Transcription Initiation in *Streptomyces coelicolor* A3(2)

Laurence J. Humphrey

Thesis submitted for the degree of Doctor of Philosophy

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

September 2015

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.

.....

Laurence J. Humphrey

Abstract

Recent studies into the stringent response and the discovery of a number of RNA polymerase binding proteins suggests that the model for bacterial transcription initiation in Actinobacteria may differ from that in *Escherichia coli*. In *E. coli*, the alarmone ppGpp, together with DksA, binds to RNA polymerase to elicit the stringent response. However, the ppGpp binding site on RNA polymerase is not conserved in *S. coelicolor*, although the organism possesses a DksA homologue. Deletion of DksA did not affect the growth and development of *S. coelicolor*, although its overexpression stimulated antibiotic production. Evidence is presented that suggests that this occurs through binding to the RNA polymerase secondary channel. The biological role of this protein remains unknown. CarD and RbpA are two RNA polymerase-binding proteins present in all Actinobacteria, including *S. coelicolor* and *M. tuberculosis*. Both proteins are critical for growth and have been identified as transcriptional activators from σ^{HrdB}-dependent promoters *in vitro*. Here it was demonstrated that CarD and RbpA activate transcription from rRNA promoters with a poorly conserved -35 element. Surprisingly it was also found that both proteins can inhibit transcription from synthetic promoters with highly conserved -35 elements. Chromatin immunoprecipitation followed by high throughput sequencing (ChIP-seq) experiments revealed that CarD and RbpA are found exclusively at promoter regions. RbpA is localised only at promoters recognised by σ^{HrdB}, whereas CarD also co-localises with the alternative sigma factor σ^{R} during oxidative stress indicating that it lacks RNA polymerase holoenzyme specificity. The sigma specificity of RbpA was tested by the generation of sigma mutants that were defective in binding. *In vivo, in vitro* and ChIP-seq data presented in this study suggest that CarD and RbpA have an overlapping role in transcription initiation at σ^{HrdB}-dependent promoters in *S. coelicolor*.

Acknowledgements

First and foremost, I would like to thank my supervisor Dr Mark Paget for his excellent guidance, patience and support throughout this PhD project. I really believe he has helped me develop as a scientist both technically and mentally, and I will sincerely miss working with him as a member of his research group.

I would like to extend my thanks to members past and present of the Paget lab, especially Dr Richard Lewis and Dr Aline Tabib-Salazar for their guidance and collaboration on a number of projects, and Heather Macklyne and Heena Jagatia for their support. Outside of our research group I would like to thank my second supervisor Dr Neil Crickmore, Sue Searle for her technical support to the lab, and members of the Morley lab for helping me through this final year.

Outside of the scientific bubble, I would like to thank my fellow PhD students at Sussex, mostly for getting me out of the lab at 5pm on a Friday; my family and friends from Birmingham, Macclesfield, Surrey, Finland, Brighton and further afield; and my team mates from the various hockey and running clubs - you've all helped keep me going in one way or another.

Finally, I would like to thank my family Rosie, Maurice, Nick and Tessa for everything they have done for me. From childhood summer science projects to more recently late night phone calls listening to me rant (usually about misbehaving cultures or disintegrating acrylamide gels), they have been unwavering in their support, belief and encouragement and I would not have got this far without them.

Publications

Hubin, E.A., Tabib-Salazar, A., Humphrey, L.J., Flack, J.E., Olinares, P.D.B., Darst, S.A., Campbell, E.A., Paget, M.S (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 112:7171–7176.

Table of Contents

Table of Figures

Abbreviations

2-DOG: 2-deoxyglucose

aa: amino acids

BACTH: bacterial two-hybrid

BLAST: basic local alignment search tool

bp: base pairs

BSA: bovine serum albumin

CFU: colony-forming unit

ChIP: chromatin immunoprecipitation

CTD: C-terminal domain

dH_2O: distilled water

ECF: extracytoplasmic function

EtOH: ethanol

FPLC: fast protein liquid chromatography

GC: guanine-cytosine

gDNA: genomic DNA

GDP: guanosine diphosphate

GTP: guanosine triphosphate

h: hours

IGB: Integrated Genome Browser

IMG: Integrated Microbial Genomes

IPTG: Isopropyl β-D-1-thiogalactopyranoside

kb: kilobase pairs

kDa: kilodalton

MCS: multi-cloning site

min: minutes

MS: mannitol soya flour

MW: molecular weight

NCBI: National Centre for Biotechnology Information

nt: nucleotides

NTD: N-terminal domain

NTP: nucleoside triphosphate

OD: optical density

P:C:IAA: phenol:chloroform:isoamyl alcohol

PCR: polymerase chain reaction

PMSF: phenylmethylsulfonyl fluoride

qPCR: quantitative polymerase chain reaction

RNAP: RNA polymerase

RPc: RNA polymerase closed complex

RPo: RNA polymerase open complex

RT: room temperature

s: seconds

SDS-PAGE: Sodium dodecyl sulphate polyacrylamide gel electrophoresis

SID: sigma-interaction domain

spp.: species

TB: tuberculosis

TIR: terminal inverted repeat

UTR: untranslated region

WT: wild-type

X-gal: 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

Chapter 1:

Introduction

1 Introduction

1.1 Actinobacteria

The Actinobacteria, one of the largest phyla of bacteria, are characterised as Gram positives with a high guanine-cytosine (GC) content ranging from 51% to over 70% in some species of *Streptomyces* and *Frankia* (Ventura et al., 2007). The Actinobacteria are highly pleomorphic, with morphologies ranging from coccoid forms (e.g. *Micrococci*) to fragmenting, rod-shaped forms (e.g. *Nocardia*) to branching, filamentous forms (e.g. *Streptomyces*) (Tortora et al., 2012). They are found in a diverse range of environments from soil and aquatic environments (e.g. *Streptomyces*, *Rhodococci*) as well as pathogens (e.g. *Mycobacteria*, *Corynebacteria*, *Nocardia*) and gastrointestinal commensals (e.g. *Bifidobacteria*).

1.1.1 *Streptomyces*

Within the Actinobacteria exists a group of mycelial bacteria including *Streptomyces*, *Frankia*, *Nocardia* and *Actinomyces* species informally known as the actinomycetes, due to their branching, filamentous morphology. The most studied and best-characterised members of this group are the soil-dwelling streptomycetes.

Unlike the vast majority of bacteria, *Streptomyces* exhibit a complex multicellular lifestyle. For a number of years, the true phylogeny of *Streptomyces* genus was unknown. Their lifestyle is similar to filamentous fungi and they were originally considered an intermediate between fungi and bacteria (Hopwood, 1999). Evidence of this misunderstanding is still present in the name "*Streptomyces*" which translates to "twisted or chain-like fungus". Closer observations of the organism, for instance the discovery of a bacterial cell wall and the susceptibility to a number of antibacterial compounds, began to reveal its true phylogeny as a bacterium. The issue was concluded following electron microscopy studies on germinated *Streptomyces* spores revealing the absence of a nuclear membrane making them by definition prokaryotes (Hopwood and Glauert, 1960).

One reason for such interest in this group of bacteria is their ability to thrive in a challenging environment as complex and variable as soil. They are an essential member of the ecosystem due to their ability to degrade insoluble compounds such as chitin and

lignocellulose from the remains of other organisms. This can be attributed to their ability to secrete proteins such as extracellular enzymes and import the degraded products as a major source of nutrients (reviewed by Chater et al. 2010). In addition to these systems, they are metabolically highly adapted to exist in an oligotrophic environment. Soil is typically rich in carbon but nitrogen and phosphate-poor, and streptomycetes are adapted to exploit this (Hodgson, 2000). For example, streptomycetes typically invest in a number of carbohydrate catabolic pathways which reflects the presence of multiple carbon sources within the soil (Hodgson, 2000).

Another feature of streptomycetes that allow them to survive in this challenging environment is their ability to produce and secrete an array of bioactive compounds or "antibiotics" that may inhibit the growth of competing organisms and provide a selective advantage in the local environment. This phenomenon, first identified by Alexander Fleming in his study of *Penicillium* bacteria and the effects of beta-lactam antibiotics, was further studied in actinomycete species by Selman Waksman. Unlike Fleming, who discovered the inhibitory properties of penicillin by chance after accidental contamination of a *Staphylococcus* plate by an airborne mould, Waksman took a more systematic approach, screening the inhibitory effect of thousands of soil isolates on pathogenic bacteria (Waksman and Lechevalier, 1951). Through this approach, Waksman and his colleagues at Rutgers University isolated as many as 15 novel antibiotics throughout the 1940s. The first antibiotic to be isolated through this method was actinomycin from the strain *Actinomyces antibioticus* (Waksman and Woodruff, 1941). Further work yielded the discovery of streptomycin from *Streptomyces griseus*, the first antibiotic active against tuberculosis for which Waksman won a Nobel prize (Schatz et al., 1944). Since this work started by Waksman, *Streptomyces* species have been responsible for the discovery of up to two thirds of the clinically useful natural product antibiotics in use today and consequently this remains one of the major reasons for research into the genus (Bentley et al., 2002).

In addition to these compounds secreted by streptomycetes and other actinomycetes, they have also been identified as the producers of the potent chemical geosmin that gives soil its "earthy" odour (Gerber and Lechevalier, 1965). Whilst a number of theories exist for the importance of this chemical, its true function remains unknown.

Streptomyces coelicolor A3(2)

Streptomyces coelicolor (A3)2 is the most widely studied and the model organism for the *Streptomyces* genus. First genetic studies on the isolate were published in the 1950s, originally characterised and named "coelicolor" after its ability to produce a bright blue pigment, later identified as the polyketide antibiotic actinorhodin (Sermonti and Spada-Sermonti, 1955) (Figure 1.1A). In 2002 its 8.7 Mb genome was sequenced (Bentley et al., 2002). The genome has a high GC content (72.1 %) predicted to encode 7,825 coding sequences on a single linear chromosome. This linear chromosome has a centrally located origin of replication (*oriC*) and terminal inverted repeats (TIRs) with covalently bound terminal proteins (TPs) at the 5' ends of the chromosome (Bentley et al., 2002). In addition to this large linear chromosome the organism possesses two additional plasmids: the 356 kb linear SCP1 and the 31 kb circular SCP2 (Kinashi and Shimaji-Murayama, 1991; Haug et al., 2003). The most commonly used strain, M145, is a prototrophic derivative of *S. coelicolor* A3(2) lacking both SCP1 and SCP2 plasmids (Chater et al., 1982).

