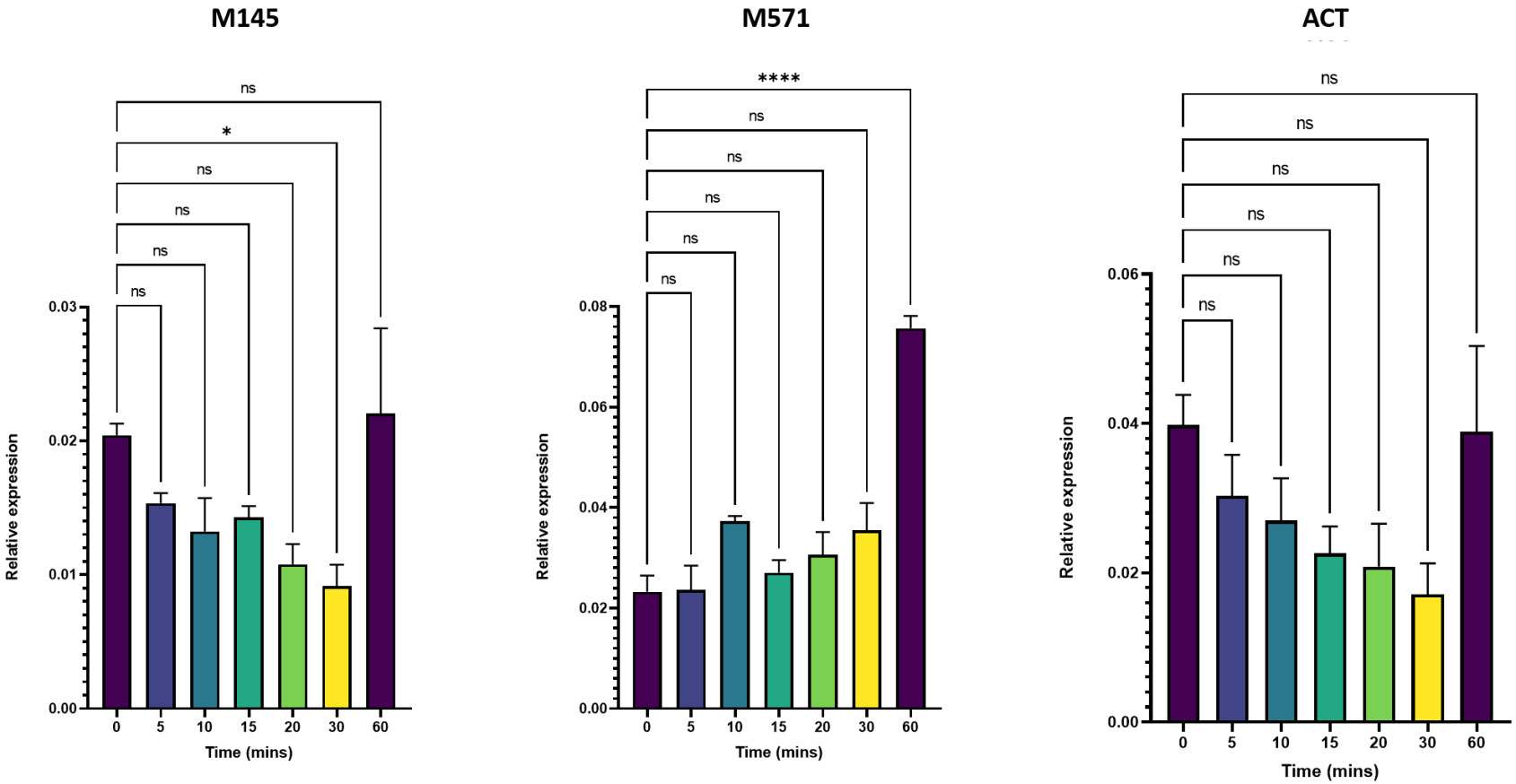


Figure 4.4.6.1: *rpoB* expression during stringent conditions in M145, M571 and the mutant M145_*rpoB*_ACT strain. The expression of *rpoB* was determined using qRT-PCR and normalised to stable 16s rRNA using primer pairs 71 and 72 and 67 and 68, respectively (Table 2.1.5.3). cDNA copy number was determined using a standard curve of known gDNA copies. Error bars shown represent the S.E.M of the data for at least biological triplicate. Statistical significance was calculated using an ordinary one-way ANOVA, where equal distribution between data and standard deviations is assumed. (P< 0.05 = *, P < 0.0001 = ****).

4.4.7 *rpoB* is downregulated in response to oxidative stress conditions in the M145_*rpoB*_ACT mutant

The induction of this mutation omitting RT events at the p1 TSS for the *rpoBC* operon was shown to have little effect on the response to stringent conditions (Section 4.4.8). During oxidative stress however, the levels of ppGpp did not significantly increase upon the addition of the oxidising reagent, diamide, hence global changes in transcription are likely elicted using differing mechanisms to those apparent in the stringent response (Paget and Hesketh, unpublished; Kallifidas *et al*., 2010).

As shown in Section 3.3.3, the *rpoBp1* promoter is downregulated in response to diamide-induced oxidative stress and therefore is likely to be responsible for the overall down regulation of the *rpoBC* operon (Kallifidas *et al.,* 2010). To investigate whether the *rpoBp1* T-tract plays a role in this response, M145_*rpoB*_ACT was subjected to this stress with *rpoB* expression quantified using qRT-PCR. Disulphide stress was induced as previously (Sections 2.2.1.3 and 3.3.3), and RNA extracted and analysed for *rpoB* expression, normalised to stable 16s rRNA.

In comparison to M145, where *rpoB* is down-regulated ~2.4 fold, diamide oxidative stress caused an even greater and highly significant 8.8-fold decrease in expression in M145_*rpoB*_ACT after 5 min (Fig 4.4.7.1). In both strains, *rpoB* expression returned to pre-stress levels 60 min after the addition of diamide (Fig 4.4.7.1), consistent with microarray data and the fact that diamide is fully consumed within 1 h (Kallifidas *et al.,* 2010). These data indicate that the *rpoBp1* T-tract is not required for down-regulation of *rpoBp1* during oxidative stress.

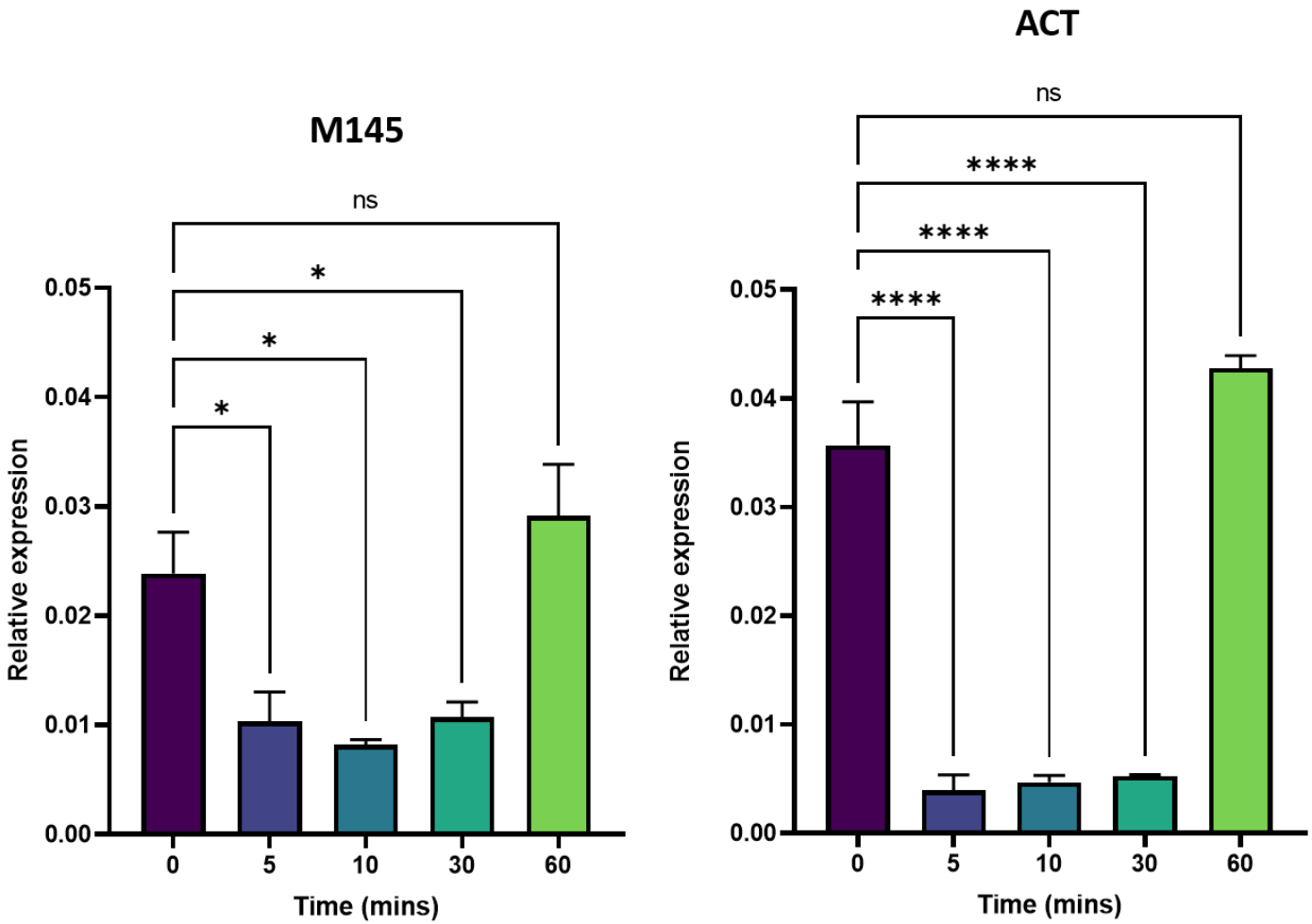


Figure 4.4.7.1: *rpoB* expression during diamide-induced oxidative stress in M145 and mutant strain M145_*rpoB*_ACT. Strains were grown in biological triplicate before treatment with 0.5 mM diamide. qRT-PCR was used to determine *rpoB* expression which was normalised to stable 16s rRNA and quantified using primer pairs 71 and 72 and 67 and 68, respectively (Table 2.1.5.3). Changes in *rpoB* expression were determined by comparison to time zero readings (before stress) using statistical analysis (S.E.M) and an ordinary one-way ANOVA. Statistical significance = $p < 0.05$ = *, $p < 0.0001$ = ****.

4.4.8 Construction of mutant pIJ5972*::rpoBp* fusions

The mutagenesis of the *rpoBp1* TSS site was shown to increase expression of *rpoB* during exponential growth in the M145_*rpoB*_ACT strain, when compared to the unmutated M145 strain. However, as this mutation was stably integrated into the *S. coelicolor* genome, *rpoB* expression may also be subjected to additional control by upstream elements, including the two upstream promoters (*rpoBp2* and *rpoBp3*) characterised previously (Section 3.2.4); indeed, these promoters might mask *rpoBp1* regulatory outcomes. Therefore, to investigate the potential role of the *rpoBp2* and *rpoBp3* promoters, the TTT>ACT mutation was introduced to the differing *rpoBp* constructs used previously (Fig 3.2.3.1, Section 3.2.3). The ACT mutations were introduced into the *rpoBp1*, *rpoBp1U*, *rpoBp1-2U* and *rpoBp1-3U* constructs using the pIJ12738*::rpoB_ACT* plasmid and primer pairs 1 and 4, 1 and 2, 2 and 5, and 2 and 6, respectively. These constructs were then cloned into pIJ5972 and conjugated into *S. coelicolor* M145 (Section 3.2.3).

4.4.9 Mutation of *rpoBp1* TSS results in increased expression at the p1 promoter

Luciferase assays were used to determine expression in both mutant and wildtype *rpoBp* promoter fusions of differing lengths. Strains were grown in biological triplicate to early exponential phase (OD450 of 0.8-1.0) before the luciferase assay was performed (Section 2.2.9; Figs 4.4.9.1.and 4.4.9.2). The ACT mutant data was compared with the previously reported wildtype data (Section 3.2.4) and revealed firstly that the ACT mutation slightly increased expression in constructs containing *rpoBp1* promoter alone (*rpoBp1* and *rpoBp1U*) (Fig 4.4.9.1) consistent with earlier qRT-PCR data (Section 4.4.5). Strikingly however, while the additional presence of the UTR causes a drop in expression in the wildtype promoter (Section 3.2.3, Fig 3.2.3.2) it has no effect on the ACT mutant, with *rpoBp1* activity remaining high. This itself suggests slippage events occurring at the *rpoBp1* TSS are interacting with a potential cis-acting elements present in this 5'UTR.

The effects of this ACT mutation on expression when additional promoters/ upstream DNA are included was also determined (Fig 4.4.9.2), with the previously displayed

rpoBp1U data included for reference. The ACT mutant TSS significantly increased expression in all constructs compared to equivalent fusions containing wildtype *rpoBp1*, again emphasising the negative effect of the T-tract or the positive effect of an +1ATP.

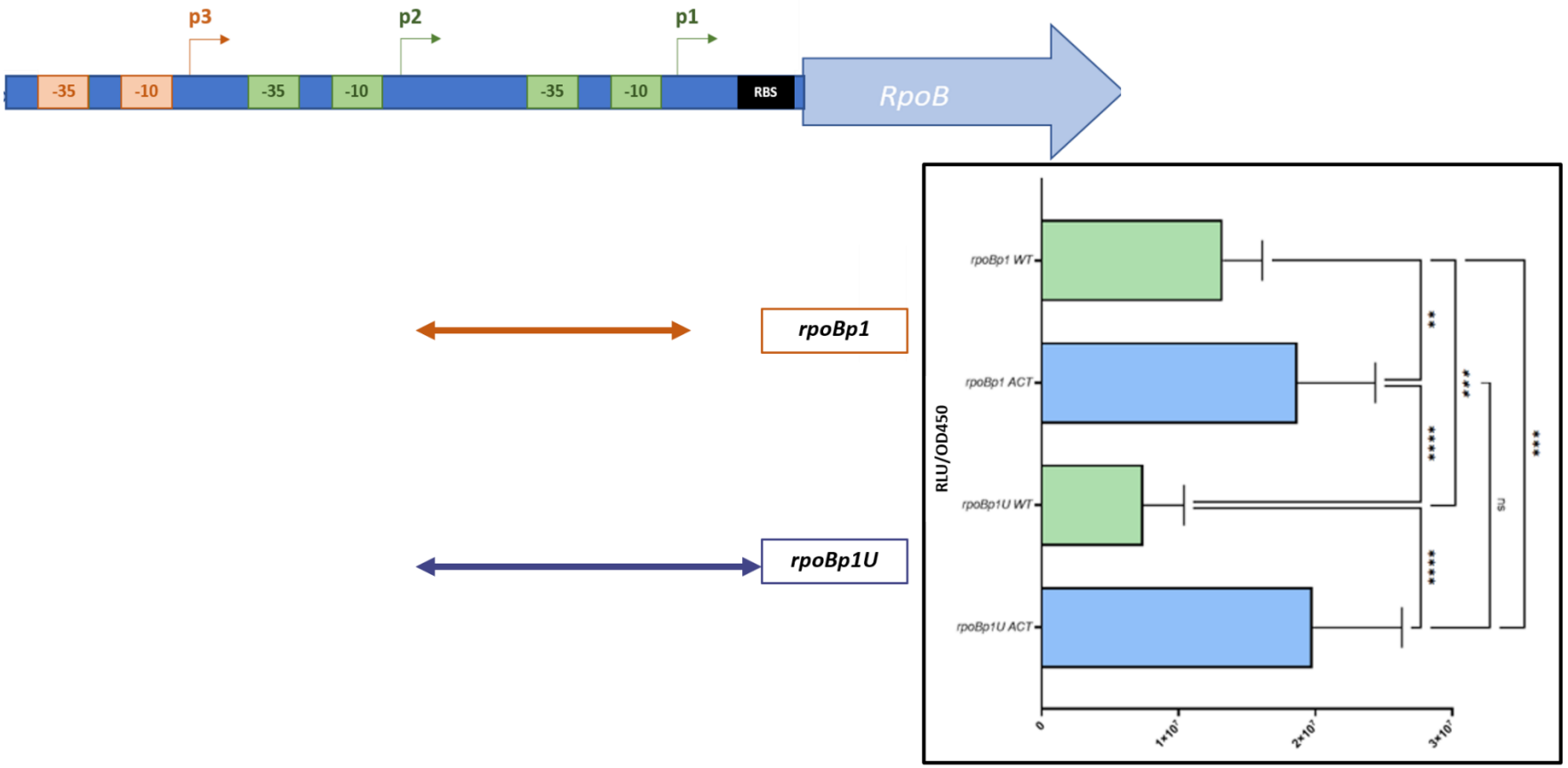


Figure 4.4.9.1: Luciferase activity of the mutant and wildtype *rpoBp1U* and *rpoBp1* truncations during exponential growth. Strains were grown in biological triplicate to early exponential phase (OD450 of 0.8-1.0) in NMMP medium before samples were taken in technical triplicate and luciferase activity (RLU) was determined and normalised to optical density. WT (green) and ACT (blue) represent the wildtype and mutant ACT promoter regions respectively. Error bars represent the S.E.M. Significance values were determined using a Brown-Forsythe and Welch ANOVA; significance values: $P < 0.05$ = *, $P < 0.01$ = **, $P < 0.0005$ = ***, $P < 0.0001$ = ****).

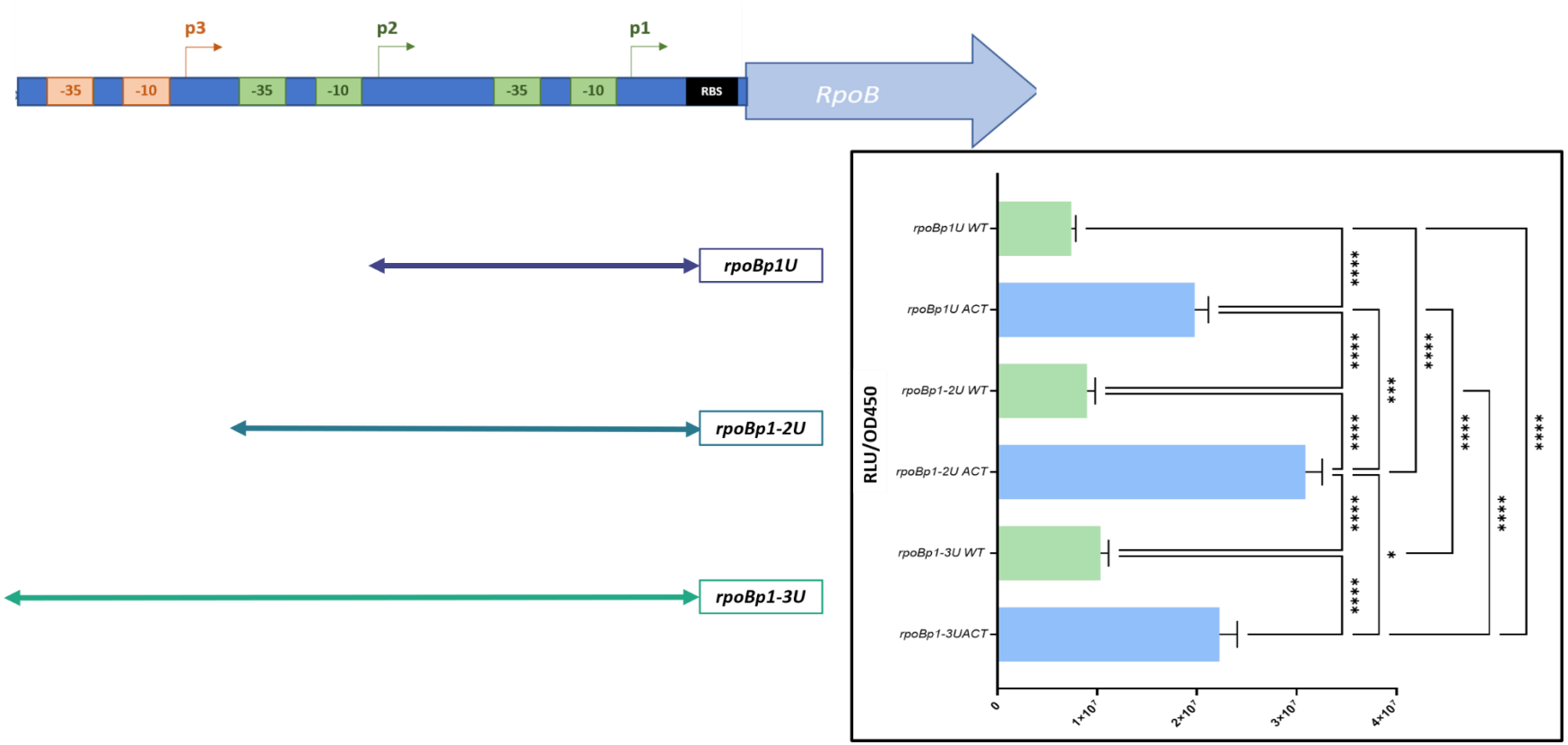


Figure 4.4.9.2: Luciferase activity of the mutant and wildtype promoter truncations containing the 5'UTR region during exponential growth. Strains were grown in biological triplicate to early exponential phase (OD450 of 0.8-1.0) before samples were taken in technical triplicate with luciferase activity (RLU) determined and normalised to optical density. WT (green) and ACT (blue) represent the wildtype and mutant ACT promoter regions. Error bars represent the S.E.M. Significance values were determined using a Brown-Forsythe and Welch ANOVA; significance values: $P < 0.05$ = *, $P < 0.0005$ = ***, $P < 0.0001$ = ****).

4.4.10 Analysis of *rpoB* expression in mutant pSNypet*:: RpoBp* reporter fusions

Due to the observed stability of the luciferase enzyme (Appendix Section 7.3), it was decided to investigate the effect of the ACT mutation using the more dynamic *ypet** reporter plasmid pSNypet (see Section 3.2.4). in which the promoter is fused to an untranslated *ypet* reporter. Promoter constructs were amplified using primers stated previously (Section 3.2.4), however with the use of pIJ12738*::rpoB_ACT* as template, cloned in pSNypet, then conjugated into M145 before testing of *ypet** expression by qRT-PCR (Section 3.2.4).

Consistent with the luciferase data, the ACT mutation again caused an increase in expression for all fusions most fusions. Furthermore, the presence of the UTR led to down regulation of the WT *rpoBp1* but not the ACT *rpoBp1* fusions (Fig 4.4.10.1). An exception to this was the ACT *rpoBp1-3U* fusion where the UTR did appear to reduce expression, suggesting the presence of a negative regulatory element upstream of *rpoBp2*. Interestingly, the activity of the ACT *rpoBp1* alone was not significantly higher than TTT *rpoBp1* implying that the intrinsic strength of the promoter was not changed by the mutation, and that the differences seen are due to the UTR.

Taken together these data provide further evidence that RT events occurring at the TSS are interacting with the 5’ UTR to down-regulate expression. Although the 5’ UTR region of *rpoBC* was characterised to a limited extent in Section 3.4, the nature of this interaction remains unknown.

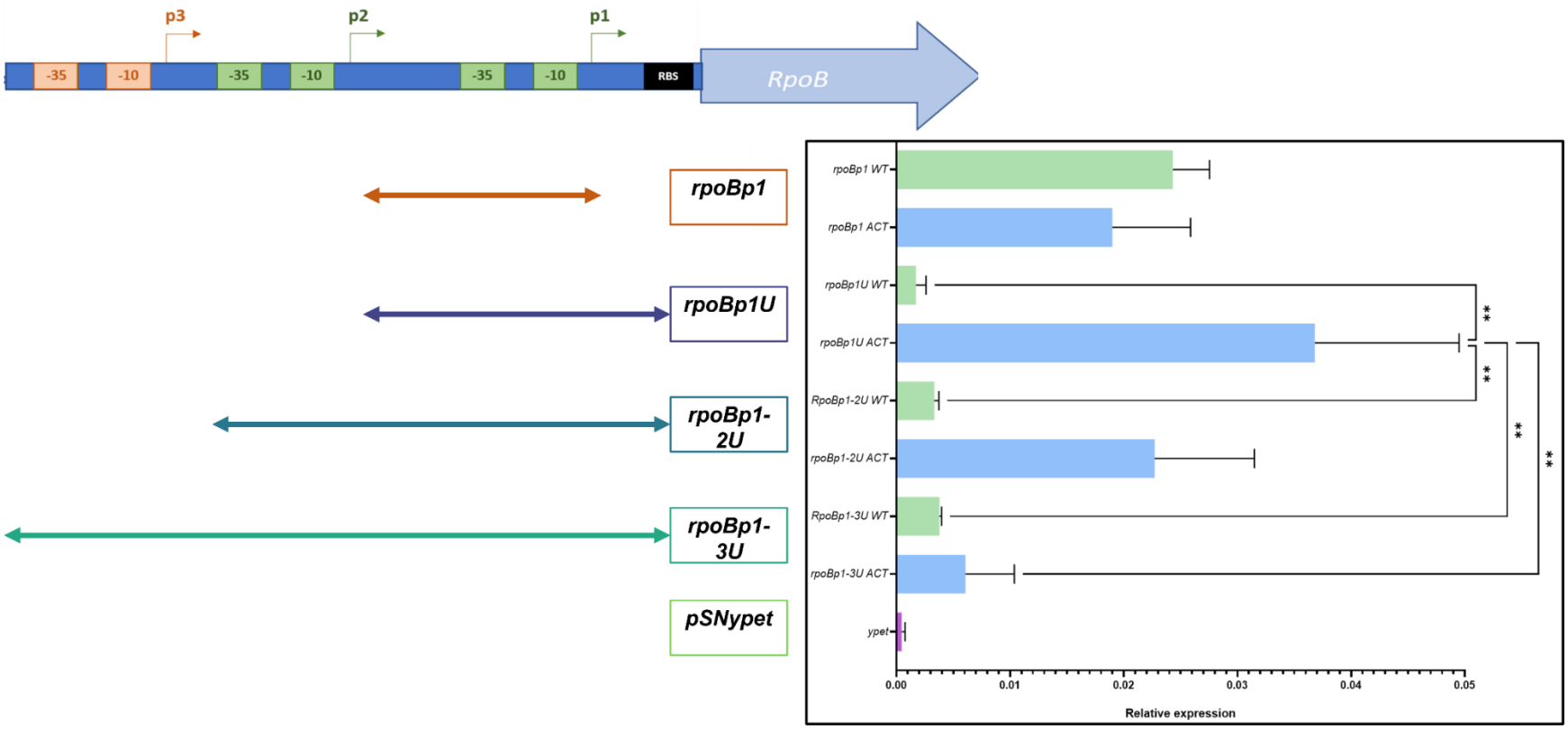


Figure 4.4.10.1: *ypet* reporter expression of wildtype and mutant *rpoBp* truncations during exponential growth. Strains were grown to an OD450 of ~0.8-1.0 in NMMP medium with *ypet** expression normalised to 16s stable rRNA, detected using primer pairs 65 and 66, and 67 and 68, respectively (Table 2.1.5.3). WT and ACT are the expression from promoter regions containing the wildtype T-tract (green) and mutant ACT TSS (blue) for *rpoBp1*, respectively. *pSNypet* represents empty vector expression. Data is from at least biological triplicates and error bars represent the standard deviation for each promoter region. Significance values were calculated using a one-way ANOVA; significance P < 0.01 = ** . Note that previous significance values of differences between wildtype *ypet::rpoBp* fusions, as seen in Figure 3.2.4.1, are not presented here.

4.4.11 The role of *rpoBp1* +4G in response to stringent conditions

Previously it was shown that RT at *rpoBp1* might contribute to stringent control of *rpoBC*, however the *rpoBp1U_ACT* promoter region, lacking the T-tract required for slippage, was still downregulated during amino acid starvation (Fig 4.4.11.1) Although the *rpoBp1* T-tract does not appear to be vital for stringent control, the +4 GTP, which is present in both tested wildtype and mutant ACT *rpoBp1U* constructs, might alternatively be responsible for downregulation (Figs 3.3.5.1 and 4.4.11.1); this is considering that rapid GTP depletion has been hypothesised to cause the downregulation at growth-related promoters, during stringent conditions in Gram-positive bacteria (Strauch *et al.*, 1991; Krásný *et al.*, 2008; Kriel *et al.*, 2014).