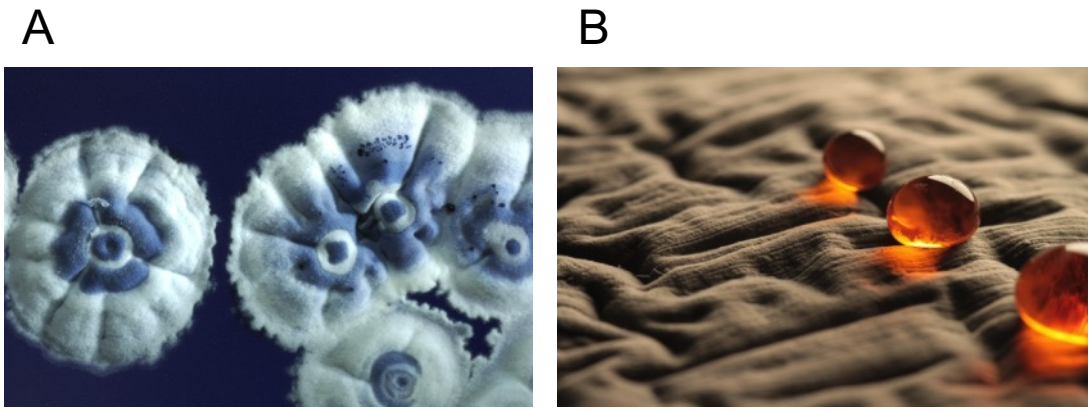


Figure 1.1 - **Colonies of *S. coelicolor* producing actinorhodin and coelimycin**. (A) Colonies of *S. coelicolor* producing actinorhodin (ACT). Photo by Mervyn Bibb and Andrew Davis (John Innes Centre). (B) Close up of *S. coelicolor* producing coelimycin (CPK). Photo by Marco Gottelt (University of Groningen).

S. coelicolor life cycle and regulation

The life cycle of *S. coelicolor* (and indeed most streptomycetes) begins with formation of a vegetative, multinucleate mycelium that colonises the local environment. Once a feeding substrate has been established, the mycelium differentiates into aerial hyphae that grow upwards from the surface into the air. These hyphae develop into coils, followed by septation, which leads to production of unigenomic, reproductive spores. Following environmental distribution, these spores can give rise to new mycelial networks and the streptomycete life cycle continues.

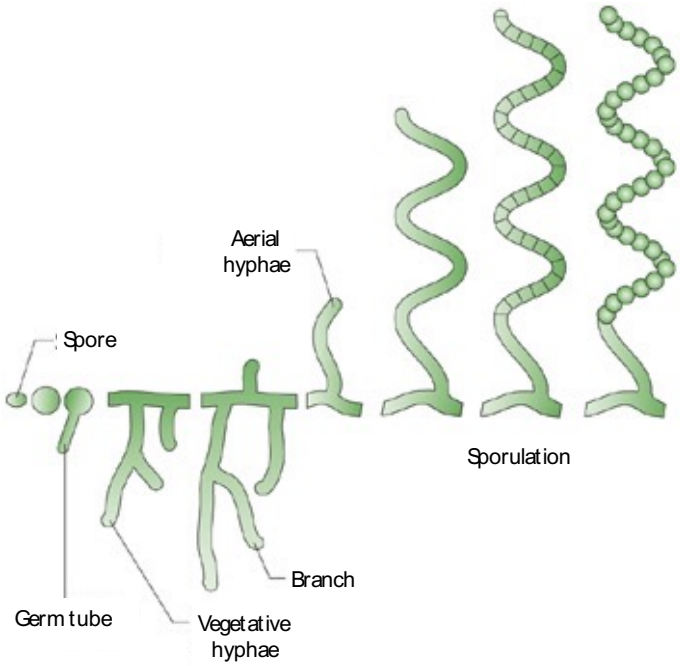


Figure 1.2 - **Developmental life cycle of *S. coelicolor*.** Illustration of the developmental life cycle of *S. coelicolor*, showing the germination of a spore, production of a vegetative and aerial hyphae through tip extension and branching, and the process of sporulation. Figure adapted from Flärdh and Buttner (2009), copyright license number: 3700341005397.

When a spore encounters conditions suitable for growth, germination occurs leading to formation of one or two germ tubes per spore. Each germ tube grows by tip extension and branching leading to formation of a vegetative, mycelial network. This type of growth, starkly different from typical cell division observed in model bacteria such as *E. coli* and *B. subtilis*, is a defining feature of streptomycetes. Notably, streptomycete cell division does not require the bacterial actin homologue MreB for formation of a new cell wall and elongation (Flärdh and Buttner, 2009). In contrast, streptomycete cell growth occurs at the cell poles by tip extension, dependent on the essential protein DivIVA

(Hempel et al., 2008). In order to grow apically, meaning unevenly from one pole, the cell must localize the necessary components for cell wall biosynthesis to that particular locus. DivIVA acts as the key factor that localises to the cell poles, interacts with the cell membrane and acts as a recruitment factor for the peptidoglycan biosynthetic machinery among other proteins which direct cell growth (Flärdh et al., 2012). The importance of this protein was particular evident during protein overexpression experiments which led to irregular growth including multiple sites of tip growth and hyper-branching (Flärdh, 2003; Hempel et al., 2008).

Formation of aerial hyphae is complex and highly regulated by a series of genes known as the *bld* cascade. The 12 *bld* genes were named as "bald" following mutagenesis studies identifying their essentiality in formation of the aerial hyphae that give *S. coelicolor* colonies their "fuzzy" colony morphology. SapB (spore associated protein B) is a surfactant peptide absolutely required by *S. coelicolor* to break the surface tension of the medium and form aerial hyphae (Flärdh and Buttner, 2009). It is a 21-amino acid peptide produced from the *ramCSAB* operon and is often described as "lantibiotic-like" due to its similarities to the class of ribosomally-synthesised oligopeptide antibiotics (Kodani et al., 2004). Despite this, it seems to have no antimicrobial functionality but acts solely as a bio-surfactant secreted by *S. coelicolor* that coats the hyphae in a hydrophobic sheath. As evidence of its importance in formation of aerial hyphae and sporulation, mutants unable to produce SapB have a pronounced bald phenotype under all growth conditions (Capstick et al., 2007). Whilst all *bld* genes are required for production of SapB, overexpression of *ramR*, the activator of SapB production, complements all *bld* mutants (Willey et al., 1991; Nguyen et al., 2002). Surprisingly, different *bld* mutants grown in proximity to one another are able to complement each other in a phenomenon known as "*bld-bld* extracellular complementation" (Willey et al., 1991). Together, these data have been used to develop a model for an extracellular signalling cascade that ultimately controls the production of SapB, inducing formation of aerial hyphae.

Following the formation of aerial hyphae, further differentiation occurs in the process of sporulation. Similar to the *bld* cascade, a number of the genes identified as essential for sporulation are named *whi* following mutagenesis studies where colonies appeared white, arrested in the later stages of aerial hyphae formation and absent of grey spore

pigment. These *whi* genes can be further divided into "early-*whi* genes", of which mutants cannot form pre-spore compartments, and "late-*whi* genes" of which mutant spores do not fully divide or produce grey spore pigment. *whiG* is the earliest gene involved in the process of sporulation which encodes an alternative sigma factor, σ^{WhiG}. Activation of this sigma factor is considered the key step for commitment to sporulation from aerial hyphae.

Antibiotic production and regulation

Streptomyces species are responsible for producing over two-thirds of natural product antibiotics used in modern medicine today, as well many other compounds used as anti-fungals, chemotherapeutics and immunosuppressants (Bentley et al., 2002). In a period of discovery in the 1940s-1960s known as the golden age of antibiotics, massive screening programmes isolated as many 12,000 bioactive compounds, with 160 reaching clinical markets (Marinelli, 2009). Notable antibiotics isolated from *Streptomyces* and related species include streptomycin, neomycin, vancomycin and chloramphenicol, with diverse targets and mechanisms of action including inhibition of protein, cell wall and DNA/RNA synthesis. Screening bioactive compounds for cytotoxity identified a number of compounds that are effective as anti-cancer chemotherapy agents including doxorubicin, bleomycin and mitomycin C, due to their ability to bind or damage DNA.

In producing strains, antibiotic biosynthetic genes are found throughout genomes organised in large clusters, tens of kilobases in size. Conserved features and organisation of such clusters has enabled mining and discovery of otherwise unexpressed, "silent" biosynthetic operons (Challis, 2008). To activate these otherwise unexpressed clusters and for increasing production yields for industrial purposes, much research has been invested into understanding the regulation of antibiotic production.

As the model organism for the streptomycetes, *S. coelicolor* produces five structurally diverse antibiotics (Figure 1.3). The blue-pigmented polyketide actinorhodin (ACT) (Wright and Hopwood, 1976b), the red-pigmented, tripyrrole undecylprodigiosin (RED) (Rudd and Hopwood, 1980), the nonribosomal lipopeptide calcium-dependent antibiotic (CDA) (Hopwood and Wright, 1983), the cyclopentanone methylenomycin (MM) encoded on the plasmid SCP1 (Wright and Hopwood, 1976a) and the recently discovered cryptic polyketide coelimycin P1 (CPK) (Pawlik et al., 2010) (Figure 1.1B).

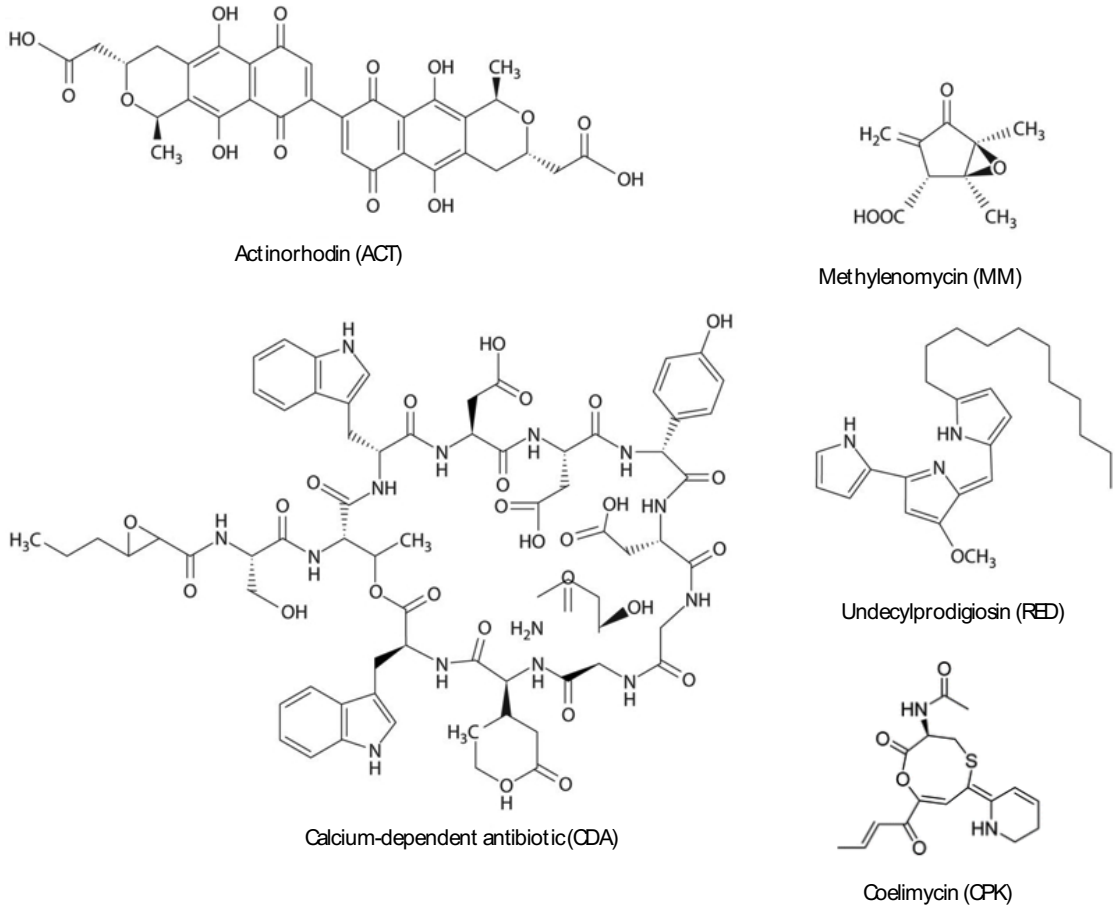


Figure 1.3 - **Structure of five antibiotics produced by *S. coelicolor* A3(2).** Structures of actinorhodin (ACT), methylenomycin (MM), undecylprodigiosin (RED), calcium-dependent antibiotic (CDA) and coelimycin (CPK). Figure modified (with addition of coelimycin structure) with permission from Liu et al. (2013) Microbiol. Mol. Biol. Rev. vol. 77 no. 1 112-143.

One feature conserved in antibiotic biosynthetic operons is the presence of a cluster-situated regulator (CSR) which acts as a specific activator for transcription of the operon and subsequent antibiotic production. In *S. coelicolor,* actII-ORF4, redD and *cdaR* serve as the CSR for ACT, RED and CDA synthesis, respectively. The regulation of CSR transcription, translation or protein activity serves as a single point of regulation for biosynthesis, over which a number of conditions, pathways and global regulators can influence.

Control of ACT production has been widely studied and currently serves as the best model, revealing the true complexity of regulation of antibiotic biosynthesis. Transcription of actII-ORF4 is a target of at least eight known regulatory proteins (Liu et al., 2013). DasR, a transcriptional repressor regulated by carbon and nitrogen source N-acetylglucosamine (GlcNAc), represents an elegant example of nutritional control over

actII-ORF4 (Rigali et al., 2008). When GlcNAc enters the cytoplasm, it is phosphorylated and deacetylated to form GlcN-6P, an allosteric inhibitor of DasR. As DasR represses actII-ORF4, the presence of N-acetylglucosamine results in derepression of actII-ORF4 and stimulates actinorhodin synthesis. Curiously and inconsistent with the known DasR system, this effect is only observed on minimal media, whilst presence of GlcNAc on rich media has the opposite effect (Rigali et al., 2008).

Other known positive and negatives regulators that directly feed into regulation of ACT production through actII-ORF4 include AbsA2, AdpA, LexA, DraR, AfsQ1, AtrA, GlnR and ROK7B7 (reviewed by Liu et al., 2013).

1.1.2 *Mycobacteria*

The mycobacteria are another well-studied genus of bacteria within the actinobacteria. They are strictly aerobic and typically rod-shaped although, as the prefix "myco-" suggests, can sometimes exhibit filamentous fungus-like morphologies (Tortora et al., 2012). Whilst mycobacteria are Gram-positive bacteria, they are characteristically resistant to Gram staining due to a distinctive lipid-rich cell wall and are consequently known as "acid-fast" bacteria (Bloch, 1953). This cell wall structure also makes this group of bacteria resistant to acids, bases, a number of other cellular stresses and is largely the reason why mycobacteria are resistant to many antibiotics (Hett and Rubin, 2008). Whilst mycobacterial species can exist in a number of different environments including soil and water sources, the better known and researched members of the genus are human and animal pathogens including *M. tuberculosis* and *M. leprae* that cause tuberculosis and leprosy, respectively. As these strains are pathogenic, slow growing and difficult to work with in a laboratory setting, the non-pathogenic, faster growing species *M. smegmatis* is often used a model organism for the genus (Reyrat and Kahn, 2001).

Mycobacterium tuberculosis

Mycobacterium tuberculosis was first identified as the causative agent of tuberculosis (TB) by Robert Koch in 1882 (Koch, 1882). Prior to this research, he established four criteria for identifying the microorganism responsible for a disease, known as "Koch's Postulates". He subsequently tested his four postulates on guinea pigs to prove that tuberculosis was caused by *M. tuberculosis*. For this ground-breaking research, he was

awarded the Nobel Prize in Physiology or Medicine in 1905 and it is one of the reasons he is considered the "founder of modern bacteriology".

In 1998, the 4.4 Mb *M. tuberculosis* H37Rv genome was sequenced (Cole et al., 1998). At the time of its publication, it was the second largest bacterial genome sequenced after *E. coli* K-12. It has a high GC content (65.6%) and is predicted to encode 3,924 coding sequences on a single circular chromosome. Most notably, the genome contains up to 250 genes responsible for lipid metabolism.

Tuberculosis is one of the world's greatest infectious killers, second only to HIV/AIDS (World Health Organisation, 2015). It is responsible for almost 9 million new cases and 1.7 million deaths per year. Pulmonary infection typically occurs following inhalation of droplets containing *M. tuberculosis,* spread from other infected humans, which colonise the lungs. Following an immune response, *M. tuberculosis* is able to survive in macrophages due to a number of survival mechanisms including inhibition of phagosome acidification and fusion with the lysosome (reviewed by Russell, 2007). Production of an inflammatory response by the infected macrophage recruits additional immune cells and results in formation of a granuloma, a hallmark of *M. tuberculosis* infection.

When diagnosed and where available, tuberculosis can be treated with antibiotics. The discovery of streptomycin was a landmark event, making tuberculosis, often recalcitrant to many antibiotics, a treatable condition. Nowadays, tuberculosis is managed with a combination of first line drugs including ethambutol, isoniazid, pyrazinamide, rifampicin and streptomycin (Zumla et al., 2013). As *M. tuberculosis* has a very slow growth rate, doubling every 24 h, effective treatment with these drugs may take months. In recent years, there has been an emergence of multi-drug resistant TB (MDR-TB), characterised as resistant to at least isoniazid and rifampicin, and extensively-drug resistant TB (XDR-TB), with additional resistance to fluoroquinolones and at least one second line injectable drug such as amikacin, kanamycin, or capreomycin (World Health Organization, 2010).

1.2 Bacterial transcription

Transcription, a key component of the "central dogma of molecular biology", is the transfer of information from DNA to RNA. It is the fundamental first step in the process of gene expression and is evolutionarily conserved across all known life forms. This process of RNA synthesis is catalysed by the multi-subunit enzyme RNA polymerase (RNAP). Unlike in eukaryotes, prokaryotic transcription occurs in the cytoplasm and therefore can occur simultaneously with translation. It is a process that is highly regulated by both intra- and extracellular signals and is consequently regarded as the major focus of lifestyle regulation in bacteria.

1.2.1 Promoters

Upstream of every RNA-coding region lies a section of DNA known as a promoter, responsible for binding RNA polymerase and initiating transcription. Among a number of features intrinsic to the DNA sequence, bacterial promoters contain two 6-bp elements known as the -10 and -35 elements. The -10 element is located approximately 7 bp upstream from the +1 (the first transcribed base) and is recognised by σ regions 2.3-2.4 (see 1.2.3 Sigma Factors) (Dombroski et al., 1992; Feklistov and Darst, 2011). The -35 element is located 17 bp upstream of the -10 element and is recognised by σ region 4.2 (Dombroski et al., 1992). The -10 and -35 consensus sequences recognised by the *E. coli* primary sigma factor σ^{70} have been identified as TATAAT and TTGACA, respectively (Hawley and McClure, 1983; Harley and Reynolds, 1987). Whilst consensus sequences vary between primary and alternative sigma promoters, studies in a diverse range of organisms have revealed that the -10 and -35 consensus sequence for primary sigma factors is largely conserved throughout eubacteria (Helmann, 1995; Rodrigue et al., 2006). In addition to recruiting RNA polymerase, features of the promoter including the "spacer" between the -10 and -35 and elements upstream of the -35 such as the "UP element" may have an impact on transcription initiation efficiency and consequently may have a regulatory function (Haugen et al., 2008a).

1.2.2 RNA polymerase

RNA polymerase is the central component of transcription which catalyses the synthesis of RNA. In bacteria, core RNAP, capable in principle of initiating transcription at any point on a DNA molecule, is composed of five subunits – two alpha subunits (α_I and α_{II}, encoded by *rpoA*), beta (β, encoded by *rpoB*), beta prime (β', encoded by *rpoC*) and omega (ω, encoded by *rpoZ*) (Figure 1.4).