To determine the role of the +4 GTP nucleotide on the response of *rpoBp1* to stringent conditions, a G>C mutation was introduced into both wildtype and ACT mutant *rpoBp1U* promoter regions used previously (Section 3.2.3) producing constructs with predicted TSS of 5' TTTC and 5' ACTC, respectively; the -10 element and TSS used for each promoter construct is displayed (Fig 4.4.11.1). Mutant promoter regions were amplified by inverse PCR, using primer sets 22 and 23, and 21 and 23, respectively, and the pBluescript plasmid containing the *rpoBp1U* sequence constructed previously (Section 3.2.2). Correctly sequenced *rpoBp1U* regions were further subcloned into the pIJ5972 plasmid and conjugated into M145 (Section 3.2.3). Strains containing the mutant plasmids were grown in biological triplicate in 60 ml NMMP medium (containing amino acids)(Section 2.1.7.2), before a nutrient downshift and qRT-PCR was carried out as previously (Section 3.3.5).

Nutrient downshift led to both promoters being downregulated (Figs 4.4.11.2 and 4.4.11.3), with the mutant TTTC promoter downregulated ~1.9 fold at 5 min, less than the ~3.3 fold decrease observed for the same template containing this +4 GTP (5'TTTG) (Fig 3.3.5.1). This may provide initial evidence that +4 GTP might contribute to stringent control, which when removed, reduces sensitivity of the *rpoBp1* to surrounding concentration of GTP. The lowest point of expression for TTTC was 30 min after amino acid starvation, with a ~3.6 fold drop in expression, much less than observed for the wildtype promoter region (TTTG) where an ~11 fold drop in *luxA* expression is observed after this time, providing further evidence for a role of the +4

GTP. However, similar to the 5' TTTG promoter region, the 5'TTTC mutation is also not observed to recover to prestress *LuxA* expression levels after 60 min of stress.

LuxA expression was quantified during stringent conditions for the strain containing the 5' ACTC mutant plasmid (pIJ5972::*rpoBp1U_ACTC*) to further determine the dual effects of both RT and this +4 GTP at *rpoBp1*. A ~2.3 fold drop was observed only 5 min after the stringent response was induced with the lowest point of expression (~6.5 fold) observed 20 min after stress, showing this promoter region is still subjected to stringent control (Fig 4.4.11.3). Differing to the both the 5'TTTC and wildtype 5'TTTG regions, expression begins to increase around 60 min after stress, showing a slight recovery in transcription.

Overall, the results indicate that *rpoBp1* promoter maintains sensitivity to stringent conditions when +4 GTP is replaced, raising the question of whether in fact the *rpoBp1* promoter is sensitive to [GTP] .

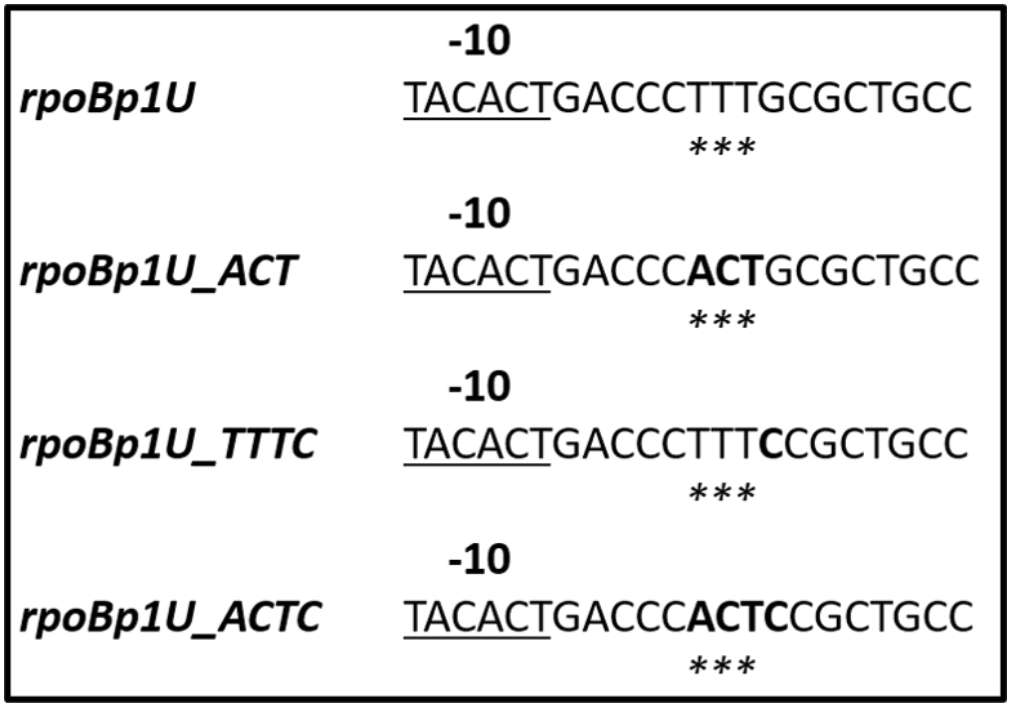


Figure 4.4.11.1: Promoter region and TSS used in mutant pIJ5972*::rpoBp1U* fusions. Part of the DNA sequence used for the differing *rpoBp1U* constructs. For each promoter region the -10 element is underlined and TSS is noted using a *. Mutations introduced can be identified in bold.

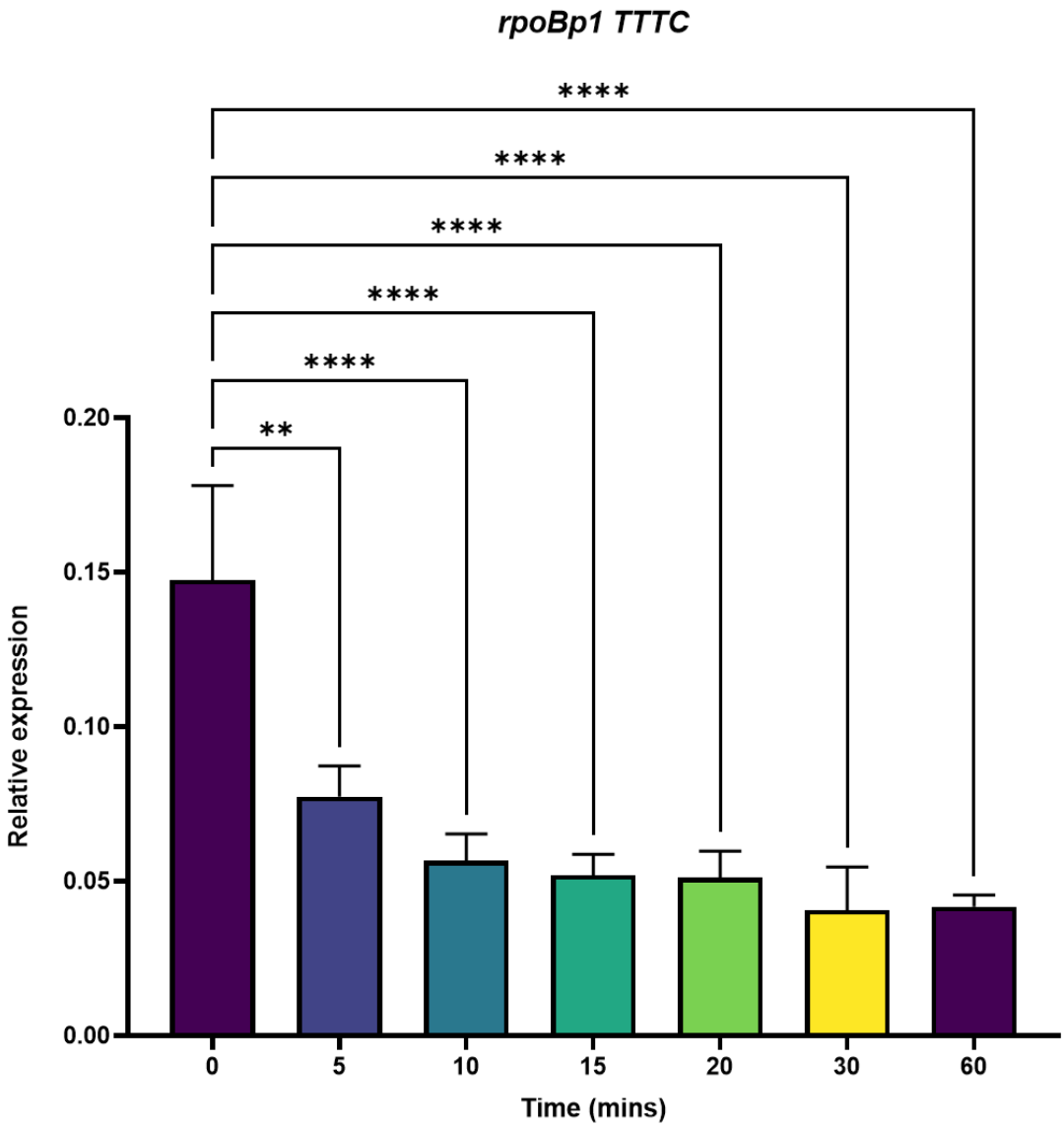


Figure 4.4.11.2: Expression from *rpoBp1_TTTC* during stringent conditions. The expression of *luxA* was determined using qRT-PCR and normalised to stable 16s rRNA using primers 73 and 74 and 67 and 68, respectively (Table 2.1.5.3). Transcripts were quantified using a standard curve of known copy number. Error bars represent the S.E.M for each data, performed in at least biological triplicate, with a one-way ANOVA used to determine significance of each data set compared to expression of *LuxA* at time zero; P < 0.0005 = ***, P < 0.0001 = ****.

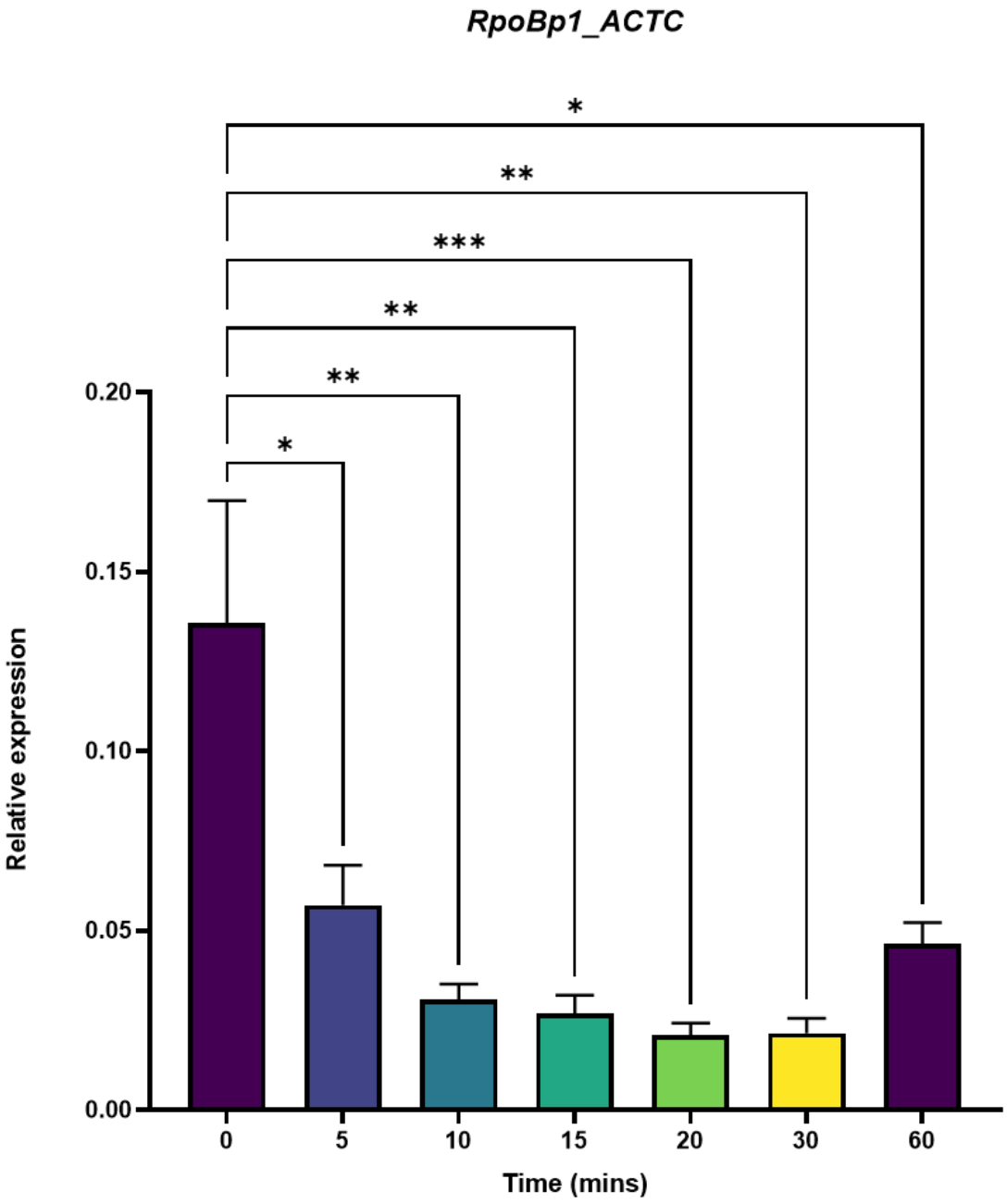


Figure 4.4.11.3: Expression from *rpoBp1_ACTC* during stringent conditions. The expression of *luxA* was determined using qRT-PCR and normalised to stable 16s rRNA using primers 73 and 74 and 67 and 68, respectively (Table 2.1.5.3). Transcripts were quantified using a standard curve of known copy number. Error bars represent the S.E.M for each data set, carried out in at least biological triplicate, with a one-way ANOVA used to determine significance of each data set compared to expression of *LuxA* at time zero; $P < 0.05$ = *, $P < 0.01$ = **, $P < 0.0005$ = ***.

4.5 Discussion

4.5.1 UTP-dependent reiterative transcription occurs at *rpoBp1*

Homopolymeric tracts consisting of three or more of the same nucleotide can be sensitive to RT during both transcription initiation and elongation, and has evolved into regulatory mechanisms in many cases (Molodtsov *et al*., 2014; Zhou *et al.*, 2013). In most documented examples, RT is used to sense nucleotide levels, and is therefore commonly associated with genes that are involved in nucleotide biosynthesis. The discovery of RT at the main promoter (*rpoBp1*) for the *rpoBC* operon in *Streptomyces* spp. was therefore intriguing. RT was detected from 5' end mapping data in both *S. coelicolor* (Jeong *et al.*, 2016a) and in *S. venezuelae* (Bush *et al.*, Pers comms, Ettwiller *et al.*, 2016), with 4 U's at the 5' end of the transcript being the most abundant form for both organisms, indicating that RT that adds at least one nucleotide occurs as part of the normal *rpoBp1* initiation. However, some transcription events result in much longer 5' polymeric tails, with up to 13 U's detected. *In vitro* transcription assays confirmed that *rpoBp1* slippage is UTP-dependent, with a ladder of transcripts, indicative of RT, observed when only UTP was included in assays (Fig 4.3.2.1), and with no RT seen when the T-tract was mutated (5'TTT > 5'ACT) (Fig 4.3.4.2). However, the role of RT, if any, in controlling *rpoB* expression remains unclear.

4.5.2 Reiterative transcription at *rpoBp1* negatively effects *rpoB* expression but has little effect on growth and morphogenesis

To determine the role of RT at the *rpoBp1* promoter, a stable natively-located TTT>ACT mutant was constructed; *in vitro* transcription assays confirmed that the ACT promoter remained active and that no slippage occurred (Fig 4.3.4.2). During exponential phase the mutant M145_*rpoB*_ACT strain displayed the highest level of *rpoB* expression when compared to both M145 and M571, suggesting that RT has a negative impact on expression. In theory, greater basal expression of *rpoB* observed in M145_*rpoB*_ACT may lead to an increase in both the concentration of β and β' subunits and, since the β and β' subunits are hypothesised to govern the overall production of the core enzyme itself (Dykxhoorn, St. Pierre and Linn, 1996), further affect RNAP levels and growth of the organism. Nonetheless, the M145_*rpoB*_ACT strain displayed little difference in both growth rate (Fig 4.4.4.1), and phenotype (Fig

4.4.4.3) when compared to M145; only a slight early onset/ increase in actinorhodin production was observed for the mutant (Fig 4.4.4.2). However, it is possible that changes in transcription could be countered by feedback autoregulation at the level of translation initiation, as was proposed for *E. coli* (Dykxhoorn *et al*., 1996), and therefore any changes in β and β' levels requires further investigation.

Confirmation that the increased *rpoB* transcription was due to increased *rpoBp1*-dependent transcription came from single cross-over experiments that used the non-replicating pSX400 plasmid to place *rpoBC* under the control of the *rpoBp1* promoter alone. A TTT>ACT mutant (S302) displayed a ~6.3 fold increase in *rpoB* expression compared to S301 containing the wildtype TSS (TTT) (Fig 4.2.2.1). However, both fusions showed lower *rpoB* transcript levels compared to M145, most likely due to the absence of the additional upstream promoters.

A simple explanation for the increased activity of the ACT *rpoBp1* mutant is that transcription initiates at a preferred +1A site and that this increases the rate of initiation. There is a strong preference for +1 purines and -1 pyrimidines at most bacterial promoters and mutations that alter e.g. +1A to +1C can drastically reduce initiation rates (Lewis and Adhya, 2004). However, this explanation does not appear to be valid because transcriptional fusion experiments with both luciferase and ypet* suggested that in the absence of the UTR, the TTT and ACT *rpoBp1* promoters had similar levels of activity (Figs 4.4.9.2 and 4.4.10.1); indeed, in the case of ypet* fusions, the TTT *rpoBp1* promoter had greater activity than the ACT mutant. Importantly, this further implies that there is a regulatory relationship between RT and the UTR.

4.5.3 Slippage events at *rpoBp1* may interact with the 5'UTR to downregulate *rpoB* expression

Transcriptional fusions results that used both *ypet* (Fig 3.2.4.1) and luciferase- (Fig 3.2.3.2) based reporter systems demonstrated that the *rpoB* 5'UTR has a detrimental impact on *rpoB* expression, suggesting the presence of a negatively acting regulatory element, which might include *rpoBa*, the potential attenuator that was originally identified in *S. griseus* (Küster, Piepersberg and Distler, 1998); however linker-based mutagenesis suggested that the region contains multiple negatively acting elements

(Section 3.4). Strikingly, the UTR did not appear to have a significant negative effect on expression for *rpoBp* truncations containing the mutant ACT TSS, for either of the reporter assays used (Figs 4.4.9.1 and 4.4.10.1). This suggests that slippage events occurring at *rpoBp1* cooperate somehow with the 5'UTR region to exert a negative regulatory effect. There are several examples where the polymeric product of RT interacts with downstream elements to control expression. For example, in *B. subtilis* RT on a poly-G tract that occurs at the promoter for the *pyrG* operon when CTP is deficient, leads to the formation of an anti-terminator, thereby preventing termination at the downstream terminator element (Meng, Turnbough and Switzer, 2004). However, in the case of *rpoB* RT acts negatively in concert with a downstream element and so might stimulate the formation of terminator. A preliminary analysis of the structure of the 5' UTR with additional U's did not reveal any obvious changes to the secondary structure (Fig 3.4.1.2); further work is required to investigate this intriguing relationship.

4.5.4 Reiterative transcription at *rpoBp1* is sensitive to the +4 nucleotide

In vitro transcription experiments confirmed the dependence on UTP for both transcription initiation and RT at the *rpoBp1* promoter. RT has been shown to not only be sensitive to tract-encoded nucleotide levels, but also those encoded directly after the tract, which are required to commit these slippage events to continued transcription of downstream DNA; taking the *B. subtilis pyrG* as an example again, slippage at the TSS is sensitive to both GTP and CTP nucleotides, with the latter required to prevent further slippage by commitment at the +4 nucleotide (5' **G**GGC, TSS in bold) (Meng, Turnbough and Switzer, 2004). A similar mechanism of control was confirmed for *rpoBp1* using *in vitro* transcription reactions lacking GTP, with a ladder of slippage products produced (Fig 4.3.4.2). The nature of +4 nucleotide did not appear to be important since a mutant 5' TTTC *rpoBp1* template showed extensive RT when CTP was removed (Fig 4.3.4.2).

4.5.5 Intracellular nucleotide levels may contribute to control of reiterative transcription at *rpoBp1* in response to stringent conditions

Given that RT occurs at *rpoBp1*, and that the +4G is highly conserved, it was hypothesised that this simple TTTG element might be of particular importance for

stringent control *(*Fig 3.3.4.2), where it is hypothesised that the production of ppGpp causes the rapid decrease in [GTP] and inhibition of transcription at promoters where GTP is present within the first few iNTP (Strauch *et al.*, 1991; Kriel *et al.*, 2014). In this chapter, the role of +4G was investigated both *in vivo* and *in vitro*. *luxA* transcript levels during nutrient limitation in M145 (pIJ5972::*rpoBp1U_TTTC*), still decreased, however, to a lesser extent compared to the wildtype promoter region; ~1.9 and ~3.6 fold decrease in transcript levels 5 and 30 min after the induction of stress, respectively, versus ~3.3 and ~11 fold decrease observed for the same time points in WT M145 pIJ5972::*rpoBp1U* (Figs 3.3.5.1 and 4.4.11.2). This suggests the +4 GTP contributes to stringent control of *rpoBp1*, but that other mechanisms are also involved. *In vitro* transcription assay results are in agreement with this, as promoters containing the TTTG (wildtype) and TTTC (mutant) TSS displayed the same rate of decreased run-off transcription during a synthetic GTP depletion (Figs 4.3.5.1 and 4.3.7.1). One possibility is that the conserved downstream +6G plays an important role in sensing and responding to [GTP]. For example, in *Staphylococcus aureus*, stringent control of rRNA promoters is effected by GTP nucleotides at both +1 and +4 (Kästle *et al.*, 2015).

Intracellular GTP concentrations alone are unlikely to be the sole effector for the down regulation of *rpoB* in response to stringent conditions. UTP levels may be important for determining the extent of slippage and thus transcriptional outcome, with elevated UTP levels increasing RT at *rpoBp1*. However, using *in vitro* transcription assays, run-off transcription at *rpoBp1* displayed the same trend in band intensity when GTP depletion assays were carried out in both high UTP (1000 μM) and equimolar concentrations of nucleotides (Fig 4.3.8.1). However, due to the partially synthetic environment of an *in vitro* transcription assay, the conditions may not be correct to accurately recreate nucleotide levels and intracellular concentrations during stringent conditions. In agreement with this, the M145_*rpoB*_ACT mutant still displayed sensitivity to stringent conditions, with *rpoB* expression still observed to decrease, however, values were not observed to be significant (Fig 4.4.6.1). This downregulation in expression for the ACT construct during stringent conditions, may be a result of the +4 GTP present and further suggest an additive role by both RT and GTP, for transcriptional control of *rpoBp1*. This was however, disputed by the fact that the

ACTC mutant (pIJ5972*_rpoBp1_ACTC*), which lacks both the slippage prone tract and the identified +4 GTP, was still subjected to stringent control (Fig 4.4.11.3).

UTP-dependent RT has, however, largely been associated with termination of transcription on these tracts (Turnbough and Switzer, 2008), which have not been quantified in this instance. Nevertheless, it is likely that multiple mechanisms of control are elicited to downregulate the *rpoBp1* promoter, aside from those highlighted above. More recently, inhibition of the polynucleotide phosphorylase (PNPase) enzyme was shown to occur during the stringent response, and resulted in increased stability of bulk RNA which may enable transcription to also be redirected due to changes in RNA turnover (Gatewood and Jones, 2010), providing an alternative to nucleotide based control. The rapid downregulation in transcription of highly expressed growth related genes, including translational apparatus such as rRNA and ribosomal proteins, likely further effects other processes intracellularly by feedback from ceased translation, and may also free up more RNAP, that is normally associated with these *rrn* genes, for activation of alternative promoters (Jin, Cagilero and Zhou, 2012; Ross *et al.*, 2014).

Nevertheless, it would be beneficial to confirm the downregulation and sensitivity of *rpoBp1 in vivo* using S1 nuclease mapping experiments. Further work characterising both GTP levels, and the expression from other promoters where GTP is within the first few iNTP during stringent conditions would also be beneficial to confirm this role of GTP.

Summary

Overall, the main promoter for *rpoBC* is confirmed to be under control by UTP-dependent RT. As far as is known, this is the first characterisation of RT in not only the *Streptomyces* genus, but also the Actinobacteria phylum. It is further likely that these slippage events play a crucial role in determining the level of expression of *rpoBC* both during standard conditions and in response to intracellular stress, which may be of particular importance in regards to the activation of silent biosynthetic gene clusters.

Chapter 5: Global analysis of reiterative transcription in the *Streptomyces* genus

5.1 Overview

Until recently, most examples of RT in bacteria involved genes involved in nucleotide metabolism, but advances in NGS technology has led to the realisation that RT is likely apparent across all kingdoms of life and might play a regulatory role in many cases (Turnbough, 2011). For example, global 5'-end mapping experiments revealed extensive RT in *Streptococcus agalactiae*, with up to 15% of TSS subjected to some form of slippage, suggesting this mechanism plays an important role in gene expression control in this organism (Rosinski-Chupin *et al.*, 2015). In another example, an NGS approach called MASTER (massively systematic transcript end readout) was applied to *E. coli* both *in vivo* and *in vitro*, and suggested that slippage occurs at the majority of promoters (Vvedenskaya *et al.*, 2015). In this chapter, we investigate if RT events are widespread in the *Streptomyces* genus, which may reveal a wider regulatory role.

5.2 RNA sequencing reveals multiple sites within the *Streptomyces* genome are subjected to RT

To conduct a preliminary analysis of the extent of RT in *Streptomyces* spp. during standard growth conditions, TSS-mapping data from *S. coelicolor* (Section 2.2.3.4) and *S. venezuelae* (Bush M, and Buttner, M, pers comms), were examined to identify further genomic sites subjected to slippage. Using a search substring containing 6 consecutive bases of each nucleotide (e.g TTTTTT), reads were extracted using the intext search (grep) of raw FASTQ reads using the Galaxy server (Afgan *et al.*, 2018). Note that this approach would not identify examples of RT where there are less than six consecutive bases at the 5' end and will also fail to identify dinucleotide slippage events that can occur at significant levels (Vvedenskaya *et al.*, 2015). All reads were extracted from the data collected from the triplicate growth of *S. coelicolor* M145 during exponential phase (time zero), and from the 10 h growth of *S. venezuelae*. Identified examples of RT were linked to their corresponding genes by applying BLAST on StrepDB. The high GC content of *Streptomyces* resulted in several false positive examples with long G/C tracts; however, several examples of RT were identified. Nonetheless, it was clear that, while RT occurs at several promoters, the extent of RT was highly variable and so it was decided to filter targets based on the presence of ≥ 3 differing transcripts of varying tract lengths (transcripts representing

at least three different slippage events). The use of TSS-mapping data in the present study also allowed TSS to be confirmed/compared to those identified previously (Jeong *et al.*, 2016a).

Due to the focus of the present study on *S. coelicolor*, all RT events confirmed in this organism alone, were extracted, with the number of the orthologous gene (if present) in *S. venezuelae* (SVEN15) genome provided if RT was also observed for this gene. In total, 29 promoters/genes were subjected to RT in *S. coelicolor* with 17 of those shared with *S. venezuelae*. Notably *S. venezuelae* revealed several additional instances of RT not seen in *S. coelicolor* (Section 7.5), suggesting that only some examples are likely to be used as part of a regulatory mechanism.

Table 5.2.1: Reiterative transcription in the *Streptomyces* genus. FASTQ reads from *S. coelicolor* (898227) and *S. venezuelae* (394939) were extracted using the intext (grep) search on Galaxy and used to identify areas of the genome subjected to slippage. StrepDB and BLASTn were used to identify the respective locus, while confirming the additional nucleotides encoded in transcripts. Gene name and number, sequence (30 bp), including tract sensitive to RT (bold), strand (+/-), location in respect to TSS and distance from start codon are provided for each example. A separate column is provided to determine if the TSS was also identified by Jeong and colleagues (2016) (identified previously) and the abundance of transcripts initiating at this TSS; (P) and (S) represent primary and secondary TSS for genes, determined by the number of transcripts identified to initiate here previously, and (N) representing TSS not identified previously. P* represents the TSS identified previously, however was not associated to newly discovered gene SCO2898a. The sequence provided is that from the *S. coelicolor* genome, with the *S. venezuelae* gene number provided for reference if RT observed in both organisms.

No	Gene	Sequence (5' -> 3')	Strand	Location of RT	Distance from start codon (bp)	Identified previously? (P, S or N)
UTP-dependent						
1	***rpoB* – RNA polymerase beta chain** (SCO4654, SVEN15_4249)	**TTT**GCGCTGCCTGTTAGCTGCCCCCTGCCC	+	RT on TSS	178	P
2	***pheS* – phenylalanyl-tRNA synthetase** (SCO1595, SVEN15_1154)	**TTT**GGCACCTTTGCGTCCAGATCCGCAGAC	-	RT on TSS	112	P
3	**Malate oxidoreductase** (SCO5261, SVEN15_4846)	**TTT**ACACCCCTCATCAGGGGCTCAGGGCGC	+	RT on TSS	157	P
4	***cvnA13* - Histidine kinase** (SCO7463)	G**TTT**GCGGAGCGGGCGGACGGGTGAAATA	+	RT 1 bp after TSS	147	P
5	**Potential membrane and sporulation control protein** (SCO0247)	G**TTT**GTAGAAACAGGGCCGGGGGCAACCCG	+	RT 1 bp after TSS	243	P
6	**Xanthine/uracil permease** (SCO4334, SVEN15_4055)	G**TTT**GTCTCTGACCGCTTTCATTCGTAGAA	+	RT 1 bp after TSS	89	P
7	**ATP-dependent RNA helicase** (SCO4096, SVEN15_3762)	G**TTT**GATCCCATTTGCCCGGCGCCACCGCA	+	RT 1 bp after TSS	123	P
8	***cvnA13* - sensor histidine kinase** (SCO7463)	G**TTT**GCGGAGCGGGCGGACGGGTGAAAT	+	RT 1 bp after TSS	146	P

9	***accA2* - acyl-CoA carboxylase complex** (SCO4921, SVEN15_4491)	G**TTT**CAAGGAGGGAGCCTCAATCGTGCGCA	-	RT 1 bp after TSS	22	S
10	**Putative DNA recombinase** (SCO3660)	**TTT**GACGTATCTGATGTGCCGGTGTGCTCG	-	RT on TSS	56	N
11	**Hypothetical protein** (SCO6240)	ATGAAGG**TTTT**CATCGCTGGTGGGCGCGGC	-	RT on TSS	8bp within coding region	N
12	**Hypothetical membrane protein** (SCO2896 - likely operon SCO2893-2896)	**TTT**GGTGCCGGTCGGAGGGGGAGGGCTGTG	-	RT on TSS	27	N
13	**Hypothetical protein** (SCO3535)	**TTT**ACCCTCAGTGGCGCAGCGAACATGACG	+	RT on TSS	389	N
14	***aspS* - aspartyl tRNA synthetase** (SCO3795)	**TTT**CACGTGAAACGCGAAACACCATCCCCG	-	RT on TSS	55	N
15	**Possible integral membrane protein** (SCO3083)	**TTT**GCCGTTCTATTACTCTTGAGCCAAGGC	-	RT on TSS	57	N
16	**Excinuclease ABC subunit A** (SCO0918, SVEN15_6694)	G**TTT**GTCCCACCCACCGTACGGAGCCCCCA	+	RT 1 bp after TSS	32	N
17	**Putative large secreted protein** (SCO1402)	G**TTT**GTTGACGCACATGAGTACCCCCACCG	-	RT 1 bp after TSS	53	N
CTP-dependent						
18	***pyrG* - CTP synthetase** (SCO1776, SVEN15_1369)	**CCC**GTGGACCGGTGGGAGAAAAACCCCGGA	-	RT on TSS	128	P
19	**Transport protein** (SCO0742, SVEN15_6902)	**CCC**GTCCGTGTGCCGAAGGAGCGACGTGAC	+	RT on TSS	25	N
20	***sugE*- molecular chaperone** (SCO2898, SVEN15_2589)	A**CCCC**GACGAGACGCGCCGTACCGGGCTGA	-	RT 1 bp after TSS	78	N
21	***slzA* - small leucine zipper protein** (SCO5576a, SVEN15_5162)	**CCCC**GCGCACCTTAGCTGTTTGGCATGGTT	+	RT on TSS	130	P*

ATP-dependent						
22	***ftsZ* - Cell division protein** (SCO2082, SVEN15_1692)	**AAAAA**CGGGAGGTTCGGCGTGTTCGTTGA	-	RT on TSS	223	S
23	**Phosphatidyl serine decarboxylase** (SCO6468, SVEN15_6161)	**AAA**GCCGCCGTCCCCCGTCATAGCGCGGCA	-	RT on TSS	50	S
GTP - dependent						
24	**ABC transporter ATP binding subunit** (SCO5393, SVEN15_4939)	AC**GGGGG**CGTGTGCCCTGCTGTCCCTGAAC	-	RT 2 bp after TSS	47	P
25	**Membrane protein** (SCO3905, SVEN15_3594)	A**GGG**ATGTGAGGGCTAAAACACACTGCCTT	-	RT 1 bp after TSS	45	P
26	**Membrane protein** (SCO2970, SVEN15_2669)	**GGGG**AACGTTCGCGTCGTCGCGGGCGTTGT	+	RT on TSS	63	S
27	**HupA DNA binding protein** (SCO2950, SVEN15_2658).	**GGG**CGCGAGGGTCTGACGACCGACCCGGGT	-	RT on TSS	135	P
28	**Hypothetical protein** (SCO4136, SVEN15_3802)	A**GGGGGG**CGACGCCGGGCGTTCCGCAAAGGGCC	-	RT 1 bp after TSS	98	N
29	**Transcriptional regulator** (SCO0132)	AA**GGGG**CGATGCTTATGGTGAGAAAGTTGC	+	RT 2 bp after TSS	12	N

5.3 Nutrient limitation increases the extent of RT at *rpoBp1*

The data presented in Chapter 4 revealed that RT occurs at *rpoBp1* and that limitation of GTP can cause an increase in RT product formation *in vitro*. Furthermore, it was shown that the T-tract interacts somehow with the UTR to negatively control promoter activity. Our working hypothesis was that GTP limitation, caused by ppGpp production, and the resulting inhibition of GTP synthesis would lead to increased RT at *rpoBp1*, and that this would underlie the observed stringent control of the *rpoBC* operon. To investigate further it was decided to perform 5' cappable-seq mapping on M145 and M571 (ΔrelA) before and after (10 min) amino acid limitation (Section 2.2.3.4), and investigate the extent of RT at *rpoBp1* using this data. For extraction of identical reads initiating at *rpoBp1*, an intext grep search was carried on the raw sequencing data using the galaxy platform (Afgan *et al.*, 2018) and a 60 bp sequence containing DNA directly downstream from the homopolymeric T-tract (TSS) (5'GCGCTGCCTGTTAGCTGCCCCCTGCCCGTCACCAGGGGTCTACCCTCG CCCGAGCACTGACGAC). Data was counted in respect to the number of 5' T nucleotides and normalised. The sum of reads was determined for individual strains M145 and M571 (0 and 10 mins) and averaged before the sum of each individual replicate (1-3) was divided by this average to provide a normalisation factor for each biological replicate.

Prior to amino acid limitation, as was observed earlier, there is extensive RT with the majority of *rpoBp1* transcripts having 4 U additions. Nutrient limitation caused a small but significant change in RT products; after 10 min the proportion of transcripts where no RT is apparent (0-3xU) decreased, with a corresponding significant increase in the proportion of products with 7 or more 5'Us (note that it is assumed that NGS sequenced T corresponds to U although this has not been formally confirmed) (Fig 5.3.1).

This data provides preliminary evidence that RT increases during nutrient limitation and are consistent with the model that GTP limitation increase RT. However, a similar trend was seen with M571 (ΔrelA) and, although the differences seen were not significant, this suggests that the stringent response is not responsible for these changes (Fig 5.3.2).

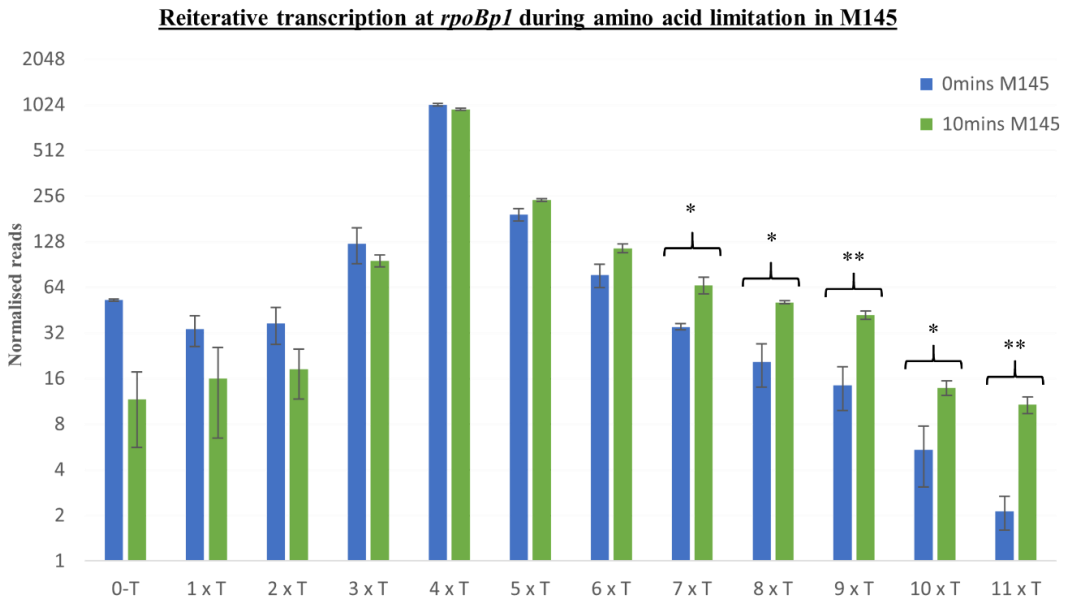


Figure 5.3.1: Reiterative transcription at *rpoBp1* TSS during exponential growth and amino acid starvation in M145. Transcripts initiating at *rpoBp1* were extracted from the TSS-mapping data (Ettwiller *et al.*, 2016) collected before and after (10 min) the induction of stringent conditions in *S. coelicolor* strain M145 (Section 2.2.3.4) using the galaxy platform (Afgan *et al.*, 2018) and a 30 bp substring homologous to *rpoBp* (Section 4.2.1). Transcripts are organised according to the number of T's at the 5' end, with reads extracted from exponential phase (0 mins) and after amino acid starvation (10 mins) shown in blue and green, respectively. Error bars represent the S.E.M from data collected in biological triplicate, with statistical analysis carried out using an unpaired two-tailed T-test; $P < 0.05$ = *, $P < 0.01$ = **. Figure and analysis carried out by Dr Murat Eravci.

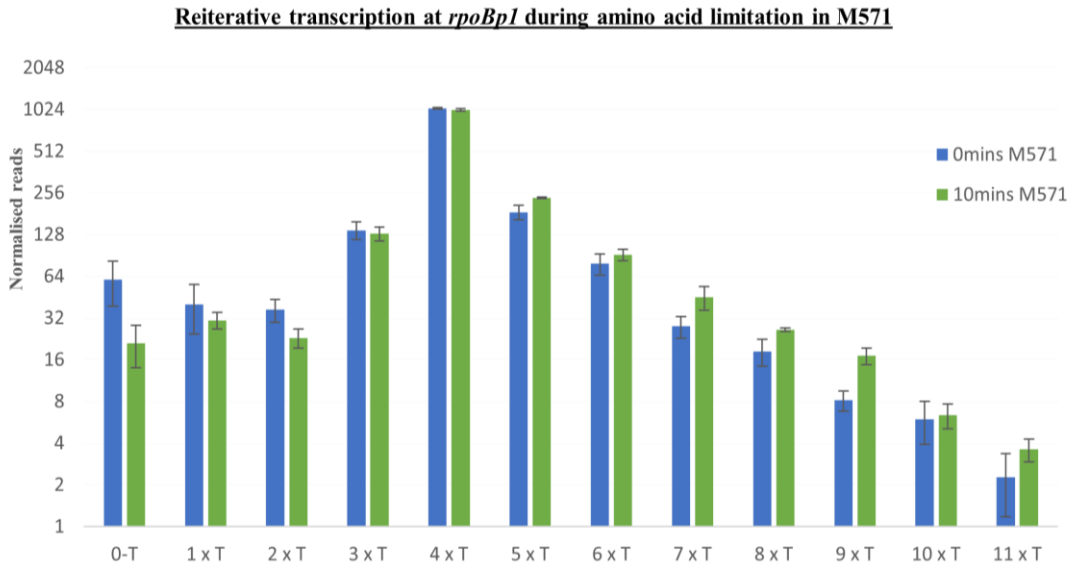


Figure 5.3.2: Reiterative transcription at *rpoBp1* TSS during exponential growth and amino acid starvation in M571. Transcripts initiating at *rpoBp1* were extracted from the TSS-mapping data (Ettwiller *et al.*, 2016) collected before and after (10 min) the induction of stringent conditions in *S. coelicolor* strain M571 (Section 2.2.3.4) using the galaxy platform (Afgan *et al.*, 2018) and a 30 bp substring homologous to *rpoBp* (Section 4.2.1). Transcripts are organised according to the number of T's at the 5' end, with reads extracted from exponential phase (0 mins) and after amino acid starvation (10 mins) shown in blue and green, respectively. Error bars represent the S.E.M from data collected in biological triplicate, with statistical analysis carried out using an unpaired two-tailed T-test. Figure and analysis carried out Dr Murat Eravci.

5.4 Discussion

5.4.1 Reiterative transcription occurs at the TSS of several genes in *Streptomyces*

The first identified case of RT, which was initially reported in *E. coli*, was simply regarded as an artefact then and for several years after (Chamberlin and Berg, 1964; Jacques and Kolakofsky, 1991). However, since this time further examples of this mechanism have come to light providing further evidence for its role in control of gene expression (Reviewed by Turnbough, 2011). RT during elongation often regulates gene expression in viruses, for example by the introduction of an RNA frame-shift and the subsequent production of differing viral proteins; however only a few instances of RT having a regulatory role have been reported in bacteria, and in the majority of cases these are for genes involved in nucleotide metabolism. High-throughout sequencing technology is now enabling genome-wide searches for RT, and this is the first time a

search for RT at initiation sites has been performed in the *Streptomyces* genus. In this limited study, together with *rpoBp1*, evidence of RT was found at an additional 28 *S. coelicolor* promoters, with UTP-dependent slippage observed to be most common (17 promoters; Table 5.2.1). It is possible that the true number is higher since the 5'cappable seq method will only detect transcripts where RNAP has escaped the promoter and entered into elongation; UTP-dependent RT most commonly leads to premature transcription termination and such an event would not be detected (Turnbough and Switzer, 2008).

Interestingly only two promoters were subjected to ATP-dependent slippage in *Streptomyces*, in contrast to *S. agalactiae* where the majority of slippage was ATP-dependent (~67% of slippage events) (Rosinski-Chupin *et al.*, 2015), however, both A- and T- associated homopolymeric tracts displayed higher instances of RT in *E.coli* (Vvedenskaya *et al.*, 2015). As well as this, up to 8 additional nucleotides were observed to be added to transcripts as a result of slippage in *E. coli,* less than the number already observed in *Streptomyces* in the present study; slippage of up to 15 and 11 additional nucleotides were identified at the 5' end of *rpoBp1* transcripts in *S. venezuelae* and *S. coelicolor,* respectively (Figs 4.2.1.1 and 4.2.1.3). A loose hypothesis for this may be due to differences in structure of RNAP and the kinetics of the reiteratively transcribing complex between the organisms, however, crystal structures monitoring slippage events have only been published for *Thermus thermophilus* RNAP (Shin, Hedglin and Murakami, 2020; Liu *et al.*, 2022). Nevertheless, a full analysis of TSS and the extent of RT at a variety of genes would be favourable to fully determine if certain homopolymeric tracts are more favourable, and if slippage is consistently higher in *Streptomyces* when compared to other organisms such as *E. coli* (Vvedenskaya *et al.*, 2015).

Notably, the initial global analysis of RT in the present study was carried out using data extracted from exponential phase cultures only. As well as this, the use of TSS-mapping data means the identification of slippage events is limited to those identified during initiation. For further work it may be favourable to determine the occurrence of slippage during both stationary phase and other intracellular stress, such as phosphate limitation, as well as perform global scale RNA-sequencing to determine areas of the genome subjected to RT during both transcription elongation and termination.

5.4.2 Genes involved in nucleotide biosynthesis and translation are among the genes subjected to reiterative transcription

Several of the reported genes have great importance for cell growth and amino-acid biosynthesis such as RNA helicase (7), serine decarboxylase (23) and two aminoacyl-tRNA synthetases *pheS* (2) and *aspS* (14). The finding that two amino-acyl tRNA synthetases might be controlled by RT is intriguing and might reveal a mechanism that links nucleotide levels with translational capacity. It was previously hypothesised that *pheS* was under slippage control due to the presence of at least 3 consecutive T-A base pairs near the beginning of the initial transcribed region in *E. coli* (Springer *et al.*, 1985; Liu, Heath and Turnbough Jr., 1994). This may suggest this mechanism of control is widespread within the bacterial kingdom, however, this is yet to be confirmed. Although both *pheS* and *aspS* initiate with a poly-U tract, the subsequent nucleotide differs, CTP and GTP for *aspS* and *pheS*, respectively, raising the possibility that the promoters respond to different signals. In lower G+C Gram positive bacteria, aminoacyl-tRNA synthetase expression is commonly controlled by T-box attenuation systems, indeed there are predicted examples that control both *pheS* and *aspS* (Henkin and Grundy, 2006; Wels *et al.*, 2008); however riboswitches appear to be rare in *Streptomyces* (Marchand *et al.*, 2021), and T-boxes have already been determined as absent from regions upstream of both *pheS* and *aspS* in Actinobacteria (Seliverstov *et al.*, 2005; Wels *et al.*, 2008) raising the possibility of alternative modes of regulation.

RT was detected at promoters for xanthine/uracil permease (6) and CTP synthetase (*pyrG*) (18), suggesting that these genes involved in nucleotide biosynthesis/salvage might use a RT-based regulatory mechanism to sense nucleotides in *Streptomyces* (see Table 5.2.1). Although slippage at the TSS of a uracil permease gene promoter has not been reported before, slippage control of *pyrG* has already been characterised in both *B. subtilis* and *S. agalactiae* (Meng, Turnbough and Switzer, 2004; Rosinski-Chupin *et al.*, 2015). However, unlike the case of RT control in *B. subtilis* and *S. agalactiae*, where low CTP levels to induce slippage encoding up to 8 GTP nucleotides leading to reduced expression, in *Streptomyces*, RT occurs on a homopolymeric C-tract, implying that slippage is likely to have a positive effect on gene expression. In the case of *B. subtilis* and *S. agalactiae* the incorporation of multiple G residues causes the formation of an anti-terminator hairpin, further preventing terminator hairpin

formation, and thus increasing transcription of *pyrG* (Shin, Hedglin and Murakami, 2020). In the case of *Streptomyces* slippage might lead to the formation of a terminator which would reduce *pyrG* expression, a hypothesis that requires testing.

5.4.3 Is GTP of particular importance for RT control in *Streptomyces?*

Expression of genes regulated by RT is seen to be strongly influenced by intracellular stress and the further effects it has on intracellular nucleotide pools (Lee *et al.*, 2009; Rosinski-Chupin *et al.*, 2015). Previously, this study discussed the likely importance of intracellular GTP levels in controlling gene expression in *Streptomyces* in response to stress, similar to that observed in the *Bacillus* genus (Section 1.2.5.2) (Strauch *et al.*, 1991). *B. subtilis* is reliant on GTP homeostasis as a result of (p)ppGpp synthesis/hydrolysis, to alter gene expression accordingly, with a rapid drop in GTP concentration that is observed during stringent conditions, causing a inhibition of promoters where GTP is the iNTP (Turnbough, 2008; Kriel *et al.*, 2013, 2014). A drop in GTP concentration during the stringent response has also been reported in *S. coelicolor* and thus may work in a similar way (Strauch *et al.*, 1991). This in addition to the fact RT is largely determined by intracellular nucleotide concentrations, may allow for increased sensitivity of certain promoters to stringent conditions, more specifically those that contain GTP as the homopolymeric tract or as the following nucleotide. Out of the 23 genes that appeared to be subjected to ATP, CTP or UTP-dependent RT, 14 (60%) contained GTP as the nucleotide directly after the homopolymeric tract, including the *rpoBp1* promoter largely characterised in this study. GTP in these cases likely influences the extent of RT at these TSS, with binding of a GTP nucleotide committing RNAP to canonical transcription. A further 6 genes were under the control of GTP-dependent RT, with high concentrations of GTP likely required for slippage events to occur on these tracts.

The study of the effects the stringent response has on RT at the *rpoBp1* promoter (Fig 5.3.1) supported this model in which slippage is influenced by [GTP]; A significant increase in the number of transcripts containing ≥7 T's at the 5' end was observed after the induction of amino acid starvation in M145, when compared to exponentially grown cultures. However, at this point further work is necessary to confirm that changes in [GTP] during the stringent response is the main instigator of transcriptional response in *Streptomyces*. As well as this, the outcome and the effects of GTP

concentration on the extent of RT likely differs between *rpoBp1* and the other genes identified, with further work necessary to characterise the influence and outcome of slippage at these varying promoters.

Summary

Overall, this work suggests the role of RT in *Streptomyces* as well as other organisms is more extensive than initially thought, yet much more work is required to elucidate the full mechanisms of action and effects on the expression of differing genes with slippage prone TSS. A more comprehensive search for RT by the inclusion of TSS mapping data and transcriptome data from other *Streptomyces* organisms and those within the Actinobacteria phylum would likely reveal further examples of this mechanism, however there is little variety of data available to analyse. Nevertheless, this work displays the dynamicity of the *Streptomyces* transcriptome and provides further evidence for the importance of nucleotide concentration in gene expression control, which could be the key to understanding transcriptional control in *Streptomyces* on a global scale.

Chapter 6: General discussion

6.1 Overview

Despite its crucial role in the first main contributor to gene expression, relatively little is known about how the production of RNAP is controlled in bacteria including the *Streptomyces* genus. *Streptomyces* organisms have long been exploited for their ability to produce a great plethora of useful secondary metabolites, which includes ~ 80% of the natural resourced antibiotics in use today (Hoskisson and Fernández-Martínez, 2018). The explosion in genome sequencing has underlined this capacity while revealing that this resource is largely untapped as many large biosynthetic gene clusters are not expressed under standard conditions. The discovery that mutations in *rpoB,* encoding the β subunit, can increase production of both known and novel antibiotics (Gomez-Escribano & Bibb, 2011; Hosaka *et al.*, 2009) provoked this study, with the proposal that a further understanding of RNAP control may ultimately give rise to new approaches to activate these silent clusters for the production of novel antibiotics.

Most studies on the regulation of RNAP production have focused on *E. coli*. Here the α-subunit was determined to be produced in >2 fold excess compared to the production of the β and β' subunits, and thus the assembly of the RNAP core is dependent on the concentration of these large subunits (Engbaek, Gross and Burgess, 1976; Dykxhoorn, St. Pierre and Linn, 1996). Hence, it follows that understanding the control of the expression of these subunits could be crucial to understanding how bacteria regulate RNAP production as a whole. Overall, this thesis, has largely contributed to the understanding of transcriptional control of the RNAP enzyme in *S. coelicolor,* opening the door for further work.

6.2 Highlights of this study

rpoBC promoter identification

Previous work in *E. coli* showed that the *rpoBC* genes were co-transcribed with the upstream ribosomal proteins, *rpIJL,* from one main promoter (*rplJp*) (Dennis, 1977; Newman, Linn and Hayward, 1979). Furthermore, specific autoregulatory control of *rpoBC* expression is thought to occur primarily at the level of translation (Passador and Linn, 1992). However, in Gram-positive bacteria, including the Actinobacteria, while the genetic organisation of *rplJL–rpoBC* is retained, expression has been

decoupled, since transcription of *rpoBC* initiates in the *rplL-rpoB* intergenic region (Boor, Duncan and Price, 1995; Küster, Piepersberg and Distler, 1998), suggesting that control might be exerted at the transcription initiation level. This indeed appears to be the case in *M. tuberculosis* where control of *rpoB* promoter activity has been found to alter tolerance to rifampicin (Zhu *et al.*, 2018). The initial aims of this study were therefore to identify promoters acting upstream of the *rpoBC* operon. Transcriptional fusion assays, along with published and new 5'-end-mapping data, revealed three promoters: *rpoBp1, rpoBp2* and *rpoBp3* (Section 3.2). The promoter elements for *rpoBp1* and *rpoBp2* resemble the consensus sequence recognised by the principal σ^{HrdB} factor, while *rpoBp3 was* hypothesised to be activated by σ^{ShbA}-containing holoenzyme. *rpoBp1* appears to be the main promoter for the *rpoBC* operon, and the dependence of this promoter on σ^{HrdB} was confirmed using *in vitro* transcription assays (Fig 4.3.1.2.1) (Jeong *et al.*, 2016). This is supported by the high level of promoter element conservation in all *Streptomyces* strains studied. Transcriptional fusions revealed slight increases in reporter expression when *rpoBp2* and *rpoBp3* were included in promoter fusions (Fig 3.2.3.2), however it is possible that these promoters contribute to a larger extent under certain growth conditions.

rpoBC 5'UTR causes downregulation of *rpoBC* expression

The *rpoBC* operon has an extensive 5'UTR of 160 bp (from *rpoBp1*), and this study detailed that it plays a key negative role in the regulation of expression. In all cases where the WT promoter was used, transcriptional reporter fusions containing *rpoBp1* plus the 5'UTR displayed a much lower transcriptional output compared to the *rpoBp1* fusion lacking this downstream region (Figs 3.2.3.2 and 3.2.4.1). An inverted repeat identified in the 5'UTR in *S. coelicolor* was similar to that previously identified in *S. griseus* (*rpoBa*) where it was hypothesised to be acting as an attenuator (Küster, Piepersberg and Distler, 1998). It is worth noting that a stem-loop attenuator is present in the *E. coli rplJL-rpoBC* intergenic region, that allows the transcriptional uncoupling from the upstream *rpIJL* genes (Dennis, 1977; Yamamoto and Nomura, 1978). However, the regulatory role of *rpoBa* differs in that it is likely to be controlling the progression of transcription that initiates at *rpoBp1-3*. Nonetheless, linker mutagenesis suggested that the role of this element was not straightforward since mutations that

disrupted the predicted secondary structure failed to significantly increase expression (Section 3.4). Nevertheless, wider mutagenesis of the 5'UTR did provide evidence that this area has several *cis*-acting regulatory elements, that when disrupted caused either increased or decreased expression. This reveals an exciting area of study for further understanding of *rpoBC* control; of particular interest is the potential interaction of this region with RT products from *rpoBp1* (see below).

rpoBC expression is affected by reiterative transcription events

In vitro transcription assays, as well as reads extracted from RNA-sequencing data, revealed that RT occurs at the TSS for *rpoBp1*. RT, in this case occurs on a run of A template nucleotides at the TSS for *rpoBp1*, causing the repeated addition of U nucleotides at the 5' end of the transcript (Sections 4.2 and 4.3). In fact, the majority of *rpoB* transcripts initiating at this promoter were subjected to at least one slippage event (>50% in all organisms studied). RT has been well studied in other organisms where it is known to control transcription based on surrounding nucleotide concentrations, including those associated with both the nascent RNA tract and the following nucleotide, which in the case of *rpoBp1* is UTP and GTP, respectively (Turnbough, 2008, 2011). Evidence for a RT response to [GTP] at *rpoBp1* was presented, with increasing levels of RT occurring at lower [GTP] (Fig 4.3.2.1). This supported our hypothesis that RT is caused by reduced [GTP] and that this might underlie the stringent control of *rpoBp1*. However, stringent control of *rpoBp1* appeared to be retained when +4 GTP was mutated to +4 CTP (Fig 4.3.7.1), which suggests alternative mechanisms or that downstream GTPs (e.g., +6 GTP) might also play a role in the commitment to *rpoBC* transcription. Increased [UTP] also had little effect on the transcription from *rpoBp1 in vitro*. This may seem unusual, however termination of transcription has not been accounted for in this assay, which is notably the most common resolution of UTP-dependent RT (Turnbough and Switzer, 2008). Preliminary evidence suggested that variation in [UTP] and [GTP] in *in vitro* experiments influenced the chosen initiating nucleotide, with high [GTP] combined with low [UTP] leading to the production of slightly smaller transcripts, presumably initiating at +4G. Initiation at +4G can also be detected in 5' mapping experiments highlighting the flexibility in initiation NTP choice. Such flexibility could contribute

to initiation control, especially if a common outcome of RT is transcription termination. For example, a drop in GTP would lead to greater rates of initiation on the T-tract and therefore greater rates of RT and possibly greater rates of termination. In support of this idea, nutrient limitation led to a decrease in +4 G initiation and a significant increase in >5T at the 5'end of *rpoBp1* transcripts. This type of regulatory mechanism is present for example in the control of *pyrC* and *pyrD* promoters in Salmonella, where changes in CTP/GTP ratio determine initiating nucleotide (+1C or +3G) (Sorensen *et al.*, 1993).