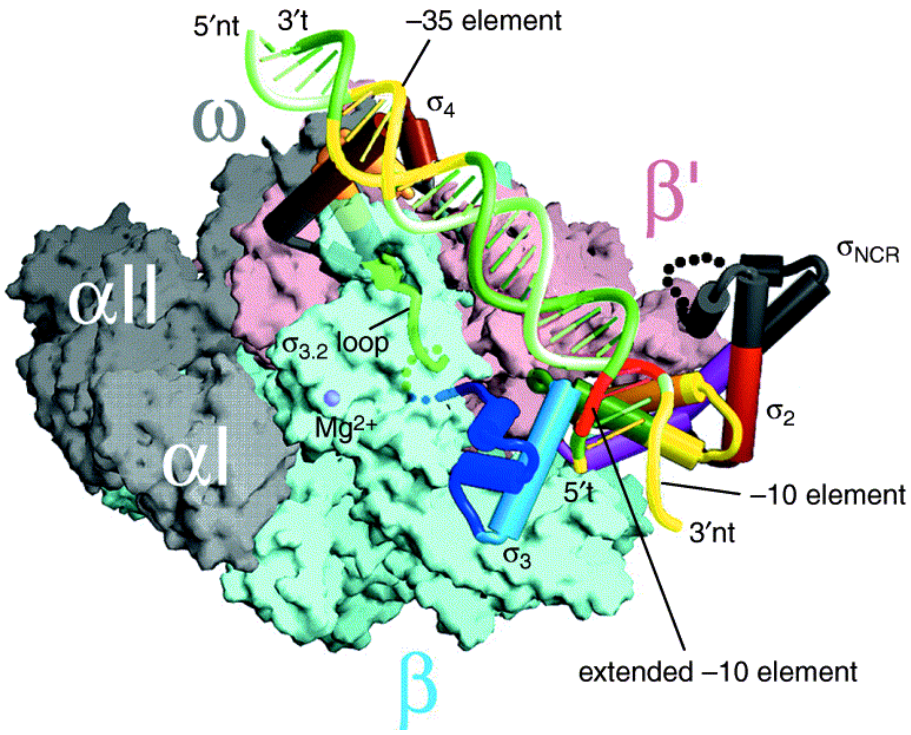


Figure 1.4 - **Structure of *T. aquaticus* RNAP holoenzyme in complex with promoter DNA**. Structure of *T. aquaticus* RNAP holoenzyme in complex with fork junction promoter DNA. The β subunit is green, the β' is pink, the αI, αII and ω subunits are grey and σ70 subunit is shown as a phosphate-backbone worm. The fork junction DNA is green with the -10 and -35 elements in yellow and extended -10 in red. Figure adapted from Murakami and Darst (2003), copyright license number: 3702491184886

Following many years of attempts to solve the structure of *E. coli* RNAP by a number of research groups, the breakthrough came in 1999 with the publication of the structure of *Thermus aquaticus* RNAP core, solved by X-ray crystallography at 3.3 Å resolution (Zhang et al., 1999). This was closely followed by the publication of the remarkably similar crystal structure of RNAP II from the eukaryote *Saccharomyces cerevisiae* (Cramer et al., 2000). Both structures revealed a molecule shaped like a crab claw, with the β and β' subunits forming two "pincers" and a central catalytic cleft. The enzyme active site is located on the back wall of the central cleft where a catalytically essential Mg^{2+} ion is bound by 3

aspartic acid residues. In addition to these, other notable features highly conserved throughout prokaryotes and eukaryotes include a large channel on the surface to accommodate double stranded DNA either side of the transcription bubble, an exit channel through which RNA leaves the enzyme and the secondary channel that allows access to the active site by substrate nucleotides, small molecules and effector proteins.

Following the publication of the structure of bacterial RNAP, it was possible to further understand the role of each subunit in detail. The α subunits are largely responsible for initial assembly of the core enzyme. Two α subunits dimerise and subsequently bring together the β and β' subunits through interactions with the N-terminal domains (αNTDs). The C-terminal domains of the α subunits (αCTDs) are typically involved interaction with upstream DNA elements as well as being a target on RNAP for transcription factors that bind upstream of the -10 and -35 promoter regions (Hochschild and Dove, 1998; Gourse et al., 2000).

The β and β' subunits, are the largest subunits of bacterial RNAP and comprise the catalytic core of the enzyme (Zhang et al., 1999). Crosslinking studies have identified that both β and β' contact the template and non-template strands as well as DNA/RNA hybrid formed in the transcription bubble.

Finally, ω is the smallest RNAP subunit and although its function is not completely understood, structural models have shown that it binds the N and C-terminal ends of the β' subunit. It is thought to have a role in RNAP core assembly although studies have identified the ω-coding *rpoZ* gene as non-essential in a number of strains including *E. coli* and *S. coelicolor* (Ghosh et al., 2001; Gentry and Burgess, 1989; Santos-Beneit et al., 2011). Recent studies have identified the interface between ω and β' as the proposed binding site for ppGpp in *E. coli* (Ross et al., 2013) (see section 1.2.6 The stringent response).

1.2.3 Sigma factors

Promoter-directed gene transcription requires the formation of an RNAP holoenzyme with the sixth subunit, sigma (σ) (Burgess et al., 1969). A sigma factor is a small, dissociable RNA polymerase subunit that performs a number of distinct functions essential to the process of promoter recognition and transcription initiation. Most bacteria possess multiple sigma factors, each specific to a different promoter sequence, which allows targeting of particular sets of genes or regulons. Following holoenzyme formation, DNA binding and transcription initiation, the sigma factor dissociates from the core in a stochastic manner as transcription elongation occurs (Raffaelle et al., 2005). This allows the free sigma molecule to bind a different core polymerase and initiate transcription at another locus.

Every prokaryotic organism possesses one primary or "housekeeping" sigma factor that is highly abundant, constitutively expressed and responsible for transcription of most genes involved in normal growth. In *E. coli* the primary sigma factor is σ^{70} (although in most organisms the homologue is called σ^{A}) (Figure 1.5) and alternative sigma factors consequently compete with this primary sigma factor for the core polymerase in a model known as "sigma factor competition" (Farewell et al., 1998; Nyström, 2004). *E. coli* possesses seven sigma factors but more are not uncommon with as many as 65 found in *S. coelicolor* (Sharma and Chatterji, 2010; Bentley et al., 2002). As alternative sigma factors are typically responsible for responding to environmental stresses it is generally thought that possession of more sigma factors correlates with the ability of an organism to exist in a more diverse or challenging environment (Mittenhuber, 2002).

Sigma factors can be categorised into two broad families based on their structural homology to two sigma factors found in *E. coli*: the aforementioned primary sigma factor, σ^{70}; and σ^{54}, which performs the same role in directing transcription but shares no domain organisation or sequence similarity to σ^{70}, in addition to requiring enhancer proteins and ATP hydrolysis to initiate transcription (Buck et al., 2000). Whilst common in other bacteria, no σ^{54} family sigma factors are present in *S. coelicolor* or indeed any members of Actinobacteria phylum (Bentley et al., 2002; Francke et al., 2011).

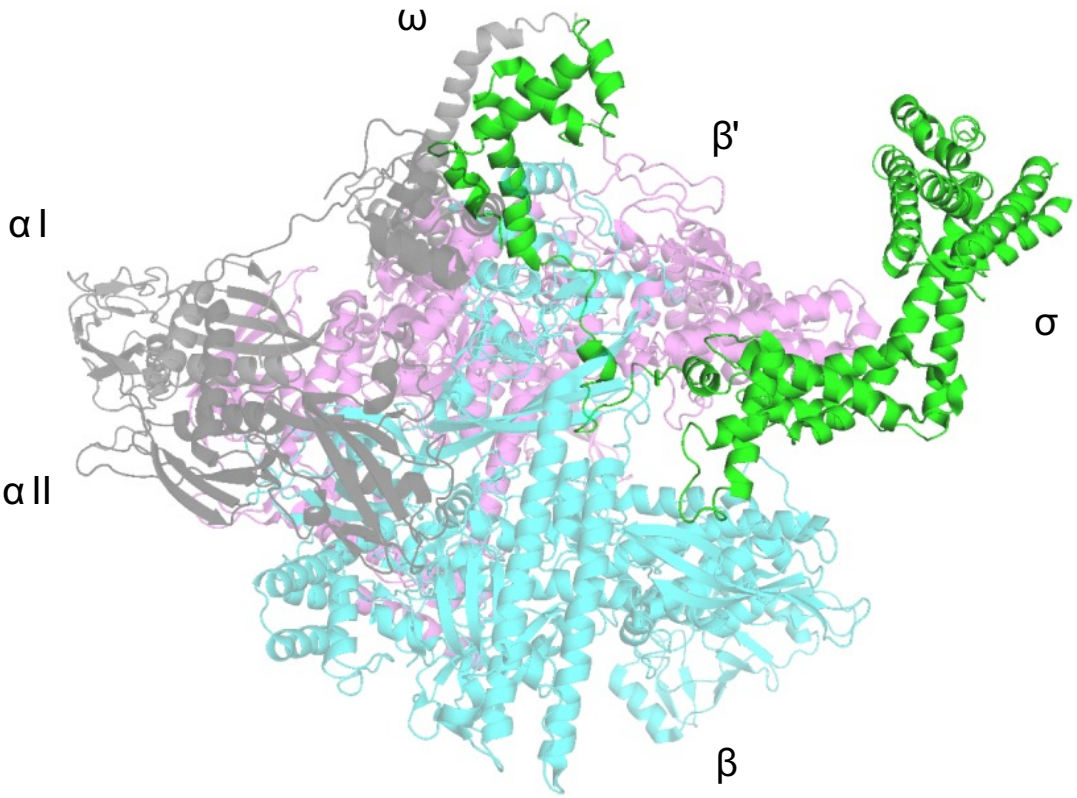


Figure 1.5 - **Structure of *E. coli* RNAP σ^{70} holoenzyme**. Structure of *E. coli* RNAP σ^{70} holoenzyme. The β subunit is cyan, the β' is pink, the αI, αII and ω subunits are grey and σ^{70} subunit is green. Structure downloaded from RCSB PDB under accession number 4YG2 and visualised with PyMOL (version 1.3).

Sigma factors belonging to the σ^{70} family can be classified into 4 groups based on structure and function. Group I contains the primary sigma factors which are all structurally similar to σ^{70} and essential for function. Group II sigma factors are most similar to primary sigma factors but are non-essential for normal bacterial growth. Group III sigma factors are less closely related and usually activate responses to general stress and heat shock or processes such as sporulation or flagella biosynthesis. Group IV sigma factors, often referred to as ECF (Extra-Cytoplasmic Function) sigma factors, are the most functionally and structurally diverse, and regulate responses to extracellular, environmental signals.