In vivo mutagenesis experiments using both transcriptional fusions and CRISPR-cas based mutations revealed that these slippage events at *rpoBp1* are detrimental to basal expression of *rpoBC* during exponential growth of *S. coelicolor* (Section 4.4). Upon removal of this T-tract, transcription of *rpoB* is observed to increase significantly, however the effect this on RNAP levels is yet to be determined. Nevertheless, this is the first report of *rpoBC* expression control by RT, and it would be interesting to investigate if this mechanism is implemented for these genes in other Actinobacteria, and among wider bacterial phyla.

How could slippage products interact with the 5'UTR to influence transcription?

One of the most striking results was that the T-tract was required for the 5'UTR to exert a negative effect on transcription. There are several examples where a 5' polynucleotide influences the formation of a structure such as an anti-terminator, and this influence transcription termination at an attenuator. Such a mechanism underlies the control of *pyrG* in *Bacillus subtilis* (Meng, Turnbough and Switzer, 2004). Preliminary analysis of the 5' extended RNA products did not reveal any obvious potential for secondary structure formation and any obvious mechanism to disrupt e.g. *rpoBa*, and so this mechanism seems unlikely. However, an alternative mechanism could be that slippage somehow pauses the RNAP and that this in turn influences downstream attenuation. This may also involve the integration of additional transcription factors such as NusA, which is already known to influence pause half-lives at RNA hairpins (Beuth *et al.*, 2005; Yakhnin *et al.*, 2019), however, such mechanisms warrant further investigation.

A wider role for slippage in transcription control

This is the first identified case of slippage in the *Streptomyces* genus, and one of the more striking results from this study was the identification of a further 28 genes subjected to RT control (Table 5.2.1). However, these RT events were only deduced from TSS-mapping data, thus characterising slippage during initiation only. Hence, the number of slippage prone genes is likely much larger in *Streptomyces* and therefore may be a great area of study in understanding nucleotide based control in this genus. Nevertheless, this study contributes to known examples of RT and suggests slippage-based transcriptional control is widespread within bacteria. In *S. agalactiae* for example, ~15% of identified TSS were subjected to some form of slippage, further displaying a vast regulatory role of RT (Rosinski-Chupin *et al.*, 2015).

Thus far, this project has shown that both this RT regulation at the *rpoBp1* TSS as well as element(s) within the 5'UTR of *rpoBC* have a negative impact on the production of these large subunits. It is likely these regulation mechanisms tightly control the expression of these genes in response to intracellular conditions, further determining the level of RNAP and thus the transcriptional capacity of the cell.

6.3 Future directions

Collectively this study has furthered the understanding of how *Streptomyces* controls the production of RNAP, however, has presented many more questions. In particular, is an increase in RNAP enzyme level able to bind to alternative promoters and activate the cryptic biosynthetic gene clusters? And does the increase in β and β' collectively increase the amount of RNAP available for transcription in the cell? In *E.coli* an increase in either or both of these subunits resulted in decreased translation of *rpoBC* transcripts, but had little effect on expression at the level of transcription (Dykxhoorn, St. Pierre and Linn, 1996). More recent work has also displayed a direct relationship between *rpoBC* expression and growth-rate in *E. coli*, with increased transcription of the large subunits resulting in increased growth, and vice versa (Izard *et al.*, 2015). Due to differences in transcriptional control outlined in this study, this mechanism in feedback control may differ in *Streptomyces*. This could be deduced using

overexpression constructs of the differing RNAP subunits to determine whether there is feedback control at the level of transcription and/or translation.

A key question for the future is whether *rpoBC* expression can influence the expression of biosynthetic gene clusters and thereby antibiotic production. It is possible that any autoregulatory control mechanism present, similar to that present in *E. coli*, would need to be disabled in order to influence RNAP levels significantly. Future investigation into the regulation of *rpoBC* and its influence on RNAP levels may reveal new ways to artificially alter RNAP levels and potentially allow the reactivation of silent clusters and the production/identification of novel antibiotics.

The novel identification of RT in *Streptomyces*, eliciting control of one of the most essential enzymes, RNAP, may provide evidence for the importance of nucleotide concentration for the control of gene expression on a global scale in this genus. Determination of homopolymeric tracts in *Streptomyces* genomes as well as the use of further RNA-sequencing data sets during differing growth phases, may provide insight into regulation elicited by RT in this genus (Baranov *et al.*, 2005; Rosinski-Chupin *et al.*, 2015; Imashimizu *et al.*, 2020).

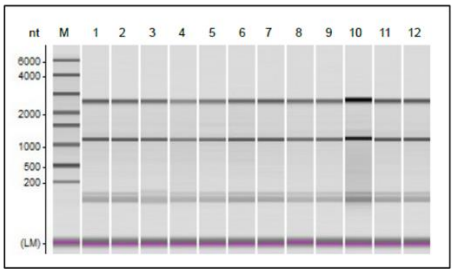
This study also touched on the potential importance of changes in nucleotide concentration in response to stress. In *Bacillus*, GTP and ATP have previously been shown to decrease and increase rapidly upon the induction of the stringent response resulting in the repression and activation at promoters where these are present as iNTPs (Kriel *et al.*, 2013). Similar changes in NTP concentrations were observed in *Streptomyces*, thus a similar mechanism of action may be apparent (Strauch *et al.*, 1991). Notably, *rrn* promoters, that initiate on a +1G nucleotide, were down-regulated as a result of amino acid starvation. As well as this, the current work also showed that RT increases as a result of this stress/decreased GTP levels. Global analysis of promoter sequences in *Streptomyces* and a focus on the response of genes to stringent conditions would provide insight into regulation during this stress. It may also be interesting to investigate changes in NTP levels during oxidative stress, however it has already been determined that ppGpp levels remained unchanged, hence nucleotide based control during this stress may be less likely (Paget and Hesketh, unpublished results; Kallifidas *et al.*, 2010).

Additionally, the involvement of attenuation mechanisms within the 5'UTR, that have been shown to interact with the slippage prone TSS to elicit negative effects on *rpoBC* expression, also suggests regulation here is likely much more complex. Much more research is necessary to identify and confirm both the presence of cis/trans acting elements within this downstream region, as well as the mechanisms by which RT influences attenuation here, during both exponential growth and stress. Hence, further investigations of nucleotide fluctuations and the RT mechanism as a whole in *Streptomyces*, may assist with the understanding of nucleotide based control in this genus, and would be beneficial to determine the full extent of *rpoBC* regulation.

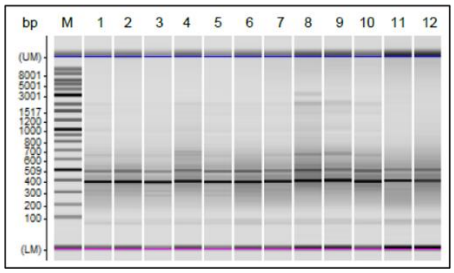
Appendices

7.1 Quality control results from Cappable-seq data from Vertis Biotechnology

A



B



D

No.	Sample	Conc. (ng/µl)	Amount (µg)	Conc. (ng/µl)	Amount (µg)	Ratio 23S/16S	Yield Cappable-seq enriched RNA (%)
		Customer-specified		Own measurements (see Fig.1)			
1	M145-3-0 mins	2106	21	779,6	14,8	1,3	0,1
2	M145-3-10 mins	1189	11	406,7	7,7	1,2	0,2
3	M145-4-0 mins	1396	13	519,1	9,9	1,2	0,3
4	M145-4-10 mins	2905	29	888,5	16,9	1,2	0,3
5	M145-5-0 mins	2136	21	614,5	11,7	1,3	0,3
6	M145-5-10 mins	465	6,9	399,3	5,6	1,2	0,6
7	M571-3-0 mins	3220	32	1131,5	21,5	1,3	0,2
8	M571-3-10 mins	2114	21	610,2	11,6	1,5	0,1
9	M571-4-0 mins	1640	16	502,5	9,5	1,4	0,2
10	M571-4-10 mins	1900	19	1113,9	21,2	1,7	0,1
11	M571-5-0 mins	3492	34	1295,9	24,6	1,3	0,1
12	M571-5-10 mins	2376	23	750,0	14,3	1,4	0,1

E

Total Reads	139.090.095
Cluster(PF)	128.128.968
% Cluster	92,1
Reads ident.	114.925.180
% Index	89,7

F

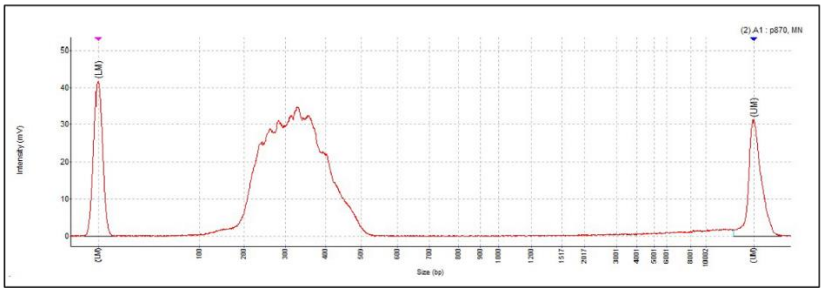
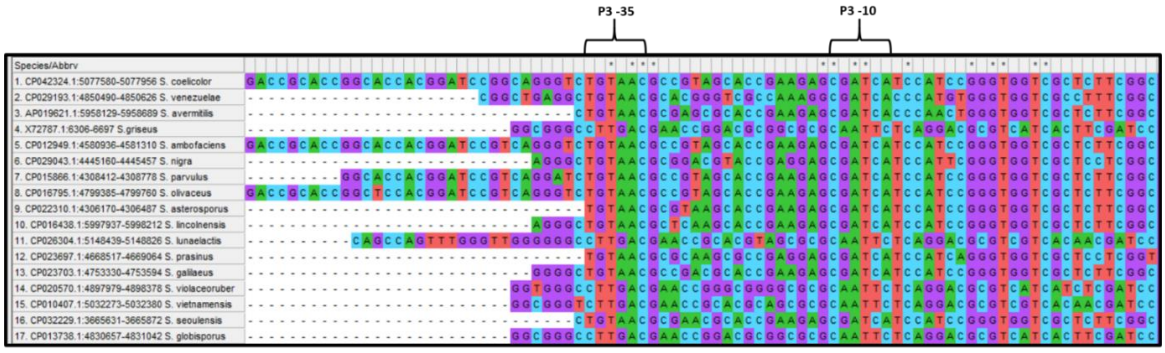
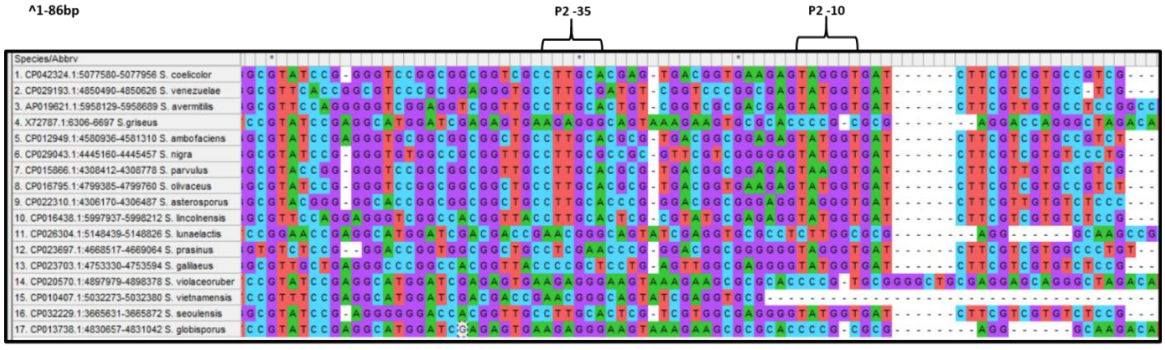


Figure 7.1.1: Quality control results for the *S. coelicolor* RNA samples used for cappable-seq by Vertis Biotechnology. A:Analysis of the total RNA samples on a Shimadzu MultiNA microchip electrophoresis system. M indicates the RNA marker used. B: Analysis of the PCR-amplified full-length cDNA on a Shimadzu MultiNA microchip electrophoresis system. M indicates the 100 bp ladder. C: Analysis of the PCR-amplified 5'fragment cDNAs on a Shimadzu MultiNA microchip electrophoresis system. M indicates the 100 bp ladder. D: A table of information of the 12 RNA samples from *S. coelicolor* used for cappable-seq. E: Summary of sequencing run. F: The NGS library pool analysed on a Shimadzu MultiNA microchip electrophoresis system.

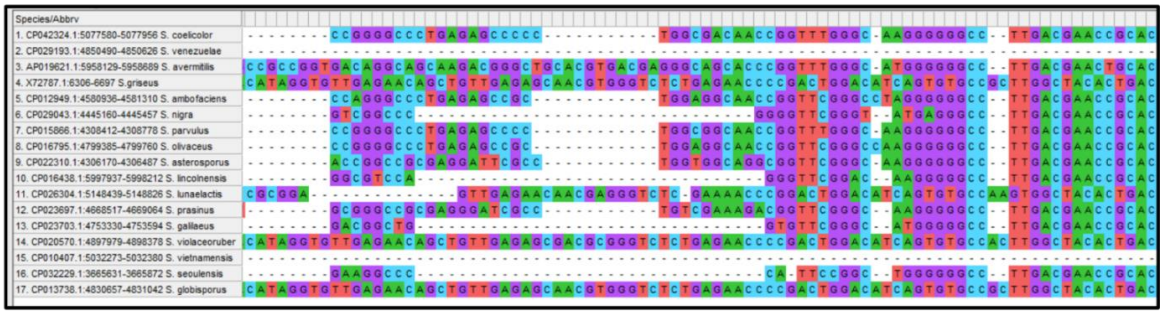
7.2 The full alignment of the *rpoBC* promoter region



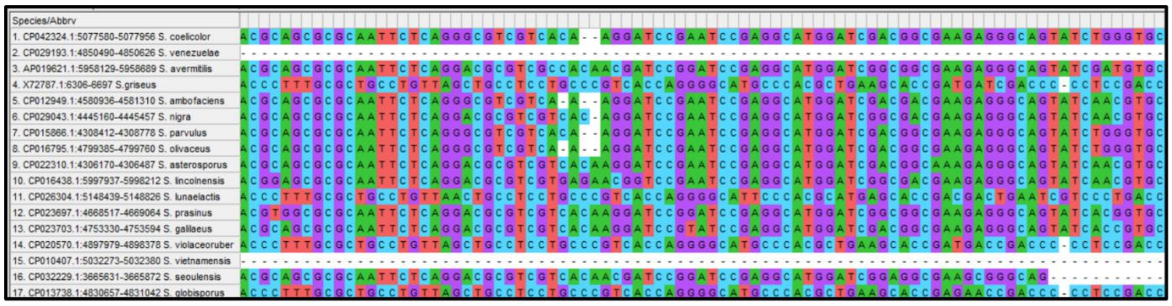
^1-86bp



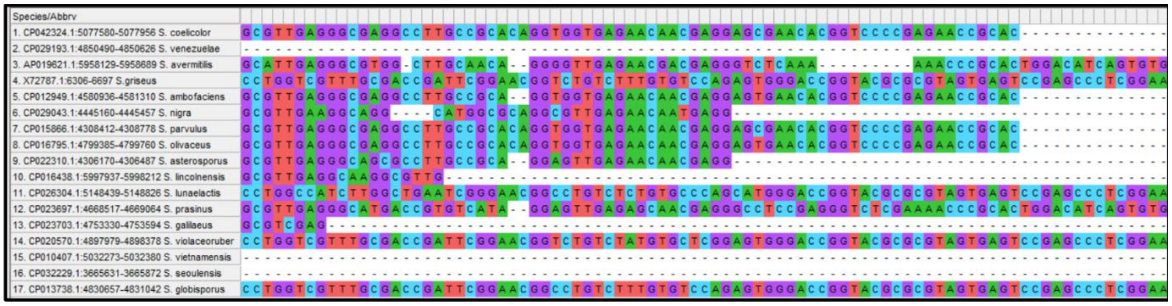
^84-168bp



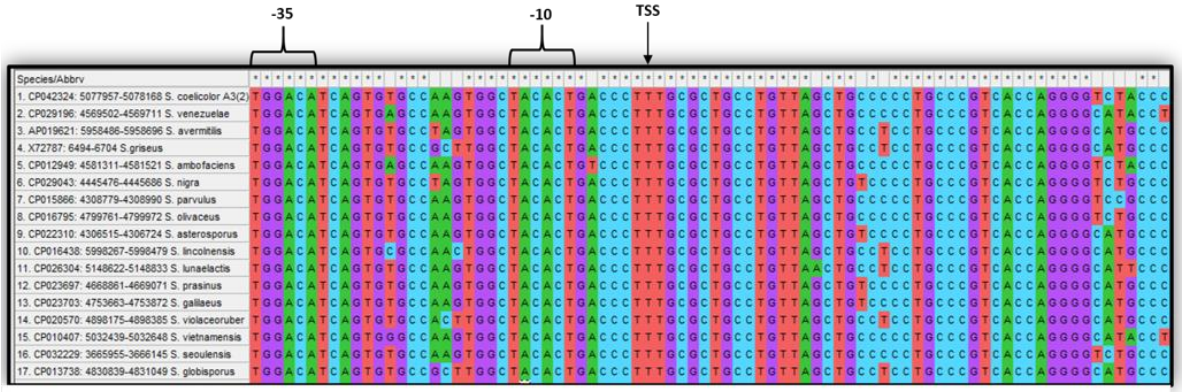
^166-250bp



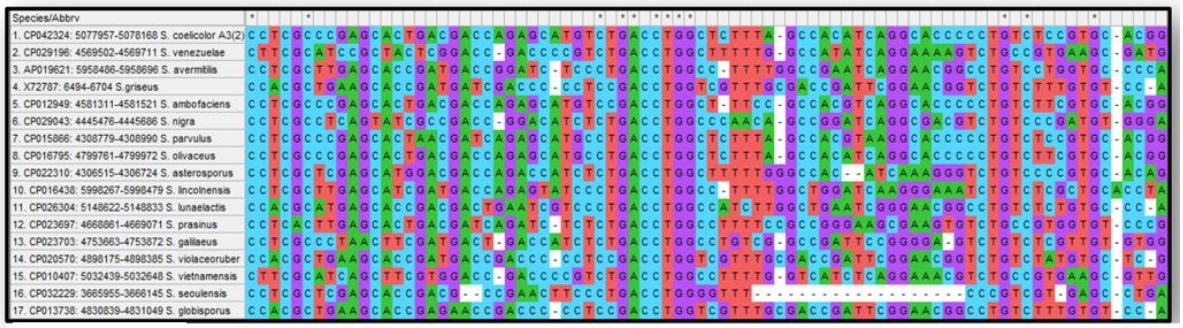
^248-332bp



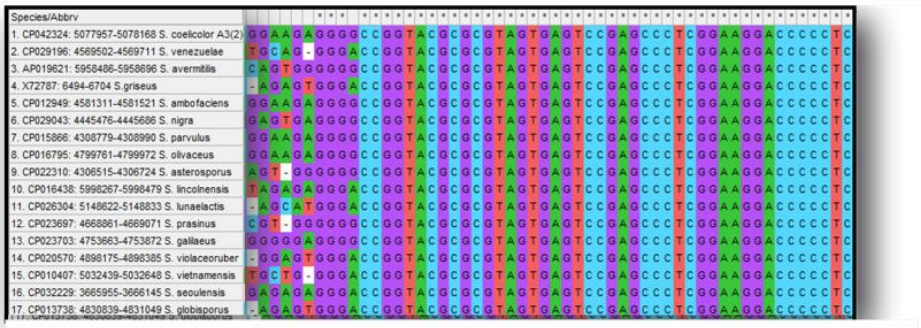
^330-414bp



415-497 bp



495-577 bp



575-629 bp

Figure 7.2.1: Alignment of the *rpoBC* promoter region. Seventeen organisms belonging to the *Streptomyces* genus were aligned to the previously identified *rpoBp1* promoter; *S. coelicolor, S. griseus, S. ambofaciens, S. olivaceus, S. parvulus, S. calvus, S. prasinus, S. violaceoruber, S. globisporus, S. venezuelae, S. avermitilis, S. nigra, S. lincolnensis, S.galilaeus, S. vietnamensis* and *S. seoulensis* were aligned using CLUSTALW and the MEGA-X programme, with organism accession numbers stated in Figure. The conserved -35 and -10 elements and TSS for all three identified promoters are indicated. For alignment to multiple *Streptomyces* organisms, the *S. coelicolor* sequence from 5077472 to 5081724 (directly before *rpoB* start codon) was used as a basis. The number of bases covered in individual alignment boxes are provided underneath.

7.3 Determination of stability of the luciferase enzyme

Due to a requirement of a rapidly dynamic system when monitoring changes in gene expression, the stability of the luciferase enzyme was determined, to confirm if this assay was suitable to detect rapid changes in transcription. The luciferase based assay relies on the expression and translation of the *luxAB* genes, with expression itself dependent on the promoter placed in control of the downstream genes. This promoter activity determines the amount of active luciferase enzyme produced which can then be quantified by the amount of light produced (in RLU), after the addition of the substrate, N-decanal, to cultures containing the pIJ5972 constructs. However, due to the reliance of this assay on the level of protein, rather than sole transcript level, the stability of this luciferase enzyme needed to be determined.

The decay of the activity of the luciferase proteins (*LuxAB*) was determined by the addition of hygromycin into cultures, causing inhibition of translation initiation. The pIJ5972:: *rpoBp1_ACT* construct was used for this optimisation with cultures pre-germinated and grown overnight in duplicate, in 60 ml of NMMP to an OD450 of ~0.8-1.0 (as seen in Section 2.2.9). Cultures were then split into two, with hygromycin added to one of the duplicate cultures at a final concentration of 50 µg/ml. Flasks were then returned to growth with RLU readings taken for the cultures in triplicate, using the conditions stated above 5, 15, 30, 45, 60, 90, 120, 150, 210, and 270 min after the initial addition of hygromycin to cultures.

Results from the duplicate growth were averaged and plotted to determine the potential stability of the luciferase enzyme (Fig 7.3.1.1). A clear decrease in the luciferase activity compared to cultures where hygromycin wasn't added was seen, indicating

the antibiotic was working correctly, with the continued increase of average RLU value as the cultures were left growing. It should be noted that the RLU value of the cultures containing hygromycin still had apparent residual activity, especially when compared to the M145 parent strain (data not shown) with reads still seen to be ~5000000 RLU. This may suggest the hygromycin concentration used was not high enough, however an obvious decrease was still observed. As well as this, compared to other results collected, the results seen in Fig 7.3.1.1 were RLU values only, normalisation was not carried out using OD450 or protein normalisation, due to the effects hygromycin may have on these readings themselves due to the inhibition of protein synthesis.

Initial results suggest that hygromycin doesn't effect luciferase activity till at least 30 min after its addition to cultures, with the RLU value observed to drop around 2-fold after ~60 min.

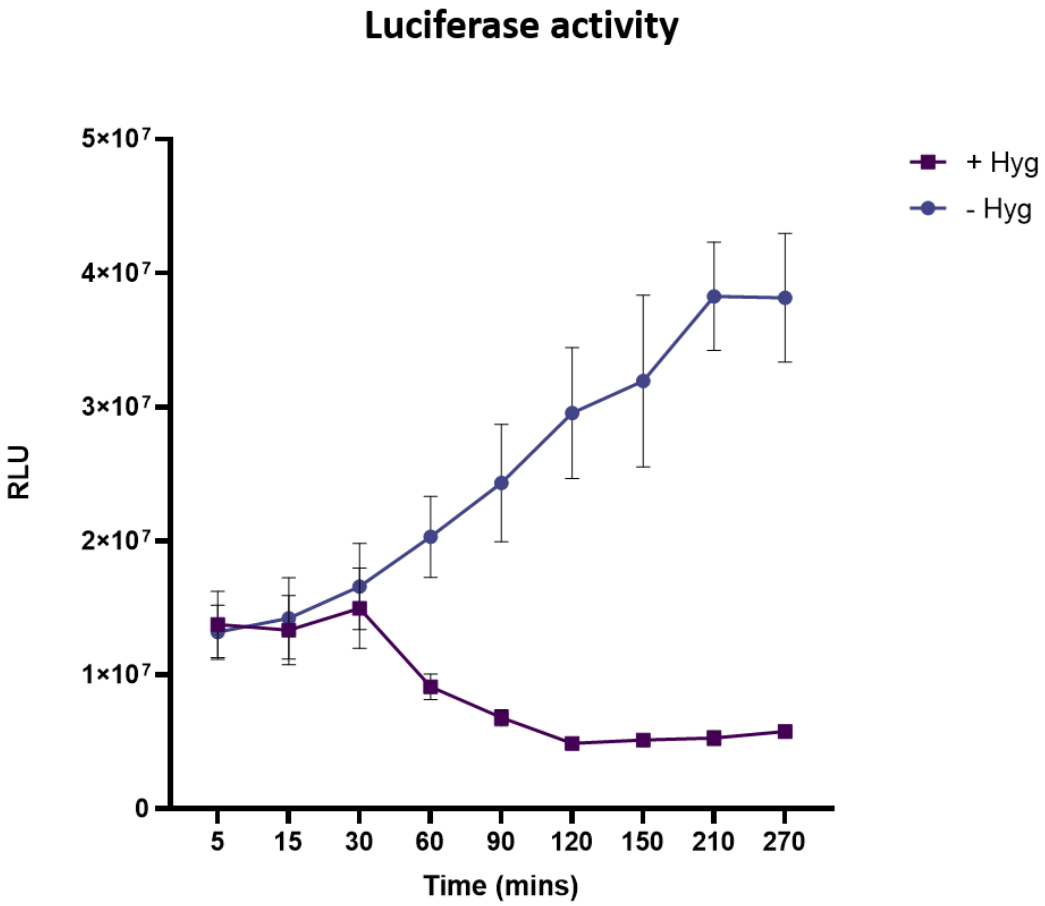


Figure 7.3.1.1: Determination of the stability of the luciferase enzyme. M145 strains containing the pIJ5972::*rpoBp1U* ACT plasmid were grown in biological triplicate to exponential phase before the addition of hygromycin to cultures and continued growth. Samples were collected 5, 15, 30, 45, 60, 90, 120, 150, 210, and 270 min after the addition of hygromycin. The graph displays the average RLU observed for strains both with (purple) and without (blue) the addition of hygromycin (50 μg/ml). RLU readings were taken in technical triplicate, and averaged. Error bars represent the S.E.M (standard error of the mean).

Growth curve of M145 and strains containing pIJ5972 reporter plasmids

To confirm the presence of the pIJ5972 constructs have no effect on the growth of the strain in comparison to the M145 stain alone, several growth curves were performed using M145 strains both the pIJ5972 empty vector, as well as the *pIJ5972:rpoBp* constructs *rpoBp1U* WT and ACT, and the M145 parent strain. All strains were inoculated at a starting OD450 of 0.06 in duplicate in 50 ml NMMP medium supplemented with amino acid (see section). Strains were left to grow overnight with OD450 readings then taken at 15.5, 18.5, 20.5, 22.5, 24.5, 42 and 48h after initial inoculation.

Figure 7.3.1.2 displays the OD450 readings and growth curves for the 4 differing strains grown in duplicate. From this, the presence of the pIJ5972 stably integrated plasmid, was confirmed to have no effect on the growth rate of the differing strains, when compared to the M145 parent strain.

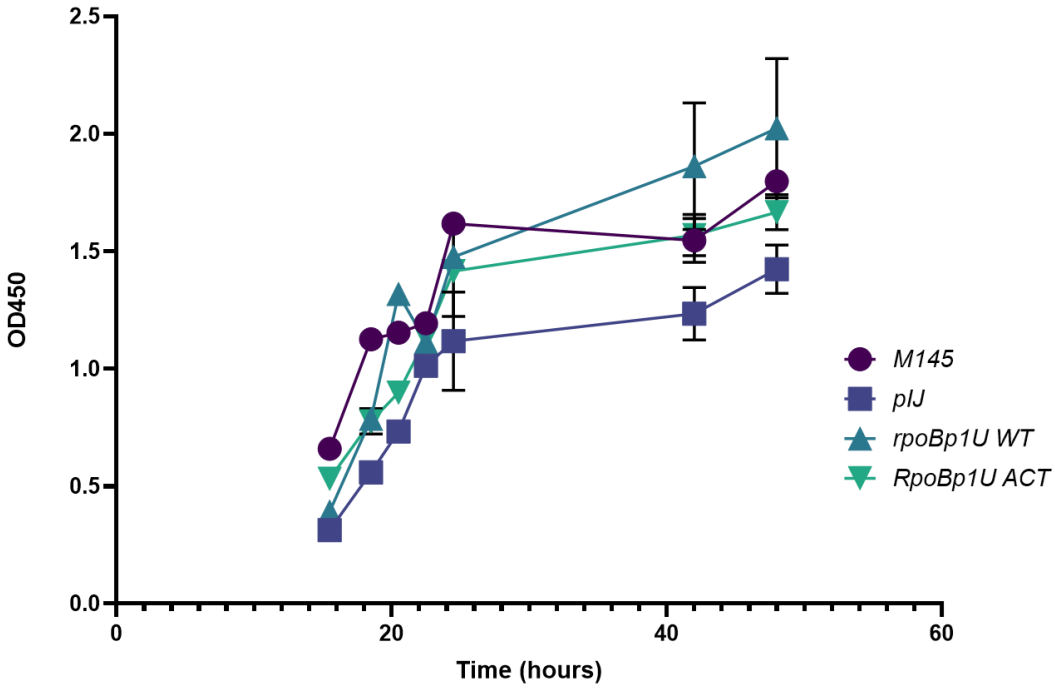


Figure 7.3.1.2: The growth curves of M145 alone, and pIJ5972 containing strains. Strains were grown in biological duplicate with OD450 readings taken after 15.5, 18.5, 20.5, 22.5, 24.5, 42 and 48h of growth. Error bars represent the S.E.M of the duplicate values. pIJ represents the empty pIJ5972 vector alone (in M145). M145 strains containing both the *rpoBp1U* constructs containing either the wildtype TSS (WT) or the ACT mutated TSS were also grown alongside M145 and the pIJ5972 empty vector control.

7.4 Reiterative transcription at *rpoBp1* after the induction of the stringent response

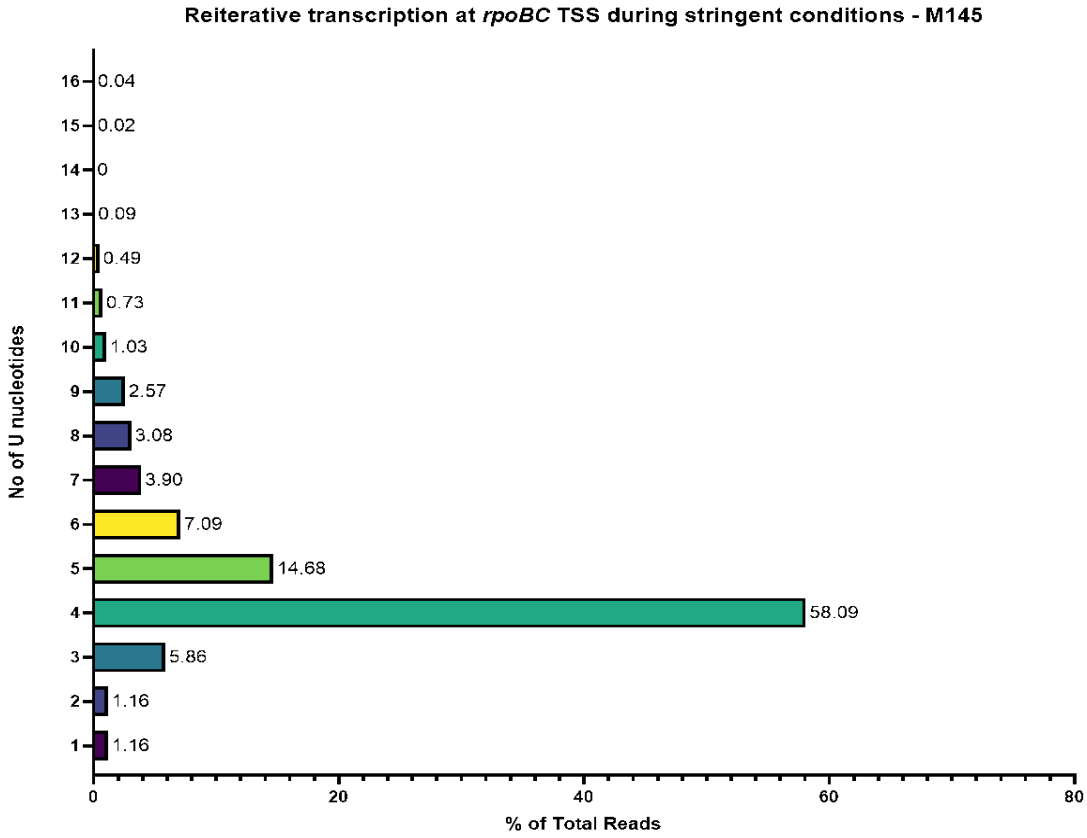


Figure 7.4.1: Reiterative transcription at *rpoBp1* after the induction of the stringent response in M145. Transcripts (5328) from the raw TSS-mapping data (Ettwiller *et al*., 2016), taken 10 min after the induction of a stringent response in M145 (Section 2.2.3.4), were extracted using the galaxy platform (Afgan *et al.*, 2018) and an intext (grep) search with 30bp of *rpoBp* sequence. Transcripts are organised according to the number of U's at the 5' end (TSS) of the transcript, as a percentage of total reads analysed.

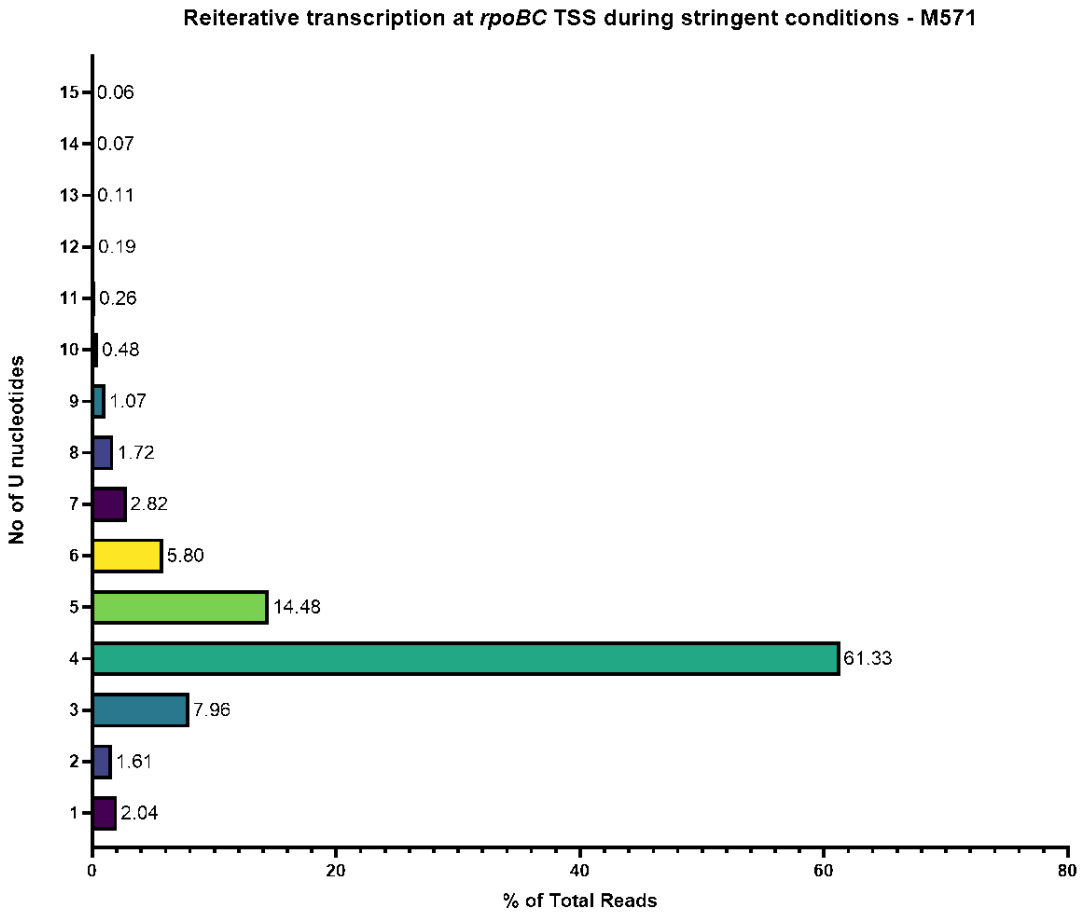


Figure 7.4.2: Reiterative transcription at *rpoBp1* after the induction of the stringent response in M571. Transcripts (5399) from the raw TSS-mapping data (Ettwiller *et al.*, 2016), taken 10 min after the induction of a stringent response in M145 (Section 2.2.3.4), were extracted using the galaxy platform (Afgan *et al.*, 2018) and an intext (grep) search with 30bp of *rpoBp* sequence. Transcripts are organised according to the number of U's at the 5' end (TSS) of the transcript, as a percentage of total reads analysed.

7.5 Identification of reiterative transcription in *S. venezuelae*

Table 7.5.1: Reiterative transcription in *S. venezuelae*. Reads from *S. venezuelae* (394939) were extracted using the intext (grep) search on galaxy, and used to identify areas of the genome subjected to slippage. StrepDB and BLASTn was used to identify locus of DNA sequence, as well as confirming additional nucleotides encoded in transcripts. Gene name and number, sequence (30bp), including tract sensitive to RT (bold), strand (+/-), location in respect to TSS and distance from start codon are provided for each example. The corresponding sequence was extracted from the complete *S. venezuelae* genome on StrepDB (NCBI, NZ_CP018074.1).

No	Gene	Sequence (5' -> 3')	Strand	Location of RT	Distance from Start Codon (bp)
ATP slippage					
1	**DNA-binding protein** (SVEN15_3754)	**AAA**GAACCGAGAAGGTTCGGTTCTCCCGAG	+	Slippage on TSS	40
2	***cydA* - Cytochrome D ubiquinol oxidase** (SVEN15_3623)	**AAA**CTGCCGATGTGACCACGGCGATGTAAG	+	Slippage on TSS	41
GTP slippage					
3	**Membrane protein** (SVEN15_4524)	**GGGGG**AGTGAAAATTCCGCCACCCCCGGCT	+	Slippage on TSS	45
4	**Topoisomerase IV subunit B** (SVEN15_5392)	**GGGGG**CTCGCGGGGGTCAGAACAGTAGTCA	+	Slippage on TSS	452
5	**DNA binding protein** (DskA homology) (SVEN15_1685)	**GGGG**TCGACGACCGTCTCGTAGCCCTGGGC	-	Slippage on TSS	1166
6	**Alkanesulfonates transport system** (SVEN15_0745)	**GGGGG**TGGTTGCGGTGACGGCCATGGTGCT	+	Slippage on TSS	2381
7	**FtsK/SpoIII family protein** (SVEN15_5273)	AT**GGGGG**AAGTCCGTACGCCCGGCTGCCGG	-	Slippage 2 bp after TSS	206
8	**Endoribonuclease** (SVEN15_6040)	AA**GGGG**CCTACCTGCAACTACGGCCCCGCG	+	Slippage 2 bp after TSS	50
9	**Transport protein** (SVEN15_6902)	A**AGGGGG**TGTCGTGGCTGTAGACGAGCGCG	+	Slippage 2 bp after TSS	1821

No	Gene	Sequence (5’ -> 3’)	Strand	Location of RT	Distance from Start Codon (bp)
10	**Electron transfer flavoprotein beta subunit** (SVEN15_0690)	**GGGG**CGTCCGGCGCCCCGACCGGAAACCCG	-	Slippage on TSS	97
11	**Membrane protein** (SVEN15_2591)	**GGGGG**CGGTGTCGCCGGGCTGCAGTCGCTG	-	Slippage on TSS	162
12	**Threonine 3-dehydrogenase** (SVEN15_0283)	A**GGG**ATCTTGGGCGCGCACTCGTGCCCGCA	-	Slippage 1 bp after TSS	69

7.6 Templates used for *in vitro* transcription assays

```
        p2 -10
5′ TAGGGTGATCTTCGTCGTGCCGTCGCCGGGGCCCTGAGAGCCCCCTGGCGACAACCGGTTTG

   GGCAAGGGGGGCCTTGACGAACCGCACGCAGCGCGCAATTCTCAGGGCGTCGTCACAAGGAT

   CCGAATCCGAGGCATGGATCGACGGCGAAGAGGGCAGTATCTGGGTGCGTTGAGGGCGAGGC
                                                          p1 -35
   CTTGCCGCACAGGTGGTGAGAACAACGAGGAGCGAACACGGTCCCCGAGAACCGCACTGGACA
                    p1 -10
   TCAGTGTGCCAAGTGGCTACACTGACCCTTTGCGCTGCCTGTTAGCTGCCCCCTGCCCGTCACC

   AGGGGTCTAC 3′
```

Figure 7.6.1: The *rpoBp1*(1) DNA template used for *in vitro* transcription assays. The forward (green) and reverse (yellow) primers are indicated, with this region amplified using primers 77 and 78 respectively. A 44 nt run-off transcript was produced from this template. The p2 and p1 promoters for *rpoBC*, and the TSS (bold, TTT) are indicated.

```
5′ GGAATTCGTGATCTTCGTCGTGCCGTCGCCGGGGCCCTGAGAGCCCCCTGGCGACAACCG

   GTTTGGGCAAGGGGGGCCTTGACGAACCGCACGCAGCGCGCAATTCTCAGGGCGTCGTCA

   CAAGGATCCGAATCCGAGGCATGGATCGACGGCGAAGAGGGCAGTATCTGGGTGCGTTGA

   GGGCGAGGCCTTGCCGCACAGGTGGTGAGAACAACGAGGAGCGAACACGGTCCCCGAGAA
          p1 -35                 p1 -10
   CCGCACTGGACATCAGTGTGCCAAGTGGCTACACTGACCCTTTGCGCTGCCAGATCTGGG

   GGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGTTC 3′
```

Figure 7.6.2: The *rpoBp1*(2) DNA template used for *in vitro* transcription assays. Forward (green) and reverse (yellow) primers are indicated. The template was amplified from the pBluescript plasmid containing the *p1* region (Fig 3.2.3.1) using primers 78 and 79 respectively. A run of transcript length of ~74 nt was produced from this template. The p1 promoter and TSS (TTT, bold) for *rpoBC* are indicated.

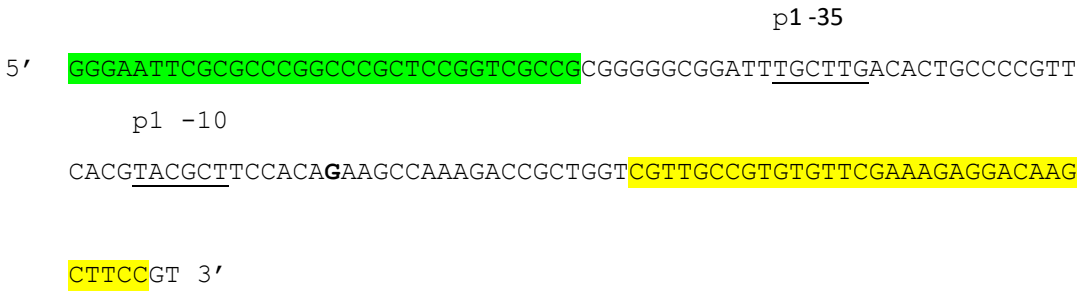


Figure 7.6.3: The *rpIJ* DNA template used for *in vitro* transcription assays. Forward (green) and reverse (yellow) primers are indicated, with the template amplified using primers 80 and 81 respectively. A 54 nt run-off transcript was produced from this template. The promoter elements and TSS (bold) for the SCO4652 gene are indicated.

7.7 RNA secondary structure for *rpoBp* and 5'UTR region

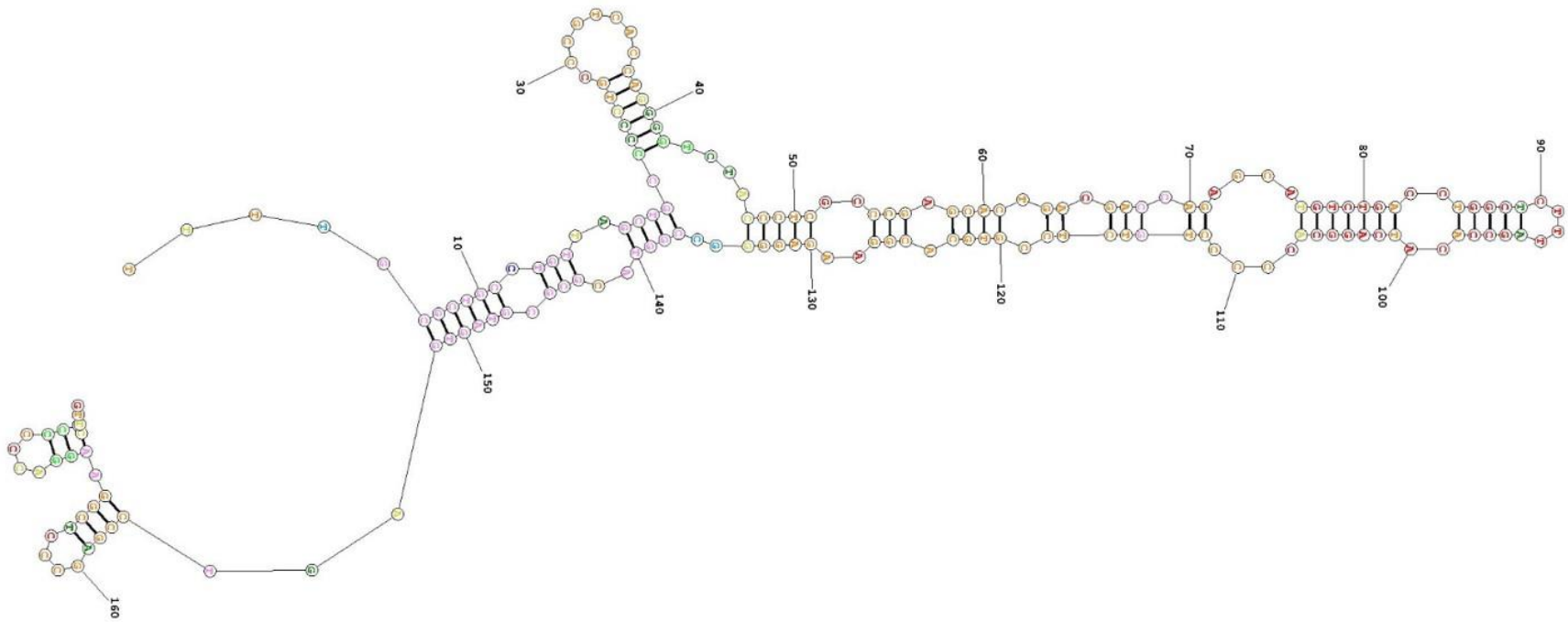


Figure 7.7: Hypothesised RNA secondary structure for the 5'UTR including RBS for *rpoB*. The RNA sequence spanning from the TSS down to 10nt after the RBS was analysed using secondary structure prediction (+1 - +178 nt) (Mathews, 2014); 4 U nucleotides were included at the 5' end of the transcript representing the sequence of the majority of *rpoB* transcripts. Default fold options were used; maximum of 10% energy difference and temperature of 310.15K. Probabilities for pairing between bases are provided as colours; red = >99%, orange = >95%, yellow = >90%, dark green = >80%, light green = >70%, light blue = >60%, dark blue = >50%, pink = <50%. The predicted folding free energy of the above structure is -73.6 kcal/mol.

7.8 Regulation of growth transitions in *Streptomyces*

The full cycle of *Streptomyces* growth using spans over 3-10 days, with various proteins contributing to this development (Fig 1.1.3.2.1). Initially, suitable environmental conditions cause the germination of dormant *Streptomyces* spores, before the production of 1 or 2 germ tubes to extend from the spore further growing by hyphal top extension producing the vegetative mycelium (Fig1.1.3.2C). This method of cell division largely differs from most bacteria, such as *E. coli*, where the MreB protein, a bacterial actin homologue, is essential for the formation of a helical cytoskeleton and assembly of peptidoglycan to the lateral cell wall, required for elongation and formation of the rod shape (Carballido-López, 2006; Flärdh and Buttner, 2009). The genes encoding MreB are not present in *S. coelicolor* and most Actinobacteria, however a coiled coil protein, DivIVA, is essential for tip extension and branching of *Streptomyces* (Flärdh, 2003; Richards *et al.*, 2012). In particular, DivIVA localises to cell poles at growing hyphal tips and assists with the recruitment of proteins necessary for cell wall biosynthesis, facilitating the asymmetrical growth of *Streptomyces* (Flärdh and Buttner, 2009). The importance of this protein was proven using both a partial deletion and overexpression of DivIVA, resulting in defective hyphal growth and hyper-branching of the mycelium, respectively (Flärdh, 2003; Hempel *et al.*, 2008).

After a decrease in the surrounding nutrients available, aerial hyphae are produced, extending upwards away from the vegetative mycelium, which is notable for its fuzzy appearance caused by the assembly of a hydrophobic sheath as a result of SapB (Spore-associated protein B) production. SapB which is produced by genes encoded in the *ramCSAB* operon, is a Lantibiotic-like surfactant peptide that reduces surface tension facilitating this growth into the air (Kodani *et al.*, 2004). Notably SapB is only produced on rich media, yet *S. coelicolor* is still able to produce aerial mycelium on minimal medium, suggesting a SapB-independent pathway exists for the formation of this mycelium (Capstick *et al.*, 2007). Alternative aerial mycelial production is resultant from the chaplins and rodlins that made up the hydrophobic sheath. There are 8 secreted proteins belonging to the chaplin family, ChpA-H, that all share a conserved hydrophobic domain, and overlap with SapB in function, enabling the production of aerial hyphae (chaplin domain) (Claessen *et al.*, 2003). Rodlins organise

the structures of these chaplins, with the chaplins themselves assembling at the air-water interface at the surface, decreasing the surface tension (Flärdh and Buttner, 2009). Strains that are unable to produce both chaplins and SapB, or SapB alone are bald in phenotype in all growth conditions, however, the overexpression of *ramR*, the activator of *SapB* expression, is sufficient to complement and restore wild type phenotype from other *bld* mutants, providing evidence for the importance of sapB in aerial hyphae formation (Willey *et al*., 1991, Nguyen *et al.,* 2002).

The *bld* genes, as discussed in Section 1.1.3.2 are also required for the formation of aerial hyphae, which initially grow to form a multinucleoidal cell containing >50 copies of the genome, before growth arrest and septation forming uninucleoidal pre-spores, before a final transition into mature spores characterised by their grey pigmentation (Schwedock *et al.,* 1997). This septation is controlled and initiated by the bacterial tubulin homologue FtsZ, that is able to assemble into helical filaments forming a ladder of rings at the sites of division forming unigenomic pre-spore compartments (Schwedock *et al.*, 1997; Van Wezel and McDowall, 2011). Mutations of *ftsZ* are not lethal to the strain, due to a continued ability grow aerial hyphae, however, the further transition of this hyphae to spores is prevented (McCormick *et al.*, 1994). As well as this the ParAB and FtsK proteins control the final portioning into uninucleoidal spores. *ParAB* encodes a cytoskeletal ATPase (*ParA*) and a DNA-binding protein (*ParB*) that binds to parS (partitioning) sites, near the oriC in the DNA, forming a compact nucleoprotein that further binds to ParA and stimulates ATP hydrolysis (Kim *et al.*, 2000; Jakimowicz, Chater and Zakrzewska-Czerwínska, 2002). Mutations of these ParAB genes however does not affect the growth phenotype of *S. coelicolor* (Jakimowicz *et al.*, 2006). *FtsK* encodes a DNA translocase that is localised to sporulation septa, and thought to pump through DNA to prevent trapping of chromosomal DNA in the septum. This prevents further consequences, such as deletion of DNA at the ends of the chromosomes, however FtsK itself is not required for septation (Flärdh and Buttner, 2009).

The process of sporulation itself also require several proteins, identified as the *whi* (white) genes (Section 1.1.3.2). These include *whiA*, *whiB*, *whiD*, *whiE*, *whiG*, *whiH*, *wihL* and *whiJ*, that are all required for spore formation (Schwedock *et al.*, 1997). These can then be divided into families of genes that affect the initial formation of pre-spore compartments (early-*whi* genes), or affect the final division or production of the

grey spore pigment (late-*whi* genes). For example, the *whiG* gene, which encodes a sigma factor, further regulates the expression of the *whiH* gene, and is the first and likely most important step for the regulation of sporulation events, and thus is identified as an early-*whi* gene (Chater *et al.*, 1989). The final step of *S. coelicolor* growth is the maturation of spores that relies on genes within the *mre* cluster, which contains the *mreB* gene required for the correct assembly of the spore wall (Wildermuth and Hopwood, 1970; Sigle *et al.*, 2015). During maturation, both a thicker and lysozyme resistant cell wall is formed, before the rounding of the pre-spores into an ovoid shape (Wildermuth and Hopwood, 1970).

References

Bioinformatic references

BLAST - https://blast.ncbi.nlm.nih.gov/Blast.cgi

Galaxy server – https://usegalaxy.org/

ImageJ Gel band quantification obtained from - http://rsbweb.nih.gov/ij/index.html

RNA folding algorithm -

https://rna.urmc.rochester.edu/RNAstructureWeb/Servers/Predict1/Predict1.html

StrepDB - http://strepdb.streptomyces.org.uk/

References

Aceti, D. J. and Champness, W. C. (1998) 'Transcriptional regulation of Streptomyces coelicolor pathway-specific antibiotic regulators by the absA and absB loci', *Journal of Bacteriology*, 180(12), pp. 3100–3106. doi: 10.1128/jb.180.12.3100-3106.1998.

Afgan, E. *et al.* (2018) 'The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update', *Nucleic Acids Research*, 46(W1), pp. W537–W544. doi: 10.1093/nar/gky379.

Aigle, B. *et al.* (2000) 'A single amino acid substitution in region 1.2 of the principal ?? factor of Streptomyces coelicolor A3(2) results in pleiotropic loss of antibiotic production', *Molecular Microbiology*, 37(5), pp. 995–1004. doi: 10.1046/j.1365-2958.2000.02022.x.

Allen, M. S. *et al.* (2007) 'A destabilized bacterial luciferase for dynamic gene expression studies', *Systems and Synthetic Biology*, 1(1), pp. 3–9. doi: 10.1007/s11693-006-9001-5.

Anikin, M. *et al.* (2010) 'Recoding: Expansion of Decoding Rules Enriches Gene Expression', 24(January 2010). doi: 10.1007/978-0-387-89382-2.

Archambault, J. and Friesen, J. D. (1993) 'Genetics of eukaryotic RNA polymerases I, II, and III', *Microbiological Reviews*, 57(3), pp. 703–724. doi: 10.1128/mmbr.57.3.703-724.1993.

Arias, P., Fernández-Moreno, M. A. and Malpartida, F. (1999) 'Characterization of the pathway-specific positive transcriptional regulator for actinorhodin biosynthesis in Streptomyces coelicolor A3(2) as a DNA-binding protein', *Journal of Bacteriology*, 181(22), pp. 6958–6968. doi: 10.1128/jb.181.22.6958-6968.1999.