The structure and function of primary sigma factors

The group I primary sigma factors all possess four helical structured domains connected by flexible linkers; $\sigma_{1.1}$, σ_2, σ_3, σ_4 comprising of regions 1.1, 1.2 - 2.4, 3.0 - 3.2, 4.1 - 4.2, respectively (Figure 1.6).

$\sigma_{1.1}$, composed of a three-helix bundle connected to region 1.2 by a flexible linker, is responsible for preventing sigma from binding to DNA in the absence of core RNAP. Biochemical studies identified that σ^{70} mutants lacking region 1.1 were more able to bind promoter DNA (Dombroski et al., 1993). Following structural and crosslinking studies it has been suggested that the mechanism behind this auto-inhibition is due to intra-domain interactions between the negatively-charged region 1.1 and the DNA binding elements σ_2 and σ_4 (Schwartz et al., 2008). Once sigma is bound to the RNAP core, region 1.1 still has an important role in transcription. FRET experiments and structural studies have identified that in the absence of DNA, region 1.1 sits in the active-site cleft downstream of the transcription start site, possibly due to possessing a similar negative charge (Mekler et al., 2002; Murakami, 2013). Consequently, for open-complex formation to occur DNA must displace region 1.1. This also provides a suitable explanation for previous results showing that region 1.1 affects the formation but not stability of open complexes (Vuthoori et al., 2001). Interestingly, region 1.1 was able to promote open complex formation at certain promoters whilst inhibit it at other promoters. Further *in vitro* experiments identified that the differential effect of region 1.1 was in fact due to the spacer sequence between the -10 and -35 elements, as opposed to the DNA binding elements themselves (Hook-Barnard and Hinton, 2009).

σ_2, composed of regions 1.2-2.4 is the most highly conserved of the 4 sigma regions and forms the largest interaction with the core RNAP through interaction between the β' subunit and the region 2.2 helix (Young et al., 2001). σ_2 is also responsible for recognition of the -10 promoter element through interaction between regions 2.3-2.4 of the sigma factor and the non-template strand. Interactions occur between σ_2 and every nucleotide in the -10 hexamer although most importantly base-specific interactions occur at A_{-11} and T_{-7}. These two nucleotides, A_{-11} and T_{-7}, are flipped out of the single-stranded base stack and into complementary hydrophobic and hydrophilic pockets of σ_2, respectively, acting as the critical step for melting to occur (see section 1.2.4 Transcription cycle) (Feklistov and Darst, 2011).

In addition to regions 2.3-2.4 responsible for -10 element recognition, region 1.2, comprised of two alpha helices oriented at 90° to one another, was most recently identified as an additional point of σ-DNA contact and transcriptional regulation (Haugen

et al., 2006; Feklistov et al., 2006; Haugen et al., 2008b). Region 1.2 interacts with a GC-rich region immediately downstream of the -10 hexamer known as a discriminator region, originally identified over 30 years ago when investigating stringently controlled promoters (Travers, 1980). Between regions 1.2-2.1 lies a non-conserved region (NCR) that varies in size and structure between organisms (Lonetto et al., 1992). Whilst uncharacterised in most organisms, in *E. coli* the NCR has been identified to contact β' and appears to play a role in promoter escape (Leibman and Hochschild, 2007).

Domain σ_3 (region 3.0-3.2), composed of three helices, interacts with double stranded DNA upstream of the -10 promoter element, in a region known as an extended -10 element. One study performed multiple sequence alignments of 554 promoter regions from *E. coli* (Mitchell et al., 2003). 20% of the analysed promoters possessed at T_{-15} G_{-14} motif, and 44% promoters possessed a G_{-14} base. Further analysis showed that promoters with extended -10 elements often had longer spacer lengths and poorly conserved -35 regions (Mitchell et al., 2003; Campbell et al., 2002). *In vitro* and structural studies also proved that the extended -10 element compensated for these features by stabilising initiation complexes (Barne et al., 1997; Zuo and Steitz, 2015).

Domain σ_4 (region 4.1-4.2) composed of four helices, interacts with the -35 element through a helix-turn-helix motif. In addition to this contact with DNA, σ_4 forms a large contact point with RNAP through the β flap. σ_4 can also serve as a contact point for DNA-bound activators of transcription (Dove et al., 2003).

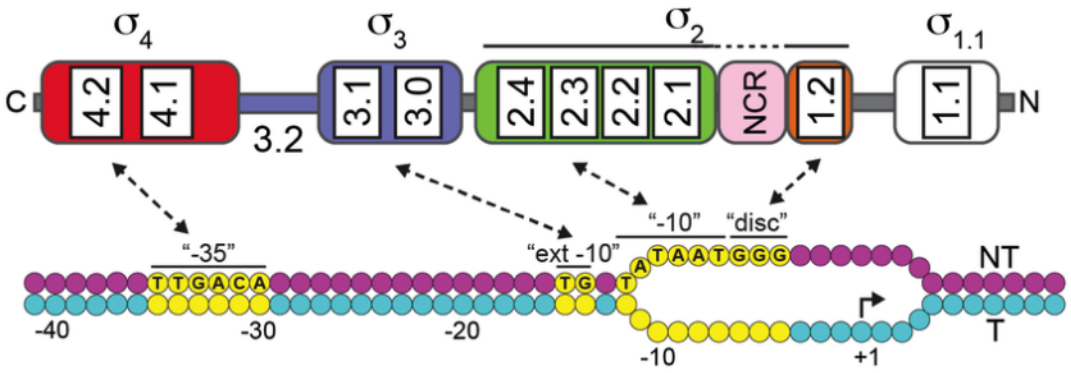


Figure 1.6 - **Domain organization and promoter recognition by the σ^{70} family**. Figure adapted from Paget (2015). This figure is licensed under a Creative Commons Attribution (CC BY) 4.0 Generic License attributed to Mark Paget and can be accessed in Biomolecules 2015, 5(3), 1245-1265.

The structure and function of alternative sigma factors

Group II sigma factors are the most closely related to primary σ^{70} sigma factors however structurally differ by the absence of region 1.1, the region which inhibits DNA binding in the absence of the RNAP core and is only found in group I sigma factors (Lonetto et al., 1992). Additionally, unlike primary sigma factors, they are non-essential for growth (Paget and Helmann, 2003). The best characterised member of the group II sigma factors is *E. coli* σ^S, the master regulator of the general stress response (Weber et al., 2005). In rapidly growing cells σ^S is expressed, although translation is inhibited and any σ^S that is produced is rapidly degraded (Lange and Hengge-Aronis, 1994). However, upon exposure to a range of stresses including nutrient limitation, pH change and DNA damage, these inhibitory mechanisms are overcome, σ^S levels increase rapidly and the σ^S regulon (approximately 500 genes or 10% of the genome) can be expressed. Some of these genes are uniquely expressed by σ^S although a large number are also controlled by the primary sigma factor, σ^{70}. The consensus sequence for promoters regulated by these two sigma factors are highly conserved and *in vitro* transcription experiments have shown that σ^{70} and σ^S holoenzymes recognise the same consensus hexamer sequences (Gaal et al., 2001). Further experiments identified that a $G/T_{-14}C_{-13}$ motif was highly conserved in σ^S-dependent promoters and led to highest expression by σ^S holoenzymes *in vitro* (Becker and Hengge-Aronis, 2001). This interaction between an extended -10 and region 3 could therefore act as a feature for σ^S promoter selection.

Group III sigma factors are structurally and functionally diverse from groups I and II sigma factors, and only possess domains σ_2, σ_3 and σ_4. They usually activate responses to general stress and heat shock or processes such as sporulation or flagella biosynthesis. Examples of well characterised group 3 sigma factors are *B. subtilis* σ^F, σ^E, σ^G and σ^K implicated in regulation of endospore formation (reviewed by Errington, 2003).

Group IV sigma factors are often referred to as ECF (extracytoplasmic function) sigma factors. They are the most functionally and structurally diverse group of sigma factors and regulate responses to extracellular, environmental signals. Group IV sigma factors differ from other σ factors by the absence of $\sigma_{1.1}$, $\sigma_{1.2}$ and σ_3, possessing only domain σ_2 and σ_4. Recent structural studies performed on σ_2 of *E. coli* σ^E, an ECF σ factor that controls the response to cell envelope stress, revealed the mechanism for promoter

melting and open complex formation differs from the studied mechanism in group I sigma factors (Campagne et al., 2014). Whilst promoter melting in group I sigma factors occurs through flipping of nucleotides A_{-11} and T_{-7} into complementary pockets on σ_2 (see section 1.2.4 Transcription cycle), a single nucleotide C_{-10} is flipped into σ_2 to initiate promoter melting at alternative σ promoters.

Anti-sigma factors

Sigma factors represent an efficient and precise method for simultaneously controlling expression of a large number of genes. To exploit this ability, the activity of the sigma factor must also be regulated and a number of mechanisms exist for this purpose. The amount of sigma factor present in the cell can be controlled by regulating transcription and translation or by actively degrading the protein; for example, all of these mechanisms are used to regulate the activity of σ^S in *E. coli* (Battesti et al., 2011).

A more common approach for controlling sigma activity, particularly for ECF sigma factors, is binding by an anti-sigma factor and inhibition of holoenzyme formation. This σ-anti-σ interaction is reversible and typically disrupted by a direct signal that releases the sigma factor and induces expression of the regulon. Typically, ECF sigma factors are positively auto-regulated and co-expressed with their anti-sigma factor; this ensures that stoichiometric levels of the sigma and anti-sigma are present in the cell. One advantage of this mechanism is that cells can respond rapidly to a stimulus by releasing an alternative sigma, which does not require a lag time for transcription and translation to occur.

Two mechanisms for anti-σ binding that effectively sequester σ and prevent holoenzyme formation have been identified. The first are anti-σ factors that bind σ between domains σ_2 and σ_4 through a conserved structural motif found in 40% of anti-σ factors including RseA, the anti-σ that binds σ^E in *E. coli* (Campbell et al., 2003, 2007). In this conserved mechanism, three alpha helices bind between σ_2 and σ_4 whilst a fourth helix binds $\sigma_{2.2}$ preventing interaction with the β' subunit of core RNAP. In addition to σ^E, this mechanism for anti-σ interaction is also used for sequestration of *S. coelicolor* σ^R, *R. sphaeroides* σ^E, and *B. subtilis* σ^W by RsrA, ChrR, and RsiW, respectively (Li et al., 2002; Campbell et al., 2007; Schöbel et al., 2004)

The second group of anti-σ factors are those that wrap around σ_2 and σ_4, maintaining free sigma in a compact form unable to bind neither RNAP nor DNA. This mechanism for anti-σ binding, identified in *Aquifex aeolicus* σ^{28}/FlgM, appears to be conserved in *Cupriavidus metallidurans* σ^{CnrH} and its anti-σ CnrY, and Alphaproteobacteria PhyR and anti-σ-like NepR despite no similarity (Sorenson et al., 2004; Campagne et al., 2012; Maillard et al., 2014).