Arnvig, K. B. *et al.* (2004) 'A high-affinity interaction between NusA and the rrn nut site in Mycobacterium tuberculosis', *Proceedings of the National Academy of Sciences of the United States of America*, 101(22), pp. 8325–8330. doi: 10.1073/pnas.0401287101.

Artsimovitch, I. and Landick, R. (2000) 'Pausing by bacterial RNA polymerase is mediated by mechanistically distinct classes of signals', *Proceedings of the National Academy of Sciences of the United States of America*, 97(13), pp. 7090–7095. doi: 10.1073/pnas.97.13.7090.

Artsimovitch, I. and Landick, R. (2002) 'The transcriptional regulator RfaH stimulates RNA chain synthesis after recruitment to elongation complexes by the exposed nontemplate DNA strand', *Cell*, 109(2), pp. 193–203. doi: 10.1016/S0092-8674(02)00724-9.

Atkinson, G. C., Tenson, T. and Hauryliuk, V. (2011) 'The RelA/SpoT Homolog (RSH) superfamily: Distribution and functional evolution of ppgpp synthetases and hydrolases across the tree of life', *PLoS ONE*, 6(8). doi: 10.1371/journal.pone.0023479.

Atling-Mees, M. . and Short, J. . (1989) 'pBluescript II: gene mapping vectors', 17(22), p. 9494.

Babcock, M. J. *et al.* (1997) 'Characterization of the rpoC gene of Streptomyces

coelicolor A3(2) and its use to develop a simple and rapid method for the purification of RNA polymerase', *Gene*, 196(1–2), pp. 31–42. doi: 10.1016/S0378-1119(97)00179-0.

Bae, B. *et al.* (2013) 'Phage T7 Gp2 inhibition of Escherichia coli RNA polymerase involves misappropriation of σ70 domain 1.1', *Proceedings of the National Academy of Sciences of the United States of America*, 110(49), pp. 19772–19777. doi: 10.1073/pnas.1314576110.

Bailey, M. J. A., Hughes, C. and Koronakis, V. (1996) 'Increased distal gene transcription by the elongation factor RfaH, a specialized homologue of NusG', *Molecular Microbiology*, 22(4), pp. 729–737. doi: 10.1046/j.1365-2958.1996.d01-1726.x.

Baranov, P. V. *et al.* (2005) 'Transcriptional slippage in bacteria: distribution in sequenced genomes and utilization in IS element gene expression.', *Genome biology*, 6(3). doi: 10.1186/gb-2005-6-3-r25.

Barka, E. A. *et al.* (2016) 'Taxonomy, Physiology, and Natural Products of Actinobacteria', *Microbiology and Molecular Biology Reviews*, 80(1), pp. 1–43. doi: 10.1128/MMBR.00019-15.

Barker, M. M. *et al.* (2001) 'Mechanism of regulation of transcription initiation by ppGpp. I. Effects of ppGpp on transcription initiation in vivo and in vitro', *Journal of Molecular Biology*, 305(4), pp. 673–688. doi: 10.1006/jmbi.2000.4327.

Barne, K. A., Bown, J. A. and Minchin, S. D. (1997) 'Region 2.5 of the Escherichia coli RNA polymerase σ70 subunit is responsible for the recognition of the "extended –10" motif at promoters', 16(13), pp. 4034–4040.

Barry, G, S. C. and S. C. L. (1985) 'Attenuation and processing of RNA from the rpIJL-rpoBC transcription unit of Escherichia coli', *Nucleic Acids Research*, 13(20), pp. 7289–7297. doi: 10.1093/nar/13.20.7289.

Barry, G., Squires, C. L. and Squires, C. (1979) 'Control features within the rplJL-rpoBC transcription unit of Escherichia coli', *Proceedings of the National Academy of Sciences of the United States of America*, 76(10), pp. 4922–4926. doi: 10.1073/pnas.76.10.4922.

Behie, S. W. *et al.* (2017) 'Molecules to ecosystems: Actinomycete natural products in situ', *Frontiers in Microbiology*, 7(JAN), pp. 1–11. doi: 10.3389/fmicb.2016.02149.

Belas, R. *et al.* (1982) 'Bacterial Bioluminescence : Isolation and Expression of the Luciferase Genes from Vibro harveyi', *Science*, 218(4574), pp. 791–793.

Belogurov, G. A. *et al.* (2007) 'Structural Basis for Converting a General Transcription Factor into an Operon-Specific Virulence Regulator', *Molecular Cell*, 26(1), pp. 117–129. doi: 10.1016/j.molcel.2007.02.021.

Bennett, P. M., Grinsted, J. and Richmond, M. H. (1977) 'Transposition of TnA does not generate deletions', *MGG Molecular & General Genetics*, 154(2), pp. 205–211. doi: 10.1007/BF00330839.

Bentley, S. *et al.* (2002) 'Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2).', *Nature*, 417(6885), pp. 141–147. doi:

10.1038/417141a.

Beuth, B. *et al.* (2005) ‘Structure of a Mycobacterium tuberculosis NusA-RNA complex’, *EMBO Journal*, 24(20), pp. 3576–3587. doi: 10.1038/sj.emboj.7600829.

Bibb, M. J. *et al.* (2012) ‘Expression of the chaplin and rodlin hydrophobic sheath proteins in Streptomyces venezuelae is controlled by σ BldN and a cognate anti-sigma factor, RsbN’, *Molecular Microbiology*, 84(6), pp. 1033–1049. doi: 10.1111/j.1365-2958.2012.08070.x.

Boor, K. J., Duncan, M. L. and Price, C. W. (1995) ‘Genetic and transcriptional organization of the region encoding the β subunit of Bacillus subtilis RNA polymerase’, *Journal of Biological Chemistry*, 270(35), pp. 20329–20336. doi: 10.1074/jbc.270.35.20329.

Borukhov, S. *et al.* (1992) ‘GreA protein: a transcription elongation factor from Escherichia coli.’, *Proceedings of the National Academy of Sciences of the United States of America*, 89(19), pp. 8899–902. doi: 10.1073/pnas.89.19.8899.

Boyaci, H. *et al.* (2018) ‘Fidaxomicin jams mycobacterium tuberculosis RNA polymerase motions needed for initiation via RBPA contacts’, *eLife*, 7, pp. 1–19. doi: 10.7554/eLife.34823.

Bralley, P. *et al.* (2006) ‘RNA 3′-tail synthesis in Streptomyces: In vitro and in vivo activities of RNase PH, the SCO3896 gene product and polynucleotide phosphorylase’, *Microbiology*, 152(3), pp. 627–636. doi: 10.1099/mic.0.28363-0.

Breaker, R. R. (2012) ‘Riboswitches and the RNA World’, *Cold Spring Harbour Perspectives in Biology*, 4(a003566).

Browning, D. F. and Busby, S. J. W. (2004) ‘The regulation of bacterial transcription initiation’, *Nature Reviews Microbiology*, 2(1), pp. 57–65. doi: 10.1038/nrmicro787.

Buck, M. and Cannon, W. (1992) ‘Specific binding of the transcription factor sigma-54 to promoter DNA’, 358(July), pp. 19–21.

Burgess, R. R. *et al.* (1969) ‘Factor stimulating transcription by RNA polymerase’, *Nature*, 221(5175), pp. 43–46. doi: 10.1038/221043a0.

Burgess, R. R. (1969) ‘Separation and characterization of the subunits of ribonucleic acid polymerase.’, *Journal of Biological Chemistry*. Â© 1969 ASBMB. Currently published by Elsevier Inc; originally published by American Society for Biochemistry and Molecular Biology., 244(22), pp. 6168–6176. doi: 10.1016/s0021-9258(18)63521-5.

Burmann, B. *et al.* (2011) ‘Domain interactions of the transcription:translation coupling factor E. coli NusG are intermolecular and transient.’, *J. Phys. Energy*, 435(3), pp. 783–789.

Burmann, B. M. *et al.* (2010) ‘A NusE:NusG complex links transcription and translation’, *Science*, 328(5977), pp. 501–504. doi: 10.1126/science.1184953.

Burmann, B. M. *et al.* (2012) ‘An α-helix to β-barrel domain switch transforms the transcription factor RfaH into a translationfactor’, *Cell*, 150(2), pp. 291–303. doi: 10.1038/jid.2014.371.

Busche, T. *et al.* (2016) 'Deciphering the transcriptional response mediated by the redox-sensing system HbpS-SenS-SenR from streptomycetes', *PLoS ONE*, 11(8). doi: 10.1371/journal.pone.0159873.

Bush, M. J. *et al.* (2013) 'Genes Required for Aerial Growth , Cell Division , and Chromosome venezuelae', *mBio*, 4(5), pp. 1–18. doi: 10.1128/mBio.00684-13.Editor.

Butt, T. R. *et al.* (2005) 'SUMO fusion technology for difficult-to-express proteins', *Protein Expression and Purification*, 43(1), pp. 1–9. doi: 10.1016/j.pep.2005.03.016.

Buttner, M. J., Chater, K. F. and Bibb, M. J. (1990) 'Cloning, disruption, and transcriptional analysis of three RNA polymerase sigma factor genes of Streptomyces coelicolor A3(2)', *Journal of Bacteriology*, 172(6), pp. 3367–3378. doi: 10.1128/jb.172.6.3367-3378.1990.

Campbell, E. A. *et al.* (2001) 'Structural mechanism for rifampicin inhibition of bacterial RNA polymerase', *Cell*, 104(6), pp. 901–912. doi: 10.1016/S0092-8674(01)00286-0.

Capstick, D. S. *et al.* (2007) 'SapB and the chaplins: Connections between morphogenetic proteins in Streptomyces coelicolor', *Molecular Microbiology*, 64(3), pp. 602–613. doi: 10.1111/j.1365-2958.2007.05674.x.

Carballido-López, R. (2006) 'The Bacterial Actin-Like Cytoskeleton', *Microbiology and Molecular Biology Reviews*, 70(4), pp. 888–909. doi: 10.1128/mmbr.00014-06.

Cashel, M. (1969) 'of Ribonucleic Acid Synthesis in Escherichia', *The Journal of biological chemistry*, 109(12), pp. 3133–3141.

Chakraburtty, R. and Bibb, M. (1997) 'The ppGpp synthetase gene (relA) of Streptomyces coelicolor A3(2) plays a conditional role in antibiotic production and morphological differentiation', *Journal of Bacteriology*, 179(18), pp. 5854–5861. doi: 10.1128/jb.179.18.5854-5861.1997.

Chamberlin, M. and Berg, P. (1964) 'Mechanism of RNA polymerase action: Characterization of the DNA-dependent synthesis of polyadenylic acid', *Journal of Molecular Biology*. Academic Press Inc. (London) Ltd., 8(5), pp. 708–726. doi: 10.1016/S0022-2836(64)80120-0.

Chater, K. F. *et al.* (1989) 'The developmental fate of S. coelicolor hyphae depends upon a gene product homologous with the motility σ factor of B. subtilis', *Cell*, 59(1), pp. 133–143. doi: 10.1016/0092-8674(89)90876-3.

Cheng, Y., Dylla, S. M. and Turnbough, C. L. (2001) 'A long T.A tract in the upp initially transcribed region is requires for regulation of upp expression by UDP-dependent reiterative transcribtion in Escherichia coli', *Journal of Bacteriology*, 183(1), pp. 221–228. doi: 10.1128/JB.183.1.221.

Cho, Y. H. *et al.* (2001) 'SigB, an RNA polymerase sigma factor required for osmoprotection and proper differentiation of Streptomyces coelicolor', *Molecular Microbiology*, 42(1), pp. 205–214. doi: 10.1046/j.1365-2958.2001.02622.x.

Claessen, D. *et al.* (2003) 'A novel class of secreted hydrophobic proteins is involved in aerial hyphae formation in Streptomyces coelicolor by forming amyloid-like fibrils', *Genes and Development*, 17(14), pp. 1714–1726. doi:

10.1101/gad.264303.

Cobb, R. E., Wang, Y. and Zhao, H. (2015) 'High-Efficiency Multiplex Genome Editing of Streptomyces Species Using an Engineered CRISPR/Cas System', pp. 2–7. doi: 10.1021/sb500351f.

Corrigan, R. M. *et al.* (2016) 'ppGpp negatively impacts ribosome assembly affecting growth and antimicrobial tolerance in Gram-positive bacteria', *Proceedings of the National Academy of Sciences*, 113(12), pp. E1710–E1719. doi: 10.1073/pnas.1522179113.

Czyz, A. *et al.* (2014) 'Mycobacterial RNA Polymerase Requires a U-Tract at Intrinsic', 5(2), pp. 1–10. doi: 10.1128/mBio.00931-14.Editor.

Dalebroux, Z. D. and Swanson, M. S. (2012) 'PpGpp: Magic beyond RNA polymerase', *Nature Reviews Microbiology*. Nature Publishing Group, 10(3), pp. 203–212. doi: 10.1038/nrmicro2720.

Darst, S. A., Kubalek, E. W. and Kornberg, R. D. (1989) 'Three-dimensional structure of Escherichia coli RNA polymerase holoenzyme determined by electron crystallography', *Nature*, 340(6236), pp. 730–732. doi: 10.1038/340730a0.

Davis, C. A. *et al.* (2007) 'Real-time footprinting of DNA in the first kinetically significant intermediate in open complex formation by Escherichia coli RNA polymerase', *Proceedings of the National Academy of Sciences of the United States of America*, 104(19), pp. 7833–7838. doi: 10.1073/pnas.0609888104.

Davis, E. *et al.* (2015) 'Mycobacterial RNA polymerase forms unstable open promoter complexes that are stabilized by CarD', *Nucleic Acids Research*, 43(1), pp. 433–445. doi: 10.1093/nar/gku1231.

Delogu, G., Sali, M. and Fadda, G. (2013) 'The biology of mycobacterium tuberculosis infection', *Mediterranean Journal of Hematology and Infectious Diseases*, 5(1). doi: 10.4084/mjhid.2013.070.

Deng, Z., Kiesei, T. and Hopwood, D. A. (1987) 'Activity of a Streptomyces transcriptional terminator in Escherichia coli', *Nucleic Acid Research*, 15(6), pp. 2665–2675.

Dennis, P. P. (1977) 'Transcription patterns of adjacent segments on the chromosome of Escherichia coli containing genes coding for four 50 S ribosomal proteins and the β and β′ subunits of RNA polymerase', *Journal of Molecular Biology*, 115(4), pp. 603–620. doi: 10.1016/0022-2836(77)90105-X.

Dombroski, A. J., Walter, W. A. and Gross, C. A. (1993) 'Amino-terminal amino acids modulate σ-factor DNA-binding activity', *Genes and Development*, 7(12 A), pp. 2446–2455. doi: 10.1101/gad.7.12a.2446.

Dove, S. L., Darst, S. A. and Hochschild, A. (2003) 'Region 4 of σ as a target for transcription regulation', *Molecular Microbiology*, 48(4), pp. 863–874. doi: 10.1046/j.1365-2958.2003.03467.x.

Dove, S. L. and Hochschild, A. (1998) 'Conversion of the ω subunit of Escherichia coli RNA polymerase into a transcriptional activator or an activation target', *Genes and Development*, 12(5), pp. 745–754. doi: 10.1101/gad.12.5.745.

Dove, S. L., Joung, J. K. and Hochschild, A. (1997) ‘Activation of prokaryotic transcription through arbitrary protein- protein contacts’, *Nature*, 386(6625), pp. 627–630. doi: 10.1038/386627a0.

Dutta, D., Chalissery, J. and Sen, R. (2008) ‘Transcription termination factor Rho prefers catalytically active elongation complexes for releasing RNA’, *Journal of Biological Chemistry*. Â© 2008 ASBMB. Currently published by Elsevier Inc; originally published by American Society for Biochemistry and Molecular Biology., 283(29), pp. 20243–20251. doi: 10.1074/jbc.M801926200.

Dykxhoorn, D. M., St. Pierre, R. and Linn, T. (1996) ‘Synthesis of the β and β′ subunits of Escherichia coli RNA polymerase is autogenously regulated in vivo by both transcriptional and translational mechanisms’, *Molecular Microbiology*, 19(3), pp. 483–493. doi: 10.1046/j.1365-2958.1996.384913.x.

Engbaek, F., Gross, C. and Burgess, R. R. (1976) ‘Quantitation of RNA Polymerase Subunits in’, 295, pp. 291–295.

Ettwiller, L. *et al.* (2016) ‘A novel enrichment strategy reveals unprecedented number of novel transcription start sites at single base resolution in a model prokaryote and the gut microbiome’, *BMC Genomics*. BMC Genomics, 17(1), pp. 1–14. doi: 10.1186/s12864-016-2539-z.

Feklistov, A. (2013) ‘RNA polymerase: In search of promoters’, *Annals of the New York Academy of Sciences*, 1293(1), pp. 25–32. doi: 10.1111/nyas.12197.

Feklistov, A. and Darst, S. A. (2011) ‘Structural basis for promoter –10 element recognition by the bacterial RNA polymerase σ subuni’, 14(4), pp. 384–399. doi: 10.1080/10810730902873927.Testing.

Fernández-Martínez, L. T. and Bibb, M. J. (2014) ‘Use of the Meganuclease I-SceI of Saccharomyces cerevisiae to select for gene deletions in actinomycetes’, *Scientific Reports*, 4, pp. 1–6. doi: 10.1038/srep07100.

Fernández-Martínez, L. T., Gomez-Escribano, J. P. and Bibb, M. J. (2015) ‘A relA-dependent regulatory cascade for auto-induction of microbisporicin production in Microbispora corallina’, *Molecular Microbiology*, 97(3), pp. 502–514. doi: 10.1111/mmi.13046.

Fernández-Moreno, M. A. *et al.* (1991) ‘The act cluster contains regulatory and antibiotic export genes, direct targets for translational control by the bldA tRNA gene of streptomyces’, *Cell*, 66(4), pp. 769–780. doi: 10.1016/0092-8674(91)90120-N.

Fillenberg, S. B. *et al.* (2016) ‘Crystal structures of the global regulator dasr from streptomyces coelicolor: Implications for the allosteric regulation of GntR/HutC Repressors’, *PLoS ONE*, 11(6), pp. 1–23. doi: 10.1371/journal.pone.0157691.

Flärdh, K. (2003) ‘Essential role of DivlVA in polar growth and morphogenesis in Streptomyces coelicolor A3(2)’, *Molecular Microbiology*, 49(6), pp. 1523–1536. doi: 10.1046/j.1365-2958.2003.03660.x.

Flärdh, K. *et al.* (2012) ‘Regulation of apical growth and hyphal branching in Streptomyces’, *Current Opinion in Microbiology*, 15(6), pp. 737–743. doi: 10.1016/j.mib.2012.10.012.

Flärdh, K. and Buttner, M. J. (2009) ‘Streptomyces morphogenetics: Dissecting differentiation in a filamentous bacterium’, *Nature Reviews Microbiology*, 7(1), pp. 36–49. doi: 10.1038/nrmicro1968.

Flentie, K., Garner, A. L. and Stallings, C. L. (2016) ‘Mycobacterium tuberculosis transcription machinery: Ready to respond to host attacks’, *Journal of Bacteriology*, 198(9), pp. 1360–1373. doi: 10.1128/JB.00935-15.

Flett F, Mersinias V, S. C. (1997) ‘High efficiency intergeneric conjugal transfer of plasmid DNA from’, *FEMS Microbiology Letters*, 155.

Gajos, M. *et al.* (2021) ‘Conserved DNA sequence features underlie pervasive RNA polymerase pausing’, *Nucleic Acids Research*, 49(8), pp. 4402–4420. doi: 10.1093/nar/gkab208.

Galaz-Davison, P., Román, E. A. and Ramírez-Sarmiento, C. A. (2021) ‘The N-terminal domain of RfaH plays an active role in protein fold-switching’, *PLoS Computational Biology*, 17(9), pp. 1–18. doi: 10.1371/journal.pcbi.1008882.

Garst, A. D., Edwards, A. L. and Batey, R. T. (2011) ‘Riboswitches: Structures and mechanisms’, *Cold Spring Harbor Perspectives in Biology*, 3(6), pp. 1–13. doi: 10.1101/cshperspect.a003533.

Gatewood, M. L. and Jones, G. H. (2010) ‘(p)ppGpp inhibits polynucleotide phosphorylase from Streptomyces but not from Escherichia coli and increases the stability of bulk mRNA in Streptomyces coelicolor’, *Journal of Bacteriology*, 192(17), pp. 4275–4280. doi: 10.1128/JB.00367-10.

Geiger, T. *et al.* (2012) ‘The Stringent Response of Staphylococcus aureus and Its Impact on Survival after Phagocytosis through the Induction of Intracellular PSMs Expression’, *PLoS Pathogens*, 8(11). doi: 10.1371/journal.ppat.1003016.

Gellert, M. *et al.* (1976) ‘DNA gyrase: an enzyme that introduces superhelical turns into DNA’, *Proceedings of the National Academy of Sciences of the United States of America*, 73(11), pp. 3872–3876. doi: 10.1073/pnas.73.11.3872.

Ghosh, P., Ishihama, A. and Chatterji, D. (2001) ‘Escherichia coli RNA polymerase subunit ω and its N-terminal domain bind full-length β′ to facilitate incorporation into the α2β subassembly’, *European Journal of Biochemistry*, 268(17), pp. 4621–4627. doi: 10.1046/j.1432-1327.2001.02381.x.

Gomez-Escribano, J. P. *et al.* (2008) ‘Streptomyces clavuligerus relA-null mutants overproduce clavulanic acid and cephamycin C: Negative regulation of secondary metabolism by (p)ppGpp’, *Microbiology*, 154(3), pp. 744–755. doi: 10.1099/mic.0.2007/011890-0.

Gomez-Escribano, J. P. *et al.* (2012) ‘Structure and biosynthesis of the unusual polyketide alkaloid coelimycin P1, a metabolic product of the cpk gene cluster of Streptomyces coelicolor M145’, *Chemical Science*, 3(9), pp. 2716–2720. doi: 10.1039/c2sc20410j.

Gomez-Escribano, J. P. and Bibb, M. J. (2011) ‘Engineering Streptomyces coelicolor for heterologous expression of secondary metabolite gene clusters’, *Microbial Biotechnology*, 4(2), pp. 207–215. doi: 10.1111/j.1751-7915.2010.00219.x.

Gopal, B. *et al.* (2001) 'Crystal structure of the transcription elongation/antitermination factor NusA from Mycobacterium tuberculosis at 1.7 Å resolution', *Journal of Molecular Biology*, 314(5), pp. 1087–1095. doi: 10.1006/jmbi.2000.5144.

Gourse, R. L. *et al.* (2018) 'Transcriptional Responses to ppGpp and DksA', *Physiology & behavior*, 176(3), pp. 139–148. doi: 10.1146/annurev-micro-090817-062444.Transcriptional.

Gourse, R. L., Ross, W. and Gaal, T. (2000) 'UPs and downs in bacterial transcription initiation: The role of the alpha subunit of RNA polymerase in promoter recognition', *Molecular Microbiology*, 37(4), pp. 687–695. doi: 10.1046/j.1365-2958.2000.01972.x.

Gregory, M. a, Till, R. and Smith, M. C. M. (2003) 'Integration Site for Streptomyces Phage ?BT1 and Development of Site-Specific Integrating Vectors', *Society*, 185(17), pp. 5320–5323. doi: 10.1128/JB.185.17.5320.

Gruber, T. M. and Gross, C. A. (2003) 'Multiple Sigma Subunits and the Partitioning of Bacterial Transcription Space', *Annual Review of Microbiology*, 57(1), pp. 441–466. doi: 10.1146/annurev.micro.57.030502.090913.

Gupta, S. and Pal, D. (2021) 'Clusters of hairpins induce intrinsic transcription termination in bacteria', *Scientific Reports*. Nature Publishing Group UK, 11(1), pp. 1–18. doi: 10.1038/s41598-021-95435-3.

Gusarov, I. and Nudler, E. (1999) 'The mechanism of intrinsic transcription termination', *Molecular Cell*, 3(4), pp. 495–504. doi: 10.1016/S1097-2765(00)80477-3.

Guthrie, E. P. *et al.* (1998) 'A response-regulator-like activator of antibiotic synthesis from Streptomyces coelicolor A3(2) with an amino-terminal domain that lacks a phosphorylation pocket', *Microbiology*, 144(3), pp. 727–738. doi: 10.1099/00221287-144-3-727.

Han, X. and Turnbough, C. L. (1998) 'Regulation of carAB expression in Escherichia coli occurs in part through UTP-sensitive reiterative transcription', *Journal of Bacteriology*, 180(3), pp. 705–713. doi: 10.1128/jb.180.3.705-713.1998.

Han, X. and Turnbough, C. L. (2014) 'Transcription start site sequence and spacing between the -10 region and the start site affect reiterative transcription-mediated regulation of gene expression in Escherichia coli', *Journal of Bacteriology*, 196(16), pp. 2912–2920. doi: 10.1128/JB.01753-14.

Haseltine, W. A. and Block, R. (1973) 'Synthesis of guanosine tetra- and pentaphosphate requires the presence of a codon-specific, uncharged transfer ribonucleic acid in the acceptor site of ribosomes.', *Proceedings of the National Academy of Sciences of the United States of America*, 70(5), pp. 1564–8. doi: 10.1073/pnas.70.5.1564.

Haug, I. *et al.* (2003) 'Streptomyces coelicolor A3(2) plasmid SCP2*: Deductions from the complete sequence', *Microbiology*, 149(2), pp. 505–513. doi: 10.1099/mic.0.25751-0.

Haugen, S. P., Ross, W. and Gourse, R. L. (2008) 'Advances in bacterial promoter

recognition and its control by factors that do not bind DNA', *Nature Reviews Microbiology*, 6(7), pp. 507–519. doi: 10.1038/nrmicro1912.

Hein, P. P. *et al.* (2014) 'RNA polymerase pausing and nascent RNA structure formation are Linked through clamp domain movement', 21(9), pp. 794–802. doi: 10.1038/nsmb.2867.RNA.

Helmann, J. D. (2002) *The extracytoplasmic function (ECF) sigma factors, Advances in Microbial Physiology*. doi: 10.1016/S0065-2911(02)46002-X.

Hempel, A. M. *et al.* (2008) 'Assemblies of DivIVA mark sites for hyphal branching and can establish new zones of cell wall growth in Streptomyces coelicolor', *Journal of Bacteriology*, 190(22), pp. 7579–7583. doi: 10.1128/JB.00839-08.

Hengst, C. D. de. *et al.* (2010) 'Genes essential for morphological development and antibiotic production in Streptomyces coelicolor are targets of BldD during vegetative growth', *Molecular Microbiology*, 78(2), pp. 361–379. doi: 10.1111/j.1365-2958.2010.07338.x.

Henkin, T. M. and Grundy, F. J. (2006) 'Sensing metabolic signals with nascent RNA tanscripts: The T box and S box riboswitches as paradigms', *Cold Spring Harbor Symposia on Quantitative Biology*, 71, pp. 231–237. doi: 10.1101/sqb.2006.71.020.

Herbert, K. M. *et al.* (2006) 'Sequence-Resolved Detection of Pausing by Single RNA Polymerase Molecules', 125(6), pp. 1083–1094.

Hesketh, A. *et al.* (2007a) 'The global role of ppGpp synthesis in morphological differentiation and antibiotic production in Streptomyces coelicolor A3(2)', *Genome Biology*, 8(8). doi: 10.1186/gb-2007-8-8-r161.

Hesketh, A. *et al.* (2007b) 'The global role of ppGpp synthesis in morphological differentiation and antibiotic production in Streptomyces coelicolor A3(2)', *Genome Biology*, 8(8). doi: 10.1186/gb-2007-8-8-r161.

Hesketh, A., Sun, J. and Bibb, M. (2001) 'Induction of ppGpp synthesis in Streptomyces coelicolor A3(2) grown under conditions of nutritional sufficiency elicits actII-ORF4 transcription and actinorhodin biosynthesis', *Molecular Microbiology*, 39(1), pp. 136–144. doi: 10.1046/j.1365-2958.2001.02221.x.

Heyduk, E. and Heyduk, T. (2018) 'DNA template sequence control of bacterial RNA polymerase escape from the promoter', *Nucleic Acids Research*. Oxford University Press, 46(9), pp. 4469–4486. doi: 10.1093/nar/gky172.

Hirschman, J. *et al.* (1985) 'Products of nitrogen regulatory genes ntrA and ntrC of enteric bacteria activate glnA transcription in vitro: Evidence that the ntrA product is a σ factor', *Proceedings of the National Academy of Sciences of the United States of America*, 82(22), pp. 7525–7529. doi: 10.1073/pnas.82.22.7525.