σ^{70} and Rsd

In *E. coli*, the switch from exponential to stationary phase requires drastic lifestyle changes that are largely governed by the alternative sigma factor σ^S taking over from the primary sigma factor σ^{70}. For a number of years this change was poorly understood, as σ^{70} levels stay constant through growth into stationary phase, at a higher concentration and binding affinity for RNAP core than σ^S (Jishage and Ishihama, 1998). The breakthrough came upon discovery of the σ^{70}-binding anti-sigma factor, Rsd, which is upregulated upon entry into stationary phase (Jishage and Ishihama, 1998, 1999). Similar to the mechanism of other anti-sigma factors, Rsd inhibits RNAP core binding through interactions with regions 2 and 4 of σ^{70} (Patikoglou et al., 2007; Yuan et al., 2008; Paget, 2015). More recently, ligand fishing experiments identified Hpr, a member of the PTS phosphotransferase system which responds to sugar availability, as a binding target Rsd (Park et al., 2013). Hpr was shown to only bind Rsd in its dephosphorylated state, suggesting that model that in the absence of the preferred sugar source Rsd can bind σ^{70} and promote σ^S activity.

Sigma factors in the actinomycetes

Following the sequencing of the *S. coelicolor* genome it was revealed that it produced as many as 65 sigma factors, at the time an unprecedented number of sigma factors (Bentley et al., 2002).

Initially, four *S. coelicolor* genes that encode a sigma factors similar to σ^{70} were cloned using primers based on the sequence of *E. coli* σ^{70} and *B. subtilis* σ^A (Tanaka et al., 1988). The amplified genes were named *hrdA*, *hrdB*, *hrdC* and *hrdD* (homologue of *rpoD* A-D). Mutagenesis studies revealed that *hrdA*, *hrdC* and *hrdD* were all non-essential genes and stable deletion mutants were "unaffected in differentiation, gross morphology, and

antibiotic production" (Buttner et al., 1990; Buttner and Lewis, 1992). In contrast, *hrdB* was identified as an essential gene. Following isolation of RNAP that can transcribe the *dagA* agarase gene *in vitro*, σ^{HrdB} was identified as a co-eluting 66 kDa polypeptide essential for transcription (Brown et al., 1992). Alignment of the -10 and -35 elements for this promoter revealed the similarity to the sequence for other major bacterial sigma factors, concluding that σ^{HrdB} is the principle sigma factor in *S. coelicolor.* The roles of σ^{HrdA}, σ^{HrdC} and σ^{HrdD} still remain largely unknown although it has been shown that whilst deletion mutants had no effect on antibiotic production, σ^{HrdD} activates transcription of antibiotic CSRs *actII-ORF4* and *redD* in vitro. Of the 61 remaining sigma factors, only a handful have been studied in further detail. Ten of these are type III sigma factors including σ^{WhiG} and σ^{F}, required for the process of sporulation (Flärdh and Buttner, 2009).

Of the 45 ECF sigma factors identified following sequencing of the genome, only 4 have been characterised. σ^{E} is an ECF sigma factor involved in maintenance of cell wall structure in response to envelope stress (Paget et al., 1999b, 1999a). The σ^{E}-regulon is induced by transcription of its gene *sigE*, regulated by a two-component signal transduction system. In response to cell wall stress, the sensor kinase CseC and response regulator CseB activate transcription of *sigE* (Paget et al., 1999a). Based on promoter sequence prediction, the σ^{E}-regulon includes genes implicated in maintaining cell wall integrity through glycan biosynthesis. σ^{BldN} is a developmentally-regulated ECF sigma required for formation of aerial mycelium (Bibb et al., 2000). Deletion mutants identified through a genetic screen present a classic bald phenotype. ChIP-chip experiments performed on *Streptomyces venezuelae* σ^{BldN} identified that in addition to the *bld* cascade, σ^{BldN} also regulates the rodlin and chaplin genes which form a hydrophobic sheath around streptomycete aerial mycelium and spores (Bibb et al., 2012). σ^{T} is the most poorly understood of the studied *S. coelicolor* ECF sigma factors. With its cognate anti-σ factor RstA, it is believed to be involved in differentiation and antibiotic production with deletion mutants showing accelerated growth and enhanced production of antibiotics ACT and RED (Mao et al., 2009; Feng et al., 2011).

<u>σ^{R} and RsrA</u>

σ^{R} is an *S. coelicolor* ECF sigma factor that regulates the response to oxidative stress (Paget et al., 1998). It was first identified as a 31 kDa protein present in purified RNAP

preparations (Kang et al., 1997). It was purified from an SDS-PAGE gel, renatured and titrated into core RNAP before use in *in vitro* transcription assays. When performed on a range of templates it directed transcription from *actIIIpx2*, *hrdDp2* and *glnRp2* promoters. Whilst a possible promoter consensus sequence could be extracted from alignments of these promoter sequences, the function of the sigma factor remained unknown. The breakthrough came when it was observed that an *S. coelicolor* Δ*sigR* mutant was sensitive to a number of redox cycling compounds, especially the thiol-specific oxidising agent diamide (Paget et al., 1998). The decreased ability of the mutant to reduce disulphide bonds caused by these agents suggested that the Δ*sigR* was deficient in the thioredoxin system. The *trxBA* operon was shown to be under the control of σ^R and induced by diamide, suggesting a homeostatic model for responding to oxidative stress. As common in many ECF sigma factor systems (Helmann, 2002), the cognate σ^R-specific anti-sigma factor RsrA is located directly downstream and co-transcribed with *sigR* (Kang et al., 1999). RsrA was shown to bind σ^R and inhibit σ^R-directed transcription when in a reduced state. Inactivation of RsrA and the release of σ^R is dependent redox-active cysteine residues present in RsrA which form intramolecular disulphide bonds upon exposure to an oxidising agent (Paget et al., 2001b). As these intramolecular disulphide bonds are a target for the thioredoxin system, this represents a homeostatic model for the σ^R-RsrA response to oxidative stress.

Using a consensus sequence derived from three previously identified σ^R-dependent promoters, Paget et al. (2001a) were able to identify and confirm 27 new genes as σ^R targets. This defined regulon was later expanded to 108 genes following microarray and ChIP-on-chip experiments (Kallifidas et al., 2010; Kim et al., 2012). In addition to the thioredoxin system, other members of the σ^R-regulon include genes responsible for protein quality degradation/control (e.g. *ssrA*, *clpX*, *lon*), ribosome-associated genes (translation initiation factors IF-2 and IF-3, 50S ribosomal proteins) and a number of transcription factors (*hrdB*, *rbpA*). The size and diversity of this regulon emphasises the importance of σ^R in controlling redox balance and maintaining protein quality in response to oxidative stress.

1.2.4 Transcription cycle

The process of transcription can be separated into 3 distinct phases: initiation, where the RNAP holoenzyme recognises the promoter, opens a 13bp transcription "bubble" and begins the process of making RNA; elongation, the major processive state of transcription; and termination, where elongating RNAP dissociates from the template DNA and releases the nascent RNA.

Initiation

Following holoenzyme assembly, initiation is the first step of the transcription cycle defined by stages from promoter recognition through to promoter escape into elongation and synthesis of RNA. Regulation of transcription is a major focus of lifestyle regulation in bacteria and initiation is considered the primary point of transcriptional regulation, effected by binding proteins and small molecules, as well as a target of antibiotics such as rifampicin.

The first step of transcription initiation is promoter binding and formation of a closed complex (RPc). When σ binds RNAP core, a conformational change exposes regions 2.3-2.4 and 4.1-4.2 allowing recognition of the promoter -10 and -35 elements, respectively (Murakami et al., 2002).

Following RPc formation, the transition to open complex (RPo) can be divided into a number of unstable isomerisation intermediates (I) (Figure 1.7). For elucidation of these short-lived conformations, much of the work was performed at low temperatures on the λP_R promoter. The first intermediate, I1,E (Early), is formed when DNA bending occurs upstream of the -35 region identified by OH footprinting experiments (Davis et al., 2007). Bending and nonspecific interaction of this upstream DNA with the αCTD is important for open complex formation, with truncated DNA fragments having 1.5-2 times less efficient isomerisation (Davis et al., 2005).

Formation of I1,L (Late) requires bending of the duplex DNA upstream of the -10 element by at least 90° into the RNAP cleft. Additionally, this step includes flipping of a highly conserved A_{-11} residue from the base stack on the non-template strand into a complementary hydrophobic pocket on σ_2 (Feklistov and Darst, 2011) (Figure 1.6). Specific interactions in this pocket account for the almost universal conservation of an

adenine base at position -11 (A_{-11}) on σ^{70} promoters. Similarly, a T_{-7} residue flips into a hydrophilic pocket on σ_2 (Feklistov and Darst, 2011). Whilst this pocket is more spacious and would accommodate the pyrimidine cytosine, specific interactions are lost with a serious effect on open complex formation. This protein-base interaction is an important requirement for transcription initiation and seemingly explains high conservation of T_{-7} to similar levels as A_{-11}. The order of DNA bending and base-flipping in open complex formation is currently unknown; it has been hypothesised that either bending of DNA nucleates base flipping into complementary pockets, or alternatively, base-flipping occurs allowing DNA to bend into the RNAP cleft (Feklistov and Darst, 2011).

Formation of the intermediate I_2 involves melting and opening of DNA to form a 13bp transcription bubble extending from -11 to +2. Whilst existing data is not conclusive, it is believed that formation of this transcription bubble begins with the flipping of A_{-11} and "unzips" downstream towards the site of transcription initiation. Through this mechanism, the exposed +1 is positioned at the RNAP active site for transcription initiation to occur. Formation of the final isomerisation intermediate, I_3, before RPo involves large conformational changes in RNAP for assembly of the β and β' clamp/jaw apparatus.