Hochgräfe, F. *et al.* (2007) 'S-cysteinylation is a general mechanism for thiol protection of Bacillus subtilis proteins after oxidative stress', *Journal of Biological Chemistry*, 282(36), pp. 25961–25985. doi: 10.1074/jbc.C700105200.

Hochschild, A. and Dove, S. L. (1998) 'Protein-protein contacts that activate and repress prokaryotic transcription', *Cell*, 92(5), pp. 597–600. doi: 10.1016/S0092-8674(00)81126-5.

Hodgson, D. A. (2000) ‘Primary metabolism and its control in Streptomycetes: A most unusual group of bacteria’, *Advances in Microbial Physiology*, 42, pp. 47–238. doi: 10.1016/s0065-2911(00)42003-5.

Hojati, Z. *et al.* (2002) ‘Structure, biosynthetic origin, and engineered biosynthesis of calcium-dependent antibiotics from Streptomyces coelicolor’, *Chemistry and Biology*, 9(11), pp. 1175–1187. doi: 10.1016/S1074-5521(02)00252-1.

Hopwood, D. A. and Wright, H. M. (1983) ‘CDA is a new chromosomally-determined antibiotic from Streptomyces coelicolor A3(2)’, *Journal of General Microbiology*, 129(12), pp. 3575–3579. doi: 10.1099/00221287-129-12-3575.

Horbal, L., Siegl, T. and Luzhetskyy, A. (2018) ‘A set of synthetic versatile genetic control elements for the efficient expression of genes in Actinobacteria’, *Scientific Reports*. Springer US, 8(1), pp. 1–13. doi: 10.1038/s41598-017-18846-1.

Hosaka, T. *et al.* (2009) ‘Antibacterial discovery in actinomycetes strains with mutations in RNA polymerase or ribosomal protein S12’, *Nature Biotechnology*, 27(5), pp. 462–464. doi: 10.1038/nbt.1538.

Hoskisson, P. A. and Fernández-Martínez, L. T. (2018) ‘Regulation of specialised metabolites in Actinobacteria – expanding the paradigms’, *Environmental Microbiology Reports*, 10(3), pp. 231–238. doi: 10.1111/1758-2229.12629.

Hu, Y. *et al.* (2014) ‘Mycobacterium RbpA cooperates with the stress-response σb subunit of RNA polymerase in promoter DNA unwinding’, *Nucleic Acids Research*, 42(16), pp. 10399–10408. doi: 10.1093/nar/gku742.

Hu, Y. and Coates, A. R. M. (2001) ‘Increased levels of sigJ mRNA in late stationary phase cultures of Mycobacterium tuberculosis detected by DNA array hybridisation’, *FEMS Microbiology Letters*, 202(1), pp. 59–65. doi: 10.1016/S0378-1097(01)00280-4.

Hubin, E. A. *et al.* (2015) ‘Structural, functional, and genetic analyses of the actinobacterial transcription factor RbpA’, *Proceedings of the National Academy of Sciences of the United States of America*, 112(23), pp. 7171–7176. doi: 10.1073/pnas.1504942112.

Hubin, E. A. *et al.* (2017) ‘Structure and function of the mycobacterial transcription initiation complex with the essential regulator RbpA’, *eLife*, 6, pp. 1–40. doi: 10.7554/eLife.22520.

Hudson, B. P. *et al.* (2009) ‘Three-dimensional EM structure of an intact activator-dependent transcription initiation complex’, *Proceedings of the National Academy of Sciences of the United States of America*, 106(47), pp. 19830–19835. doi: 10.1073/pnas.0908782106.

Huff, J. *et al.* (2010) ‘Taking phage integration to the next level as a genetic tool for mycobacteria’, *Gene*. Elsevier B.V., 468(1–2), pp. 8–19. doi: 10.1016/j.gene.2010.07.012.

Imashimizu, M. *et al.* (2020) ‘Control of transcription initiation by biased thermal fluctuations on repetitive genomic sequences’, *Biomolecules*, 10(9), pp. 1–22. doi: 10.3390/biom10091299.

Imber, M., Pietrzyk-Brzezinska, A. J. and Antelmann, H. (2019) ‘Redox regulation

by reversible protein S-thiolation in Gram-positive bacteria', *Redox Biology*. Elsevier B.V., 20(July 2018), pp. 130–145. doi: 10.1016/j.redox.2018.08.017.

Irving, S. E., Choudhury, N. R. and Corrigan, R. M. (2021) 'The stringent response and physiological roles of (pp)pGpp in bacteria', *Nature Reviews Microbiology*. Springer US, 19(4), pp. 256–271. doi: 10.1038/s41579-020-00470-y.

Izard, J. *et al.* (2015) 'A synthetic growth switch based on controlled expression of RNA polymerase', *Molecular Systems Biology*, 11(11), p. 840. doi: 10.15252/msb.20156382.

Jacques, J. P. and Kolakofsky, D. (1991) 'Pseudo-templated transcription in prokaryotic and eukaryotic organisms', *Genes and Development*, 5(5), pp. 707–713. doi: 10.1101/gad.5.5.707.

Jakimowicz, D. *et al.* (2006) 'Developmental Control of a parAB Promoter Leads to Formation of Sporulation-Associated', *Society*, 188(5), pp. 1710–1720. doi: 10.1128/JB.188.5.1710.

Jakimowicz, D., Chater, K. and Zakrzewska-Czerwínska, J. (2002) 'The ParB protein of Streptomyces coelicolor A3(2) recognizes a cluster of parS sequences within the origin-proximal region of the linear chromosome', *Molecular Microbiology*, 45(5), pp. 1365–1377. doi: 10.1046/j.1365-2958.2002.03102.x.

Jensen, D. *et al.* (2019) 'CarD and RbpA modify the kinetics of initial transcription and slow promoter escape of the Mycobacterium tuberculosis RNA polymerase', *Nucleic Acids Research*. Oxford University Press, 47(13), pp. 6685–6698. doi: 10.1093/nar/gkz449.

Jensen, D. and Galburt, E. A. (2021) 'The context-dependent influence of promoter sequence motifs on transcription initiation kinetics and regulation', *Journal of Bacteriology*, 203(8), pp. 1–19. doi: 10.1128/JB.00512-20.

Jeong, Y. *et al.* (2016a) 'The dynamic transcriptional and translational landscape of the model antibiotic producer Streptomyces coelicolor A3(2)', *Nature Communications*. Nature Publishing Group, 7, pp. 1–11. doi: 10.1038/ncomms11605.

Jeong, Y. *et al.* (2016b) 'The dynamic transcriptional and translational landscape of the model antibiotic producer Streptomyces coelicolor A3(2)', *Nature Communications*. Nature Publishing Group, 7, pp. 1–11. doi: 10.1038/ncomms11605.

Jha, R. K., Tare, P. and Nagaraja, V. (2018) 'Regulation of the gyr operon of Mycobacterium tuberculosis by overlapping promoters, DNA topology, and reiterative transcription', *Biochemical and Biophysical Research Communications*. Elsevier Ltd, 501(4), pp. 877–884. doi: 10.1016/j.bbrc.2018.05.067.

Jin, D. J., Cagilero, C. and Zhou, Y. N. (2012) 'Growth rate regulation in Escherichia coli', *Bone*, 36(2), pp. 269–287. doi: 10.1111/j.1574-6976.2011.00279.x.Growth.

Johnson, G. E. *et al.* (2020) 'Functionally uncoupled transcription–translation in Bacillus subtilis', *Nature*, 585(7823), pp. 124–128. doi: 10.1038/s41586-020-2638-5.

Kallifidas, D. *et al.* (2010) 'The σR regulon of Streptomyces coelicolor A3(2) reveals a key role in protein quality control during disulphide stress', *Microbiology*, 156(6), pp. 1661–1672. doi: 10.1099/mic.0.037804-0.

Kang, J. G. *et al.* (1999) 'RsrA, an anti-sigma factor regulated by redox change', *EMBO Journal*, 18(15), pp. 4292–4298. doi: 10.1093/emboj/18.15.4292.

Kang, J. Y. *et al.* (2018) 'Structural Basis for Transcript Elongation Control by NusG Family Universal Regulators', *Cell*. Elsevier Inc., 173(7), pp. 1650–1662. doi: 10.1016/j.cell.2018.05.017.

Kang, J. Y. *et al.* (2019) 'Mechanisms of Transcriptional Pausing in Bacteria', *Journal of Molecular Biology*. Elsevier Ltd, 431(20), pp. 4007–4029. doi: 10.1016/j.jmb.2019.07.017.

Kang, S. G. *et al.* (1998) 'Actinorhodin and undecylprodigiosin production in wild-type and *relA* mutant strains of *Streptomyces coelicolor* A3(2) grown in continuous culture', *FEMS Microbiology Letters*, 168(2), pp. 221–226. doi: 10.1111/j.1574-6968.1998.tb13277.x.

Karam, C. and Siadati, A. (2021) 'Diagnosis and Neurosurgical Management of Cerebral Nocardiosis', *Journal of Neurological Surgery Reports*, 82(02), pp. e21–e24. doi: 10.1055/s-0040-1722345.

Kästle, B. *et al.* (2015) 'rRNA regulation during growth and under stringent conditions in Staphylococcus aureus', *Environmental Microbiology*, 17(11), pp. 4394–4405. doi: 10.1111/1462-2920.12867.

Keiler, K. C., Waller, P. R. H. and Sauer, R. T. (1996) 'Role of a peptide tagging system in degredation of proteins synthesised from damaged messenger RNA', *Science*, 271(February), pp. 990–993.

Kelemen, G. H. *et al.* (1996) 'The positions of the sigma-factor genes, whiG and sigF, in the hierarchy controlling the development of spore chains in the aerial hyphae of Streptomyces coelicolor A3(2)', *Molecular Microbiology*, 21(3), pp. 593–603. doi: 10.1111/j.1365-2958.1996.tb02567.x.

Kim, H. J. *et al.* (2000) 'Partitioning of the linear chromosome during sporulation of Streptomyces coelicolor A3(2) involves an oriC-linked parAB locus', *Journal of Bacteriology*, 182(5), pp. 1313–1320. doi: 10.1128/JB.182.5.1313-1320.2000.

Kim, M. *et al.* (2012) 'Conservation of thiol-oxidative stress responses regulated by SigR orthologues in actinomycetes', *Mol microbiol*, 85(2), pp. 326–344. doi: 10.1111/j.1365-2958.2012.08115.x.Conservation.

Kinashi, H. and Shimaji-Murayama, M. (1991) 'Physical characterization of SCP1, a giant linear plasmid from Streptomyces coelicolor', *Journal of Bacteriology*, 173(4), pp. 1523–1529. doi: 10.1128/jb.173.4.1523-1529.1991.

Kodani, S. *et al.* (2004) 'The SapB morphogen is a lantibiotic-like peptide derived from the product of the developmental gene ramS in Streptomyces coelicolor', *Proceedings of the National Academy of Sciences of the United States of America*, 101(31), pp. 11448–11453. doi: 10.1073/pnas.0404220101.

Kolb, K. E., Hein, P. P. and Landick, R. (2014) 'Antisense oligonucleotide-stimulated transcriptional pausing reveals RNA exit channel specificity of RNA

polymerase and mechanistic contributions of NusA and RfaH', *Journal of Biological Chemistry*, 289(2), pp. 1151–1163. doi: 10.1074/jbc.M113.521393.

Kontur, W. S. *et al.* (2008) 'Probing DNA Binding, DNA Opening and Assembly of a Downstream Clamp/Jaw in E. coli RNA Polymerase – λPR Promoter Complexes Using Salt and the Physiological Anion Glutamate', *Bone*, 23(1), pp. 1–7. doi: 10.1021/bi100092a.Probing.

Krásný, L. *et al.* (2008) 'The identity of the transcription +1 position is crucial for changes in gene expression in response to amino acid starvation in Bacillus subtilis', *Molecular Microbiology*, 69(1), pp. 42–54. doi: 10.1111/j.1365-2958.2008.06256.x.

Krásny, L. and Gourse, R. L. (2004) 'An alternative strategy for bacterial ribosome synthesis: Bacillus subtilis rRNA transcription regulation', *EMBO Journal*, 23(22), pp. 4473–4483. doi: 10.1038/sj.emboj.7600423.

Kriel, A. *et al.* (2013) 'Direct Regulation of GTP homeostasis by (p)ppGpp: A Critical Component of Viability and Stress Resistance', 48(2), pp. 231–241. doi: 10.1016/j.molcel.2012.08.009.Direct.

Kriel, A. *et al.* (2014) 'GTP dysregulation in Bacillus subtilis cells lacking (p)ppGpp results in phenotypic amino acid auxotrophy and failure to adapt to nutrient downshift and regulate biosynthesis genes', *Journal of Bacteriology*, 196(1), pp. 189–201. doi: 10.1128/JB.00918-13.

Kurkela, J. *et al.* (2021) 'Revealing secrets of the enigmatic omega subunit of bacterial RNA polymerase', *Molecular Microbiology*, 115(1), pp. 1–11. doi: 10.1111/mmi.14603.

Küster, C., Piepersberg, W. and Distler, J. (1998) 'Cloning and transcriptional analysis of the rplKA-or f31-rplJL gene cluster of Streptomyces griseus', *Mol Gen Genet*, 257(2), pp. 219–229. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9491081.

Lai, C. *et al.* (2002) 'Genetic and physiological characterization of rpoB mutations that activate antibiotic production in Streptomyces lividans', *Microbiology*, 148(11), pp. 3365–3373. doi: 10.1099/00221287-148-11-3365.

Landick, R. (2021) 'Transcriptional Pausing as a Mediator of Bacterial Gene Regulation', *Annual Review of Microbiology*, 75, pp. 291–314. doi: 10.1146/annurev-micro-051721-043826.

Lane, W. J. and Darst, S. A. (2010) 'Molecular Evolution of Multi-subunit RNA Polymerases: Structural Analysis', *J Mol Biol*, 395(4), p. 686. doi: 10.1016/j.jmb.2009.10.063.Molecular.

Larsen, B. *et al.* (2000) 'Nonlinearity in genetic decoding: Homologous DNA replicase genes use alternatives of transcriptional slippage or translational frameshifting', *Proceedings of the National Academy of Sciences of the United States of America*, 97(4), pp. 1683–1688. doi: 10.1073/pnas.97.4.1683.

Lawson, M. R. and Berger, J. M. (2019) 'Tuning the sequence specificity of a transcription terminator', *Current Genetics*, 65(3), pp. 729–733. doi: 10.1007/s00294-019-00939-1.

Lee, L. H. *et al.* (2014) 'Streptomyces pluripotens sp. nov., A bacteriocin-producing

streptomycete that inhibits meticillin-resistant Staphylococcus aureus', *International Journal of Systematic and Evolutionary Microbiology*, 64. doi: 10.1099/ijs.0.065045-0.

Lee, S. J. *et al.* (2009) 'Cellular stress created by intermediary metabolite imbalances', *Proceedings of the National Academy of Sciences of the United States of America*, 106(46), pp. 19515–19520. doi: 10.1073/pnas.0910586106.

Lewis, D. E. A. and Adhya, S. (2004) 'Axiom of determining transcription start points by RNA polymerase in Escherichia coli', *Molecular Microbiology*, 54(3), pp. 692–701. doi: 10.1111/j.1365-2958.2004.04318.x.

Lewis, R. A. *et al.* (2019) 'Genome-wide analysis of the role of the antibiotic biosynthesis regulator AbsA2 in Streptomyces coelicolor A3(2)', *PLoS ONE*, 14(4), pp. 1–23. doi: 10.1371/journal.pone.0200673.

Li, L. *et al.* (2020) 'RNA extension drives a stepwise displacement of an initiation-factor structural module in initial transcription', *Proceedings of the National Academy of Sciences of the United States of America*, 117(11), pp. 5801–5809. doi: 10.1073/pnas.1920747117.

Li, Z., Pandit, S. and Deutscher, M. P. (1999) 'RNase G (CafA protein) and RNase E are both required for the 5' maturation of 16S ribosomal RNA', *EMBO Journal*, 18(10), pp. 2878–2885. doi: 10.1093/emboj/18.10.2878.

Lievin, V. *et al.* (2000) 'Bifidobacterium strains from resident infant human gastrointestinal microflora exert antimicrobial activity', *Gut*, 47(5), pp. 646–652. doi: 10.1136/gut.47.5.646.

Lilic, M., Darst, S. A. and Campbell, E. A. (2021) 'Structural basis of transcriptional activation by the Mycobacterium tuberculosis intrinsic antibiotic-resistance transcription factor WhiB7', *Molecular Cell*. Elsevier Inc., 81(14), pp. 2875-2886.e5. doi: 10.1016/j.molcel.2021.05.017.

de Lima Procópio, R. E. *et al.* (2012) 'Antibiotics produced by Streptomyces', *Brazilian Journal of Infectious Diseases*. Elsevier Editora Ltda, 16(5), pp. 466–471. doi: 10.1016/j.bjid.2012.08.014.

Liu, C., Heath, L. S. and Turnbough Jr., C. L. (1994) 'Regulation of pyrBI operon expression in Escherichia coli by UTP-', *Genes Dev*, 8(23), pp. 2904–2912.

Liu, G. *et al.* (2013) 'Molecular Regulation of Antibiotic Biosynthesis in Streptomyces', *Microbiology and Molecular Biology Reviews*, 77(1), pp. 112–143. doi: 10.1128/MMBR.00054-12.

Liu, K., Myers, Angela R., *et al.* (2015) 'Molecular mechanism and evolution of guanylate kinase regulation by (p)ppGpp', *Molecular Cell*. Elsevier Inc., 57(4), pp. 735–749. doi: 10.1016/j.molcel.2014.12.037.

Liu, K., Myers, Angela R, *et al.* (2015) 'Molecular mechanism and evolution of guanylate kinase regulation by (p)ppGpp', 57(4), pp. 735–749. doi: 10.1016/j.molcel.2014.12.037.Molecular.

Liu, K., Bittner, A. N. and Wang, J. D. (2015) 'Diversity in (p)ppGpp metabolism and effectors', 344(6188), pp. 1173–1178. doi: 10.1126/science.1249098.Sleep.

Liu, Y. *et al.* (2022) 'Structural and mechanistic basis of reiterative transcription initiation', *Proceedings of the National Academy of Sciences of the United States of America*, 119(5), pp. 1–12. doi: 10.1073/pnas.2115746119.

Lonetto, M. A. *et al.* (1994) 'Analysis of the Streptomyces coelicolor sigE gene reveals the existence of a subfamily of eubacterial RNA polymerase σ factors involved in the regulation of extracytoplasmic functions', *Proceedings of the National Academy of Sciences of the United States of America*, 91(16), pp. 7573–7577. doi: 10.1073/pnas.91.16.7573.

Lonetto, M., Gribskov, M. and Gross, C. A. (1992) 'The σ70 family: Sequence conservation and evolutionary relationships', *Journal of Bacteriology*, 174(12), pp. 3843–3849. doi: 10.1128/jb.174.12.3843-3849.1992.

Łyzén, R. *et al.* (2016) 'The dual role of DksA protein in the regulation of Escherichia coli pArgX promoter', *Nucleic Acids Research*, 44(21), pp. 10316–10325. doi: 10.1093/nar/gkw912.

MacNeil, D. J. *et al.* (1992) 'Analysis of Streptomyces avermitilis genes required for avermectin biosynthesis utilizing a novel integration vector', *Gene*, 111(1), pp. 61–68. doi: 10.1016/0378-1119(92)90603-M.

Magnusson, L. U., Farewell, A. and Nyström, T. (2005) 'ppGpp: A global regulator in Escherichia coli', *Trends in Microbiology*, 13(5), pp. 236–242. doi: 10.1016/j.tim.2005.03.008.

Mao, C. *et al.* (2018) 'Association of ω with the Cterminal region of the β' subunit is essential for assembly of RNA polymerase in Mycobacterium tuberculosis', *Journal of Bacteriology*, 200(12), pp. 1–9. doi: 10.1128/JB.00159-18.

Marchand, J. A. *et al.* (2021) 'TBDB: A database of structurally annotated T-box riboswitch:tRNA pairs', *Nucleic Acids Research*. Oxford University Press, 49(D1), pp. D229–D235. doi: 10.1093/nar/gkaa721.

Mathew, R. and Chatterji, D. (2006) 'The evolving story of the omega subunit of bacterial RNA polymerase', *Trends in Microbiology*, 14(10), pp. 450–455. doi: 10.1016/j.tim.2006.08.002.

Mathews, D. H. (2014) 'RNA secondary structure analysis using RNAstructure', *Current Protocols in Bioinformatics*, (SUPPL.46), pp. 1–25. doi: 10.1002/0471250953.bi1206s46.

McCormick, J. R. *et al.* (1994) 'Growth and viability of Streptomyces coelicolor mutant for the cell division gene ftsZ', *Molecular Microbiology*, 14(2), pp. 243–254. doi: 10.1111/j.1365-2958.1994.tb01285.x.

McKenzie, N. L. and Nodwell, J. R. (2007) 'Phosphorylated AbsA2 negatively regulates antibiotic production in Streptomyces coelicolor through interactions with pathway-specific regulatory gene promoters', *Journal of Bacteriology*, 189(14), pp. 5284–5292. doi: 10.1128/JB.00305-07.

Meng, Q., Turnbough, C. L. and Switzer, R. L. (2004) 'Attenuation control of pyrG expression in Bacillus subtilis is mediated by CTP-sensitive reiterative transcription.', *Proceedings of the National Academy of Sciences of the United States of America*, 101(30), pp. 10943–8. doi: 10.1073/pnas.0403755101.

Merino, E., Babitzke, P. and Yanofsky, C. (1995) ‘trp RNA-binding attenuation protein (TRAP)-trp leader RNA interactions mediate translational as well as transcriptional regulation of the Bacillus subtilis trp operon’, *Journal of Bacteriology*, 177(22), pp. 6362–6370. doi: 10.1128/jb.177.22.6362-6370.1995.

Miroux, B. and Walker, J. E. (1996) ‘#044_Miroux et al_Over-production of proteins in Escherichia coli_ mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels J_ Mol_ Biol_ 260_ 289-298 (1996)_.pdf’, *J Mol Biol*, 260, pp. 289–298.

Mitchell, J. E. *et al.* (2003) ‘Identification and analysis of “extended -10” promoters in Escherichia coli’, *Nucleic Acids Research*, 31(16), pp. 4689–4695. doi: 10.1093/nar/gkg694.

Mitra, P. *et al.* (2017) ‘Rho Protein: Roles and Mechanisms’, *Annual Review of Microbiology*, 71(July), pp. 687–709. doi: 10.1146/annurev-micro-030117-020432.

Molodtsov, V. *et al.* (2018) ‘Allosteric Effector ppGpp Potentiates the Inhibition of Transcript Initiation by DksA’, *Molecular Cell*. Elsevier Inc., 69(5), pp. 828-839.e5. doi: 10.1016/j.molcel.2018.01.035.

Molodtsov V et al. (2014) ‘Slippage Promotes Termination By T7 Rna’, 426(18), pp. 3095–3107. doi: 10.1016/j.jmb.2014.06.012.THE.

Mondal, S., Yakhnin, A. V. and Babitzke, P. (2017) ‘Modular organization of the NusA- and NusG-stimulated RNA polymerase pause signal that participates in the Bacillus subtilis trp operon attenuation mechanism’, *Journal of Bacteriology*, 199(14), pp. 1–12. doi: 10.1128/JB.00223-17.

Murakami, K. S. (2013) ‘X-ray crystal structure of escherichia coli RNA polymerase σ70 holoenzyme’, *Journal of Biological Chemistry*, 288(13), pp. 9126–9134. doi: 10.1074/jbc.M112.430900.

Murakami, K. S. *et al.* (2017) ‘X-ray crystal structure of a reiterative transcription complex reveals an atypical RNA extension pathway’, *Proceedings of the National Academy of Sciences*, 114(31), pp. 8211–8216. doi: 10.1073/pnas.1702741114.

Murakami, K. S. and Darst, S. A. (2003) ‘Bacterial RNA polymerases: The wholo story’, *Current Opinion in Structural Biology*, 13(1), pp. 31–39. doi: 10.1016/S0959-440X(02)00005-2.

Narayanan, A. *et al.* (2018) ‘Cryo-EM structure of Escherichia coli 70 RNA polymerase and promoter DNA complex revealed a role of non-conserved region during the open complex formation’, *Journal of Biological Chemistry*, 293(19), pp. 7367–7375. doi: 10.1074/jbc.RA118.002161.

Narva, K. E. and Feitelson, J. S. (1990) ‘Nucleotide sequence and transcriptional analysis of the redD locus of Streptomycescoelicolor A3(2)’, *Journal of Bacteriology*, 172(1), pp. 326–333. doi: 10.1128/jb.172.1.326-333.1990.

Newell, K. V. *et al.* (2006) ‘The RNA polymerase-binding protein RbpA confers basal levels of rifampicin resistance on Streptomyces coelicolor’, *Molecular Microbiology*, 60(3), pp. 687–696. doi: 10.1111/j.1365-2958.2006.05116.x.

Newman, A. J., Linn, T. G. and Hayward, R. S. (1979) ‘Evidence for Co-transcription of the RNA Polymerase Genes rpoBC with a Ribosomal Protein Gene

of Escherichia coli', *Molec.gen.genet*, (169), pp. 195–204.

Ochi, K. (1986) 'Occurrence of the stringent response in Streptomyces sp. and its significance for the initiation of morphological and physiological differentiation.', *Journal of general microbiology*, 132(9), pp. 2621–31. doi: 10.1099/00221287-132-9-2621.

Ochi, K. (1987) 'Changes in Nucleotide Pools during Sporulation of Streptomyces griseus in Submerged Culture', *Microbiology*, 133(10), pp. 2787–2795. doi: 10.1099/00221287-133-10-2787.

Ochi, K. and Hosaka, T. (2013) 'New strategies for drug discovery: Activation of silent or weakly expressed microbial gene clusters', *Applied Microbiology and Biotechnology*, 97(1), pp. 87–98. doi: 10.1007/s00253-012-4551-9.

Ochi, K., Tanaka, Y. and Tojo, S. (2014) 'Activating the expression of bacterial cryptic genes by rpoB mutations in RNA polymerase or by rare earth elements', *Journal of Industrial Microbiology and Biotechnology*, 41(2), pp. 403–414. doi: 10.1007/s10295-013-1349-4.

Oguienko, A. *et al.* (2021) 'Universal functions of the σ finger in alternative σ factors during transcription initiation by bacterial RNA polymerase', *RNA Biology*. Taylor & Francis, 18(11), pp. 2028–2037. doi: 10.1080/15476286.2021.1889254.

Ole Leichert, L. I., Scharf, C. and Hecker, M. (2003) 'Global characterization of disulfide stress in Bacillus subtilis', *Journal of Bacteriology*, 185(6), pp. 1967–1975. doi: 10.1128/JB.185.6.1967-1975.2003.

Opalka, N. *et al.* (2010) 'Complete structural model of Escherichia coli RNA polymerase from a Hybrid Approach', *PLoS Biology*, 8(9). doi: 10.1371/journal.pbio.1000483.

Osaka, N. *et al.* (2020) 'Novel (p)ppGpp0 suppressor mutations reveal an unexpected link between methionine catabolism and GTP synthesis in Bacillus subtilis', *Molecular Microbiology*, 113(6), pp. 1155–1169. doi: 10.1111/mmi.14484.

Otani, H. *et al.* (2013) 'An alternative sigma factor governs the principal sigma factor in Streptomyces griseus', *Molecular Microbiology*, 87(6), pp. 1223–1236. doi: 10.1111/mmi.12160.

Paget, M. S. (2015) 'Bacterial sigma factors and anti-sigma factors: Structure, function and distribution', *Biomolecules*, 5(3), pp. 1245–1265. doi: 10.3390/biom5031245.

Paget, M. S. B. (1998) 'sigma R, an RNA polymerase sigma factor that modulates expression of the thioredoxin system in response to oxidative stress in Streptomyces coelicolor A3(2)', *The EMBO Journal*, 17(19), pp. 5776–5782. doi: 10.1093/emboj/17.19.5776.

Paget, M. S. B. *et al.* (1999) 'Evidence that the extracytoplasmic function sigma factor σ(E) is required for normal cell wall structure in Streptomyces coelicolor A3(2)', *Journal of Bacteriology*, 181(1), pp. 204–211. doi: 10.1128/jb.181.1.204-211.1999.