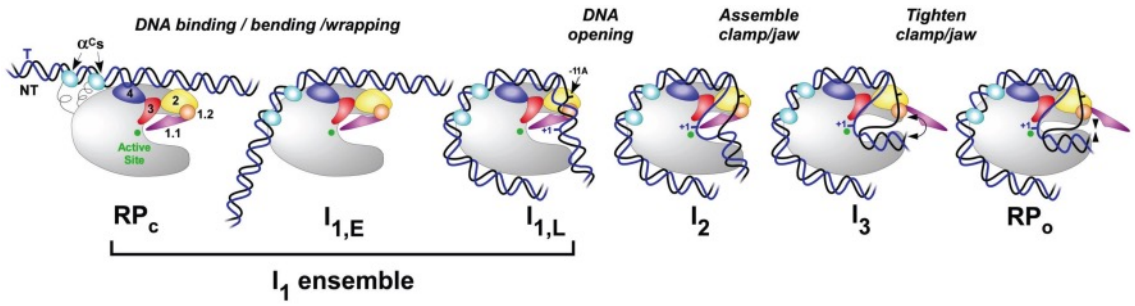


Figure 1.7 - **Schematic representation of proposed intermediates in open complex formation** Figure from Ruff et al. (2015). This figure is licensed under a Creative Commons Attribution (CC BY) 4.0 Generic License attributed to Emily F. Ruff, Thomas Record Jr, and Irina Artsimovitch and can be accessed in Biomolecules 2015, 5(2), 1035-1062.

Abortive transcription is the final process of initiation that occurs following open complex formation where short RNA transcripts, up to 9-10 nt in size are produced before RNAP escapes the promoter into elongation. RNA synthesis begins and downstream DNA is pulled into the active site. However as contacts between RNAP and promoter DNA remain, template DNA is unable to leave the active site and DNA "scrunching" occurs (Kapanidis et al., 2006). When 9-10 nt is formed, the stress caused by scrunching is

relieved either by release of a small abortive transcript, or by disruption of the RNAP-promoter contacts allowing promoter escape into elongation (Zuo and Steitz, 2015).

Elongation

Elongation is the major processive state of RNA transcription that occurs following successful promoter escape. The process of elongation is composed of three defined steps. An NTP molecule binds to the exposed complementary template strand in the active site of elongating RNAP. Upon binding, the RNA chain is extended by a reaction between the 3'-OH of the last NTP with the α-phosphate of the newly added NTP (releasing the β and γ phosphate groups) forming a phosphodiester bond. The final stage is translocation of RNAP to position the next exposed template base at the centre of the active site. Without interruption, this process of elongation occurs at an rate of 50-100 nucleotides per second (Roberts et al., 2008).

Termination

The final stage of transcription is termination, the dissociation of RNAP from the DNA template and release of the nascent RNA. There are two distinct forms of prokaryotic transcription termination: intrinsic termination and Rho-dependent termination.

Intrinsic termination occurs at specific template sequences with the requirement for no additional factors. At intrinsic terminators, the mRNA has two defined features that cause RNAP dissociation; a GC rich sequence that forms a hairpin structure 7-20 bases in length (d'Aubenton Carafa et al., 1990), closely followed by a short run of uracil residues (Gusarov and Nudler, 1999). The RNA hairpin has the function of causing RNAP to stall and the rU-dA interaction formed in the transcription bubble is unusually weak. When these two features are located in close proximity in a terminator sequence, RNAP stalls at the hairpin, the weak RNA-DNA interaction at the U-rich region separates and RNAP is released from the template DNA. Computational analyses initially suggested that intrinsic terminators were found at the termini of 50% of *E. coli* genes (Lesnik et al., 2001). However, following genome scale ChIP analysis of Rho, it was estimated that intrinsic termination (where Rho is absent) is responsible for transcriptional termination at 80% all *E. coli* genes (Peters et al., 2009).

The second mechanism of prokaryotic transcription termination is dependent on the essential protein Rho. Rho-dependent terminators do not share the features of intrinsic terminators but instead rely on the absence of RNA secondary structures allowing Rho translocation. Rho-dependent termination is initiated by binding of a Rho complex, composed of six Rho molecules arranged in an open circle (Skordalakes and Berger, 2003), to a cytosine rich *rut* (Rho utilisation) site on the nascent RNA (Chen and Richardson, 1987). Once bound, Rho translocates along RNA in a 5' to 3' direction through its ATP-dependent helicase activity until it encounters an RNAP transcription elongation complex. Upon meeting the RNAP complex, although the mechanism is not fully understood, it is assumed that Rho causes transcription termination through its helicase activity on the RNA-DNA hybrid (Epshtein et al., 2010).

1.2.5 RNA polymerase binding factors

Secondary channel binding factors

The secondary channel is a narrow pathway from the surface to the RNAP active site, roughly 10-12 Å in size, formed by two regions of the β' subunit (Zhang et al., 1999). Due to the size of the channel, it was predicted that neither double nor single stranded DNA would fit and was consequently proposed as an access point for NTP substrates. Since this suggestion, the secondary channel has also been identified as a binding site for a number of proteins collectively known as second channel-binding factors (SCBFs).

SCBFs are structurally similar, typically possessing a globular domain responsible for binding to RNAP and a coiled-coil structure which protrudes into the secondary channel and regulates transcription through interactions between the tip and the RNAP active site (Zenkin and Yuzenkova, 2015).

GreA and GreB are two SCBFs that modify transcription by rescuing backtracked elongation complexes. Backtracking occurs when RNAP slides backwards along the transcribed DNA template, leaving the 3' end of the nascent RNA away from the active site (Komissarova and Kashlev, 1997). This renders RNAP unable to extend the 3' end of nascent RNA causing an arrest of transcription. In the absence of a Gre factor, the RNAP trigger loop is able to cleave the nascent RNA and position a new 3' end in the RNAP active site (Yuzenkova and Zenkin, 2010). GreA and GreB are able to displace the trigger

loop when positioned in the secondary channel and hydrolyse RNA in a much more efficient manner (Roghanian et al., 2011).

DksA is an SCBF that acts in concert with ppGpp to regulate transcription during the stringent response. Together ppGpp and DksA are able to both positively and negatively regulate the expression of many genes, dependent on intrinsic promoter properties (see section 1.2.6).

In addition to its role in regulation of transcription initiation, DksA has also been identified to have effects on transcription elongation. Indeed, recent ChIP studies performed on DksA have identified it as located on elongation complexes as well as at promoter regions (Zhang et al., 2014). It was initially identified that upon exposure to UV light, DNA damaging agents or nutrient stress, DNA replication was inhibited in a Δ*dksA* deletion mutant (Trautinger et al., 2005; Tehranchi et al., 2010). Further experiments suggested that this inhibition was caused by arrested RNAP elongation complexes blocking replication fork progression. Interestingly, gre factors appeared to functionally overlap this role of DksA, as overexpression of *greA* rescued a Δ*dksA* deletion mutant (Tehranchi et al., 2010), presumably through its understood mechanism of rescuing backtracked/stalled elongation complexes. However, the mechanism through which DksA prevents collision between replication and elongation complexes is still not fully understood. It was initially suggested that DksA displaced stalled elongation complexes or prevented backtracking (Tehranchi et al., 2010) however *in vitro* data suggested DksA does not bind or effect backtracked or active elongation complexes (Furman et al., 2013; Roghanian et al., 2015). More recently, it has been suggested that DksA, potentiated by ppGpp, functions on elongation through increasing transcription fidelity (Roghanian et al., 2015). By reducing nucleotide misincorporation, which consequently causes RNAP stalling/backtracking, DksA is able to prevent formation of stalled RNAP complexes and therefore reduces collision of transcription and replication machinery.

CarD

CarD is an 18 kDa RNA polymerase binding protein first identified in screens for genes involved in the DNA damage response in *M. smegmatis* (Stallings et al., 2009). In these screens *carD* was one of the most upregulated genes in response to a number of DNA

damaging agents including bleomycin, ciprofloxacin, MMS and the double strand break-forming endonuclease I-SceI.

Genetic studies identified CarD as an essential protein. Authors Stallings et al. (2009) found it not possible to create a *carD* deletion strain without the presence of an additional copy of the gene under the control of an ATc-inducible promoter in both *M. smegmatis* and *M. tuberculosis*. Whole-genome microarray transcriptional profiling of these strains depleted of the inducer revealed a large number of genes upregulated in the absence of CarD including translational machinery such as ribosomal proteins and rRNA operons. This is largely reminiscent of the transcriptional profile of relaxed mutants characteristically unable to mount a stringent response by downregulation of rRNA and ribosomal proteins (see section 1.2.6). Additionally, heterologous expression of CarD rescued an *E. coli dksA* deletion mutant unable to downregulate rRNA transcription following nutrient limitation. As CarD accumulates in response to a range of stresses and regulates rRNA it was at this point thought that the protein could play a role in the mycobacterial stringent response.

However, despite these experiments suggesting CarD as a negative regulator of transcription, *in vitro* run off transcription experiments performed on rRNA promoters have since identified CarD as an activator of transcription (Srivastava et al., 2013). Further experiments performed *in vitro* have demonstrated that CarD acts on transcription initiation, specifically through stabilising open promoter complexes (Rammohan et al., 2015; Davis et al., 2015). DNAse I and $KMnO_4$ footprinting experiments demonstrated that CarD does not affect RNAP-promoter interaction or structure of the transcription bubble but rather stabilises and prevents transcription bubble collapse. Comparisons between *E. coli* and *M. bovis* systems revealed that mycobacterial RNAP holoenzymes form very unstable open complexes. Addition of CarD rescues these intrinsically unstable holoenzymes, potentially revealing the importance and essential nature of CarD (Davis et al., 2015).

CarD is composed of two functional domains. The N-terminal RNAP-interacting domain (RID) is highly similar in amino acid sequence and structure to the RNAP-interacting domain of transcription-repair coupling factor (TRCF), the product of the *mfd* gene (Stallings et al., 2009; Srivastava et al., 2013). The N-terminal RID is composed of a Tudor-

like fold consisting of four anti-parallel β strands. TRCF and CarD interact with the β1 lobe of the β subunit of RNAP through similar mechanisms however there is no known functional overlap, with TRCF unable to complement a CarD mutant (Westblade et al., 2010; Stallings et al., 2009). Weakening the CarD-RNAP interaction through site-directed mutagenesis significantly affected *M. tuberculosis* viability, highlighting the importance of CarD action on RNAP *in vivo* (Weiss et al., 2012). Despite these interaction mutants having similar effects on viability as depletion of CarD, the picture was less clear when comparing antibiotic-susceptibility profiles. Whilst depletion mutants are sensitive to ciprofloxacin but not rifampicin, mutants with attenuated RNAP-CarD interaction are particularly sensitive to rifampicin and streptomycin. This finding is of particular interest as it suggests the CarD-RNAP interaction as a novel drug target in rifampicin-resistant strains of *M. tuberculosis* (Weiss et al., 2012).