Paget, M. S. B. *et al.* (2001) 'Defining the disulphide stress response in Streptomyces coelicolor A3(2): Identification of the σR regulon', *Molecular*

Microbiology, 42(4), pp. 1007–1020. doi: 10.1046/j.1365-2958.2001.02675.x.

Paget, M. S. B. and Helmann, J. D. (2003) 'The sigma70 family of sigma factors.', *Genome biology*, 4(1), p. 203. doi: 10.1186/gb-2003-4-1-203.

Pal, M. and Luse, D. S. (2002) 'Strong Natural Pausing by RNA Polymerase II within 10 Bases of Transcription Start May Result in Repeated Slippage and Reextension of the Nascent RNA', *Molecular and Cellular Biology*, 22(1), pp. 30–40. doi: 10.1128/mcb.22.1.30-40.2002.

Park, J. H. and Roe, J. H. (2008) 'Mycothiol regulates and is regulated by a thiol-specific antisigma factor RsrA and σR in Streptomyces coelicolor', *Molecular Microbiology*, 68(4), pp. 861–870. doi: 10.1111/j.1365-2958.2008.06191.x.

Passador, L. and Linn, T. (1992) 'An internal region of rpoB is required for autogenous translational regulation of the β subunit of Escherichia coli RNA polymerase', *Journal of Bacteriology*, 174(22), pp. 7174–7179. doi: 10.1128/jb.174.22.7174-7179.1992.

Paul, B. J. *et al.* (2004) 'DksA: A Critical Component of the Transcription Initiation Machinery that Potentiates the Regulation of rRNA Promoters by ppGpp and the Initiating NTP', *Cell*, 118(3), pp. 311–322. doi: 10.1016/j.cell.2004.07.009.

Paul, B. J., Berkmen, M. B. and Gourse, R. L. (2005) 'DksA potentiates direct activation of amino acid promoters by ppGpp', *Proceedings of the National Academy of Sciences of the United States of America*, 102(22), pp. 7823–7828. doi: 10.1073/pnas.0501170102.

Pawlik, K., Kotowska, M. and Kolesiński, P. (2010) 'Streptomyces coelicolor A3(2) produces a new yellow pigment associated with the polyketide synthase Cpk', *Journal of Molecular Microbiology and Biotechnology*, 19(3), pp. 147–151. doi: 10.1159/000321501.

Perederina, A. *et al.* (2004) 'Regulation through the secondary channel - Structural framework for ppGpp-DksA synergism during transcription', *Cell*, 118(3), pp. 297–309. doi: 10.1016/j.cell.2004.06.030.

Petushkov, I. *et al.* (2017) 'Interplay between σ region 3.2 and secondary channel factors during promoter escape by bacterial RNA polymerase', *Biochemical Journal*, 474(24), pp. 4053–4064. doi: 10.1042/BCJ20170436.

Plaskon, D. M. *et al.* (2021) 'Temperature effects on RNA polymerase initiation kinetics reveal which open complex initiates and that bubble collapse is stepwise', *Proceedings of the National Academy of Sciences of the United States of America*, 118(30), pp. 1–8. doi: 10.1073/pnas.2021941118.

Prusa, J. *et al.* (2018) 'Domains within RbpA serve specific functional roles that regulate the expression of distinct mycobacterial gene subsets', *Journal of Bacteriology*, 200(13). doi: 10.1128/JB.00690-17.

Pupov, D. *et al.* (2014) 'Distinct functions of the RNA polymerase σ subunit region 3.2 in RNA priming and promoter escape', *Nucleic Acids Research*, 42(7), pp. 4494–4504. doi: 10.1093/nar/gkt1384.

Puttikhunt, C. *et al.* (1993) 'Distribution in the genus Streptomyces of a homolog to nusG, a gene encoding a transcriptional antiterminator', *FEMS Microbiology Letters*,

110, pp. 243–248. doi: 10.1016/0378-1097(93)90473-F.

Qi, F. *et al.* (1996) 'In Vitro Assay for Reiterative Transcription during Transcriptional Initiation by Escherichia coli RNA polymerase', 273(1991), pp. 71–85.

Rammohan, J. *et al.* (2015) 'CarD stabilizes mycobacterial open complexes via a two-tiered kinetic mechanism', *Nucleic Acids Research*, 43(6), pp. 3272–3285. doi: 10.1093/nar/gkv078.

Rammohan, J. *et al.* (2016) 'Cooperative stabilization of Mycobacterium tuberculosis rrnAP3 promoter open complexes by RbpA and CarD', *Nucleic Acids Research*, 44(15), pp. 7304–7313. doi: 10.1093/nar/gkw577.

Rashad, F. M. *et al.* (2015) 'Isolation and characterization of multifunctional Streptomyces species with antimicrobial, nematicidal and phytohormone activities from marine environments in Egypt', *Microbiological Research*. Elsevier GmbH., 175, pp. 34–47. doi: 10.1016/j.micres.2015.03.002.

Richards, D. M. *et al.* (2012) 'Mechanistic basis of branch-site selection in filamentous bacteria', *PLoS Computational Biology*, 8(3). doi: 10.1371/journal.pcbi.1002423.

Rigali, S. *et al.* (2008) 'Feast or famine: The global regulator DasR links nutrient stress to antibiotic production by Streptomyces', *EMBO Reports*, 9(7), pp. 670–675. doi: 10.1038/embor.2008.83.

Roberts, J. W., Shankar, S. and F (2008) 'RNA Polymerase Elongation Factors', *Annual Review of Microbiology*, 62(211), pp. 1–27. doi: 10.1016/j.currproblcancer.2009.01.001.Multiple.

Romero, D. A. *et al.* (2014) 'A comparison of key aspects of gene regulation in Streptomyces coelicolor and Escherichia coli using nucleotide-resolution transcription maps produced in parallel by global and differential RNA sequencing', *Molecular Microbiology*, 94(5), pp. 963–987. doi: 10.1111/mmi.12810.

Rosinski-Chupin, I. *et al.* (2015) 'Single nucleotide resolution RNA-seq uncovers new regulatory mechanisms in the opportunistic pathogen Streptococcus agalactiae', *BMC Genomics*, 16(1), pp. 1–15. doi: 10.1186/s12864-015-1583-4.

Ross, W. *et al.* (1993) 'A third recognition element in bacterial promoters: DNA binding by the α subunit of RNA polymerase', *Science*, 262(5138), pp. 1407–1413. doi: 10.1126/science.8248780.

Ross, W. *et al.* (2013) 'The magic spot: A ppGpp binding site on E. coli RNA polymerase responsible for regulation of transcription initiation', *Molecular Cell*. Elsevier Inc., 50(3), pp. 420–429. doi: 10.1016/j.molcel.2013.03.021.

Ross, W. *et al.* (2014) 'Polymerase Responsible for Regulation of Transcription', 50(3), pp. 420–429. doi: 10.1016/j.molcel.2013.03.021.The.

Rudd, B. A. M. and Hopwood, D. A. (1980) 'A pigmented mycelial antibiotic in Streptomyces coelicolor: control by a chromosomal gene cluster', *Journal of General Microbiology*, 119(2), pp. 333–340. doi: 10.1099/00221287-119-2-333.

Saecker, R. M. *et al.* (2002) 'Kinetic studies and structural models of the association

of E. coli σ70 RNA polymerase with the λPR promoter: Large scale conformational changes in forming the kinetically significant intermediates', *Journal of Molecular Biology*, 319(3), pp. 649–671. doi: 10.1016/S0022-2836(02)00293-0.

Saecker, R. M., Record Jr, 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.', *Artificial Intelligence in Medicine*, 412(5), p. 1118. doi: 10.1016/j.

Sanchez-Vazquez, P. *et al.* (2019) 'Genome-wide effects on Escherichia coli transcription from ppGpp binding to its two sites on RNA polymerase', *Proceedings of the National Academy of Sciences of the United States of America*, 116(17), pp. 8310–8319. doi: 10.1073/pnas.1819682116.

Santos-Beneit, F. *et al.* (2011) 'The RNA polymerase omega factor rpoZ is regulated by phop and has an important role in antibiotic biosynthesis and morphological differentiation in Streptomyces coelicolor', *Applied and Environmental Microbiology*, 77(21), pp. 7586–7594. doi: 10.1128/AEM.00465-11.

Sasaki, S. *et al.* (2001) 'Mycobacterium leprae and leprosy: A compendium', *Microbiology and Immunology*, 45(11), pp. 729–736. doi: 10.1111/j.1348-0421.2001.tb01308.x.

Saxena, S. *et al.* (2019) 'Escherichia coli transcription factor NusG binds to 70S ribosomes', *Methods Molecular Biology*, 176(5), pp. 139–148. doi: 10.1111/mmi.13953.Escherichia.

Schwartz, E. C. *et al.* (2005) 'A Full Length Group 1 Bacterial Sigma Factor Adopts a Compact Structure Incompatible with DNA Binding', *Bone*, 23(1), pp. 1–7. doi: 10.1016/j.chembiol.2008.09.008.A.

Schwedock, J. *et al.* (1997) 'Assembly of the cell division protein FtsZ into ladder-like structures in the aerial hyphae of Streptomyces coelicolor', *Molecular Microbiology*, 25(5), pp. 847–858. doi: 10.1111/j.1365-2958.1997.mmi507.x.

Sclavi, B. *et al.* (2005) 'Real-time characterization of intermediates in the pathway to open complex formation by Escherichia coli RNA polymerase at the T7A1 promoter', *Proceedings of the National Academy of Sciences of the United States of America*, 102(13), pp. 4706–4711. doi: 10.1073/pnas.0408218102.

Seliverstov, A. V. *et al.* (2005) 'Comparative analysis of RNA regulatory elements of amino acid metabolism genes in Actinobacteria', *BMC Microbiology*, 5, pp. 1–14. doi: 10.1186/1471-2180-5-54.

Seyfzadeh, M., Keener, J. and Nomura, M. (1993) 'spoT-dependent accumulation of guanosine tetraphosphate in response to fatty acid starvation in Escherichia coli', *Proceedings of the National Academy of Sciences of the United States of America*, 90(23), pp. 11004–11008. doi: 10.1073/pnas.90.23.11004.

Sheeler, N. L., MacMillan, S. V. and Nodwell, J. R. (2005) 'Biochemical activities of the absA two-component system of Streptomyces coelicolor', *Journal of Bacteriology*, 187(2), pp. 687–696. doi: 10.1128/JB.187.2.687-696.2005.

Shepherd, N., Dennis, P. and Bremer, H. (2001) 'Cytoplasmic RNA polymerase in Escherichia coli', *Journal of Bacteriology*, 183(8), pp. 2527–2534. doi:

10.1128/JB.183.8.2527-2534.2001.

Shin, Y. *et al.* (2021) 'Structural basis of ribosomal RNA transcription regulation', *Nature Communications*. Springer US, 12(1), pp. 1–12. doi: 10.1038/s41467-020-20776-y.

Shin, Y., Hedglin, M. and Murakami, K. S. (2020) 'Structural basis of reiterative transcription from the pyrG and pyrBI promoters by bacterial RNA polymerase', *Nucleic Acids Research*. Oxford University Press, 48(4), pp. 2144–2155. doi: 10.1093/nar/gkz1221.

Sigle, S. *et al.* (2015) 'Synthesis of the spore envelope in the developmental life cycle of Streptomyces coelicolor', *International Journal of Medical Microbiology*. Elsevier GmbH., 305(2), pp. 183–189. doi: 10.1016/j.ijmm.2014.12.014.

Sivapragasam, S. and Grove, A. (2019) 'The link between purine metabolism and production of antibiotics in streptomyces', *Antibiotics*, 8(2). doi: 10.3390/antibiotics8020076.

Skordalakes, E. and Berger, J. M. (2003) 'Structure of the Rho transcription terminator: Mechanism of mRNA recognition and helicase loading', *Cell*, 114(1), pp. 135–146. doi: 10.1016/S0092-8674(03)00512-9.

Sohlberg, B., Huang, J. and Cohen, S. N. (2003) 'The Streptomyces coelicolor Polynucleotide Phosphorylase Homologue, and Not the Putative Poly(A) Polymerase, Can Polyadenylate RNA', *Journal of Bacteriology*, 185(24), pp. 7273–7278. doi: 10.1128/JB.185.24.7273-7278.2003.

Sorensen, K. I. *et al.* (1993) 'Nucleotide pool-sensitive selection of the transcriptional start site in vivo at the Salmonella typhimurium pyrC and pyrD promoters', *Journal of Bacteriology*, 175(13), pp. 4137–4144. doi: 10.1128/jb.175.13.4137-4144.1993.

Springer, M. *et al.* (1985) 'Attenuation control of the Escherichia coli phenylalanyl-tRNA synthetase operon', *Journal of Molecular Biology*, 181(4), pp. 467–478. doi: 10.1016/0022-2836(85)90420-6.

Srivastava, D. B. *et al.* (2013) 'Structure and function of CarD, an essential mycobacterial transcription factor', *Proceedings of the National Academy of Sciences of the United States of America*, 110(31), pp. 12619–12624. doi: 10.1073/pnas.1308270110.

St-Onge, R. J. *et al.* (2015) 'Nucleotide Second Messenger-Mediated Regulation of a Muralytic Enzyme in Streptomyces', *Molecular Microbiology*, 96(4), pp. 779–795. doi: 10.1111/mmi.12971.

Stallings, C. L. *et al.* (2009) 'CarD Is an Essential Regulator of rRNA Transcription Required for Mycobacterium tuberculosis Persistence', *Cell*. Elsevier Ltd, 138(1), pp. 146–159. doi: 10.1016/j.cell.2009.04.041.

Stallings, C. L. and Glickman, M. S. (2011) 'CarD: A new RNA polymerase modulator in mycobacteria', *Transcription*, 2(1), pp. 15–18. doi: 10.4161/trns.2.1.13628.

Stephens, J. C., Artz, S. W. and Ames, B. N. (1975) 'Guanosine 5'-diphosphate 3'-diphosphate (ppGpp): positive effector for histidine operon transcription and general

signal for amino-acid deficiency.', *Proceedings of the National Academy of Sciences of the United States of America*, 72(11), pp. 4389–93. doi: 10.1073/pnas.72.11.4389.

Steward, K. L. and Linn, T. (1991) 'In vivo analysis of overlapping transcription units in the rplKAJLrpoBC ribosomal protein-RNA polymerase gene cluster of Escherichia coli', *Journal of Molecular Biology*, 218(1), pp. 23–31. doi: 10.1016/0022-2836(91)90870-C.

Strauch, E. *et al.* (1991) 'The stringent response in Streptomyces coelicolor A3(2).', *Molecular microbiology*, 5(2), pp. 289–298. doi: 10.1111/j.1365-2958.1991.tb02109.x.

Strauß, M. *et al.* (2015) 'The two domains of Mycobacterium tuberculosis NusG protein are dynamically independent', *Journal of Biomolecular Structure and Dynamics*, 34(2), pp. 352–361. doi: 10.1080/07391102.2015.1031700.

Stubbendieck, R. M., Vargas-Bautista, C. and Straight, P. D. (2016) 'Bacterial communities: Interactions to scale', *Frontiers in Microbiology*, 7(AUG), pp. 1–19. doi: 10.3389/fmicb.2016.01234.

Sudha Kalyani, B. *et al.* (2015) 'A NusG paralogue from Mycobacterium tuberculosis, Rv0639, has evolved to interact with ribosomal protein s10 (Rv0700) but not to function as a transcription elongation–termination factor', *Microbiology (United Kingdom)*, 161(1), pp. 67–83. doi: 10.1099/mic.0.083709-0.

Sun, J., Hesketh, A. and Bibb, M. (2001) 'Functional analysis of relA and rshA (relA/SpoT homologues) of Streptomyces coelicolor A3 (2)', *Society*, 183(11), pp. 3488–3498. doi: 10.1128/JB.183.11.3488.

Sutherland, C. and Murakami, K. S. (2018) 'An Introduction to the Structure and Function of the catalytic core enzyme of Escherichia RNA polymerase', *EcoSal Plus*, 8(1), pp. 100–106. doi: 10.1128/ecosalplus.ESP-0004-2018.An.

Sy, J. and Lipmann, F. (1973) 'Identification of the synthesis of guanosine tetraphosphate (MS I) as insertion of a pyrophosphoryl group into the 3'-position in guanosine 5'-diphosphate.', *Proceedings of the National Academy of Sciences of the United States of America*, 70(2), pp. 306–309. doi: 10.1073/pnas.70.2.306.

Tabib-Salazar, A. *et al.* (2013) 'The actinobacterial transcription factor RbpA binds to the principal sigma subunit of RNA polymerase', *Nucleic Acids Research*, 41(11), pp. 5679–5691. doi: 10.1093/nar/gkt277.

Takizawa, M., Colwell, R. R. and Hill, R. T. (1993) 'Isolation and diversity of actinomycetes in the Chesapeake Bay', *Applied and Environmental Microbiology*, 59(4), pp. 997–1002.

Tan, H. *et al.* (1998) 'The Streptomyces coelicolor sporulation-specific σ(WhiG) form of RNA polymerase transcribes a gene encoding a ProX-like protein that is dispensable for sporulation', *Gene*, 212(1), pp. 137–146. doi: 10.1016/S0378-1119(98)00152-8.

Tezuka, T. and Ohnishi, Y. (2014) 'Two glycine riboswitches activate the glycine cleavage system essential for glycine detoxification in Streptomyces griseus', *Journal of Bacteriology*, 196(7), pp. 1369–1376. doi: 10.1128/JB.01480-13.

Toulokhonov, I., Artsimovitch, I. and Landick, R. (2001) 'Allosteric control of RNA

polymerase by a site that contacts nascent RNA hairpins', *Science*, 292(5517), pp. 730–733. doi: 10.1126/science.1057738.

Travis, B. A. and Schumacher, M. A. (2022) 'Diverse molecular mechanisms of transcription regulation by the bacterial alarmone ppGpp', *Molecular Microbiology*, 117(2), pp. 252–260. doi: 10.1111/mmi.14860.

Tripathi, L., Zhang, Y. and Lin, Z. (2014) 'Bacterial sigma factors as targets for engineered or synthetic transcriptional control', *Frontiers in Bioengineering and Biotechnology*, 2(SEP), pp. 1–7. doi: 10.3389/fbioe.2014.00033.

Turnbough, C. L. (1992) 'Regulation of bacterial gene expression', *Reviews in Medical Microbiology*, 3(4), pp. 217–226. doi: 10.1007/978-1-4613-0803-4_6.

Turnbough, C. L. (2008) 'Regulation of bacterial gene expression by the NTP substrates of transcription initiation', *Molecular Microbiology*, 69(1), pp. 10–14. doi: 10.1111/j.1365-2958.2008.06272.x.

Turnbough, C. L. (2011) 'Regulation of gene expression by reiterative transcription', *Current Opinion in Microbiology*, 14(2), pp. 142–147. doi: 10.1016/j.mib.2011.01.012.

Turnbough, C. L. and Switzer, R. L. (2008) 'Regulation of Pyrimidine Biosynthetic Gene Expression in Bacteria: Repression without Repressors', *Microbiology and Molecular Biology Reviews*, 72(2), pp. 266–300. doi: 10.1128/mmbr.00001-08.

Ul-Hassan, A. and Wellington, E. M. (2009) *Bacteria*.

Urem, M. *et al.* (2016) 'Intertwining nutrient-sensory networks and the control of antibiotic production in Streptomyces', *Molecular Microbiology*, 102(2), pp. 183–195. doi: 10.1111/mmi.13464.

Vassylyev, D. G. *et al.* (2007) 'Structural basis for transcription elongation by bacterial RNA polymerase', *Nature*, 448(7150), pp. 157–162. doi: 10.1038/nature05932.

Ventura, M. *et al.* (2007) 'Genomics of Actinobacteria: tracing the evolutionary history of an ancient phylum.', *Microbiology and molecular biology reviews : MMBR*, 71(3), pp. 495–548. doi: 10.1128/MMBR.00005-07.

Vinella, D. *et al.* (2005) 'Iron limitation induces SpoT-dependent accumulation of ppGpp in Escherichia coli', *Molecular Microbiology*, 56(4), pp. 958–970. doi: 10.1111/j.1365-2958.2005.04601.x.

Vuthoori, S. *et al.* (2001) 'Domain 1.1 of the σ70 subunit of Escherichia coli RNA polymerase modulates the formation of stable polymerase/promoter complexes', *Journal of Molecular Biology*, 309(3), pp. 561–572. doi: 10.1006/jmbi.2001.4690.

Vvedenskaya, I. O. *et al.* (2015) 'Massively Systematic Transcript End Readout, "MASTER": Transcription Start Site Selection, Transcriptional Slippage, and Transcript Yields', *Molecular Cell*. Elsevier Inc., 60(6), pp. 953–965. doi: 10.1016/j.molcel.2015.10.029.

Wang, G., Hosaka, T. and Ochi, K. (2008) 'Dramatic activation of antibiotic production in Streptomyces coelicolor by cumulative drug resistance mutations', *Applied and Environmental Microbiology*, 74(9), pp. 2834–2840. doi:

10.1128/AEM.02800-07.

Watve, M. G. *et al.* (2001) 'How many antibiotics are produced by the genus Streptomyces?', *Archives of Microbiology*, 176(5), pp. 386–390. doi: 10.1007/s002030100345.

Wels, M. *et al.* (2008) 'An in silico analysis of T-box regulated genes and T-box evolution in prokaryotes, with emphasis on prediction of substrate specificity of transporters', *BMC Genomics*, 9. doi: 10.1186/1471-2164-9-330.

Van Wezel, G. P. and McDowall, K. J. (2011) 'The regulation of the secondary metabolism of Streptomyces: New links and experimental advances', *Natural Product Reports*, 28(7), pp. 1311–1333. doi: 10.1039/c1np00003a.

Wildermuth, H. and Hopwood, D. A. (1970) 'Septation during sporulation in Streptomyces coelicolor.', *Journal of general microbiology*, 60(1), pp. 51–59. doi: 10.1099/00221287-60-1-51.

Winkelman, J. T. *et al.* (2016) 'Open complex scrunching before nucleotide addition accounts for the unusual transcription start site of E. Coli ribosomal RNA promoters', *Proceedings of the National Academy of Sciences of the United States of America*, 113(13), pp. E1787–E1795. doi: 10.1073/pnas.1522159113.

Wright, L F and Hopwood, D. A. (1976) 'Actinorhodin is a Chromosomally-determined Antibiotic in Streptomyces coelicolor A3 (2)', *Journal of General Microbiology*, 96, pp. 289–297.

Wright, L. F. and Hopwood, D. A. (1976) 'Identification of the antibiotic determined by the SCP1 plasmid of Streptomyces coelicolor A3(2)', *Journal of General Microbiology*, 95(1), pp. 96–106. doi: 10.1099/00221287-95-1-96.

Xiong, X. and Reznikoff, W. (1993) 'Transcriptional Slippage During the Transcription Initiation Process at a Mutant lac Promoter in vivo', *Journal of Molecular Biology*, 231, pp. 569–580.

Xu, J. *et al.* (2002) 'A rifampicin resistance mutation in the rpoB gene confers ppGpp-independent antibiotic production in Streptomyces coelicolor A3(2)', *Molecular Genetics and Genomics*, 268(2), pp. 179–189. doi: 10.1007/s00438-002-0730-1.

Yadav, P. *et al.* (2021) 'Clinical Features, Radiological Findings, and Treatment Outcomes in Patients with Pulmonary Nocardiosis: A Retrospective Analysis', *Cureus*, 13(8). doi: 10.7759/cureus.17250.

Yakhnin, A. V. *et al.* (2020) 'NusG controls transcription pausing and RNA polymerase translocation throughout the Bacillus subtilis genome', *Proceedings of the National Academy of Sciences of the United States of America*, 117(35), pp. 21628–21636. doi: 10.1073/pnas.2006873117.

Yakhnin, A. V., Murakami, K. S. and Babitzke, P. (2016) 'NusG is a sequence-specific RNA polymerase pause factor that binds to the non-template DNA within the paused transcription bubble', *Journal of Biological Chemistry*, 291(10), pp. 5299–5308. doi: 10.1074/jbc.M115.704189.

Yakhnin, A. V., Yakhnin, H. and Babitzke, P. (2008) 'Function of the Bacillus subtilis transcription elongation factor NusG in hairpin-dependent RNA polymerase

pausing in the trp leader', *Proceedings of the National Academy of Sciences of the United States of America*, 105(42), pp. 16131–16136. doi: 10.1073/pnas.0808842105.

Yakhnin, H. *et al.* (2019) 'Nusg-dependent rna polymerase pausing and tylosin-dependent ribosome stalling are required for tylosin resistance by inducing 23s rrna methylation in bacillus subtilis', *mBio*, 10(6), pp. 1–14. doi: 10.1128/mBio.02665-19.

Yamamoto, M. and Nomura, M. (1978) 'Cotranscription of genes for RNA polymerase subunits β and β' with genes for ribosomal proteins in Escherichia coli', *Proceedings of the National Academy of Sciences of the United States of America*, 75(8), pp. 3891–3895. doi: 10.1073/pnas.75.8.3891.

Young, B. A. *et al.* (2001) 'A coiled-coil from the RNA polymerase β′ subunit allosterically induces selective nontemplate strand binding by σ70', *Cell*, 105(7), pp. 935–944. doi: 10.1016/S0092-8674(01)00398-1.

Yu, L. *et al.* (2017) 'The mechanism of variability in transcription start site selection', *eLife*, 6(iii), pp. 1–22. doi: 10.7554/eLife.32038.

Zachrdla, M. *et al.* (2017) 'Solution structure of domain 1.1 of the σA factor from Bacillus subtilis is preformed for binding to the RNA polymerase core', *Journal of Biological Chemistry*, 292(28), pp. 11610–11617. doi: 10.1074/jbc.M117.784074.

Zhang, G. *et al.* (1999) 'Crystal structure of thermus aquaticus core RNA polymerase at 3.3 å resolution', *Cell*, 98(6), pp. 811–824. doi: 10.1016/S0092-8674(00)81515-9.

Zhang, J. and Landick, R. (2016) 'A Two-Way Street: Regulatory Interplay between RNA Polymerase and Nascent RNA Structure', *Trends in Biochemical Sciences*. Elsevier Ltd, 41(4), pp. 293–310. doi: 10.1016/j.tibs.2015.12.009.

Zhang, N. and Buck, M. (2015) 'A perspective on the enhancer dependent bacterial RNA polymerase', *Biomolecules*, 5(2), pp. 1012–1019. doi: 10.3390/biom5021012.

Zhang, Y. *et al.* (2005) 'DksA Guards Elongating RNA Polymerase Against Ribosome- Stalling-Induced Arrest', *Bone*, 23(1), pp. 1–7. doi: 10.1016/j.molcel.2014.02.005.DksA.

Zhang, Y. *et al.* (2012) 'Structural basis of transcription initiation', *Science*. Springer US, 338(6110), pp. 1076–1080. doi: 10.1126/science.1227786.

Zhou, Y. N. *et al.* (2013) 'Isolation and characterization of RNA polymerase rpoB mutations that alter transcription slippage during elongation in Escherichia coli', *Journal of Biological Chemistry*, 288(4), pp. 2700–2710. doi: 10.1074/jbc.M112.429464.

Zhu, D. X. *et al.* (2019) 'CarD contributes to diverse gene expression outcomes throughout the genome of Mycobacterium tuberculosis', *Proceedings of the National Academy of Sciences of the United States of America*, 116(27), pp. 13573–13581. doi: 10.1073/pnas.1900176116.

Zhu, J. H. *et al.* (2018) 'Rifampicin can induce antibiotic tolerance in mycobacteria via paradoxical changes in rpoB transcription', *Nature Communications*. Springer US, 9(1). doi: 10.1038/s41467-018-06667-3.

Zuo, Yuhong, Wang, Y. and Steitz, T. A. (2013) ‘The Mechanism of E. coli RNA Polymerase Regulation by ppGpp Is Suggested by the Structure of Their Complex’, 50(3), pp. 430–436. doi: 10.1016/j.molcel.2013.03.020.The.

Zuo, Y. and Steitz, T. A. (2015) ‘Crystal Structures of the E. coli Transcription Initiation Complexes with a Complete Bubble’, *Physiology & behavior*, 58(3), pp. 534–540. doi: 10.1016/j.molcel.2015.03.010.Crystal.