The CarD-CTD has no sequence or structural similarity to any previously characterised protein or fold (Srivastava et al., 2013). It is a compact domain, composed of five alpha helices. Among these alpha helices is a surface exposed tryptophan residue, almost universally conserved in all CarD homologues, surrounded by a patch of basic residues. Based on the structure of TRCF-RID in complex with the RNAP β1 lobe it was possible to model CarD onto a structure of the *T. thermophilus* RNAP open complex. Due to the structural rigidity of CarD, this model revealed that when bound to RNAP the CarD-CTD is positioned to directly interact with the upstream edge of the -10 element of the promoter DNA (Srivastava et al., 2013). Closer inspection revealed that this interaction potentially occurs through the highly conserved tryptophan residue. When the structure of CarD was modelled onto an RNAP closed-complex structure, the tryptophan residue and surrounding basic patch clashed with the promoter DNA. Site-directed mutagenesis revealed that this tryptophan residue was essential for the function of CarD as a transcriptional activator.

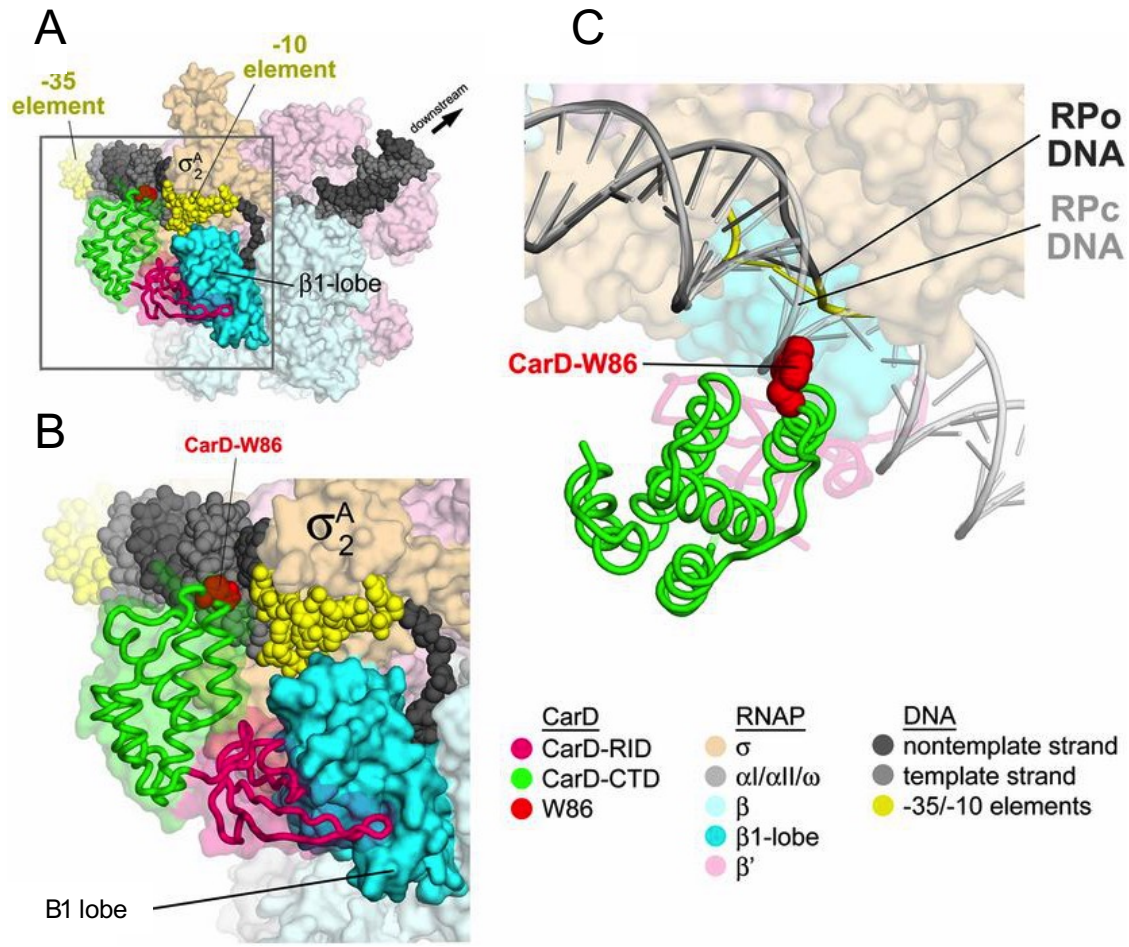


Figure 1.8 - **Structure of CarD modelled onto *T. thermophilus* RNAP open complex.** (A, B) Structure of *T. thermophilus* CarD modelled onto β1-lobe *T. thermophilus* RNAP open complex. (C) Close up view of CarD/RPo model showing potential steric clash and close approach of CarD-W86 with DNA in RPc and RPo models, respectively. Figure adapted with permission from Srivastava et al (2013) can be accessed in Proc Natl Acad Sci USA. 2013, 110(31), 12619-24.

RbpA

RbpA is a 14 kDa RNA polymerase binding protein (RNA polymerase binding protein A) first identified through its consistent presence in purified *S. coelicolor* RNAP preparations (Paget et al., 2001a). It is widespread throughout but exclusive to the Actinobacteria. Studies into the oxidative stress response identified it as a member of the σ^R regulon although at this point its function was still unknown (Paget et al., 2001a).

Deletion mutants demonstrated the importance of RbpA in growth, with *rbpA* mutants exhibiting a slow growth phenotype (Newell et al., 2006). Colonies appear significantly smaller on all studied solid media and have extended lag and slow exponential phases in liquid culture. Despite such significant growth defects, *rbpA* mutants appear morphologically unaffected, with colonies eventually forming aerial hyphae and sporulating comparable to wild-type strains.

Multi-round *in vitro* transcription studies identified RbpA as an activator of transcription from σ^{HrdB}-dependent promoters (Newell et al., 2006; Tabib-Salazar et al., 2013) and single-round reactions suggest that this occurs through the initiation step of transcription (Hu et al., 2012).

RbpA binds RNAP through interactions with the σ subunit. Bacterial two-hybrid (BACTH) experiments have shown that this interaction is specific to the primary σ factors, σ^{HrdB} and σ^A, in *S. coelicolor* and *M. tuberculosis* respectively, in addition to some group II sigma factors (Hu et al., 2012; Tabib-Salazar et al., 2013). Further binding and structural studies have shown that the flexible C-terminal region of RbpA binds to σ_2. The structure of this sigma-interaction domain (SID) in complex with domain 2 of σ^A was solved by Hubin et al. (2015). When modelled upon the *Thermus aquaticus* RNAP open complex structure, it unexpectedly revealed that the RbpA basic linker contacts DNA on the upstream edge of the -10 promoter element. Indeed, crosslinking studies confirmed interaction between the basic linker and promoter DNA and mutagenesis of interacting residues identified this DNA interaction as essential for RbpA function. It had previously been suggested that RbpA could regulate transcription through binding to the β/β' active-site channel near the rifampicin binding site (Dey et al., 2011) or through binding to the β subunit (Hu et al., 2012). Structural data predicting distances of 53 Å and 80 Å from the RbpA binding site on σ suggest these interactions are unlikely (Hubin et al., 2015).

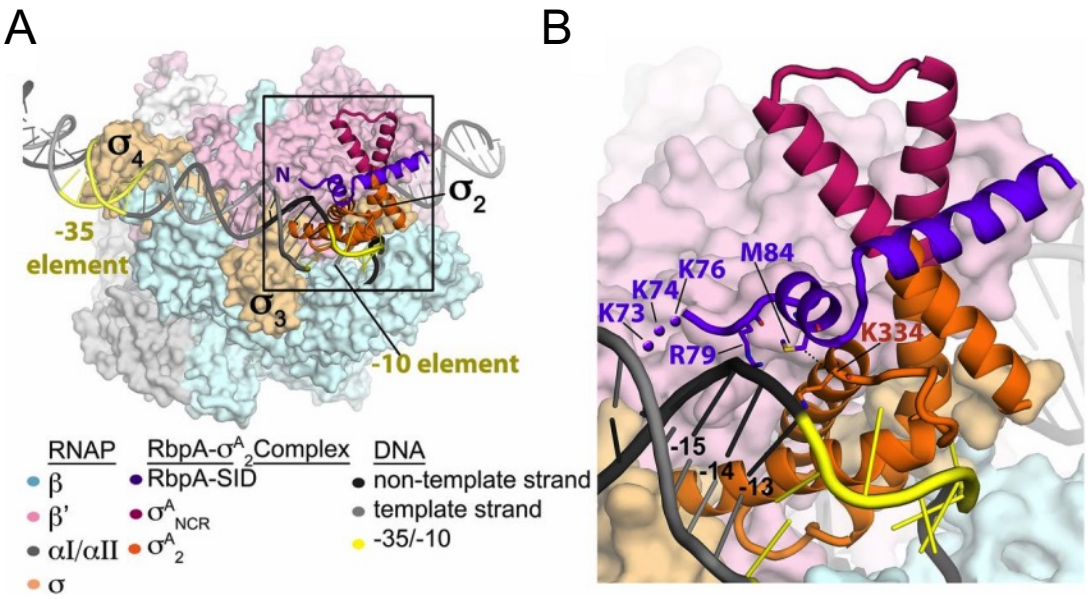


Figure 1.9 - **Structural model of *M. tuberculosis* RbpA on *T. thermophilus* RPo.** Figure adapted with permission from Hubin et al. (2015) can be accessed in Proc Natl Acad Sci USA. 2015, 112(23), 7171-7176.

Crl

Crl is a γ-proteobacterial regulatory protein that binds to the group II sigma factor σ^S, the so-called "master regulator" of the stress response. Crl functions as an activator of transcription from σ^S-dependent promoters, purportedly through facilitating assembly of σ^S with core RNAP (Pratt and Silhavy, 1998; Gaal et al., 2006). In addition to this role it has been suggested that Crl can potentiate promoter binding and open complex formation (Bougdour et al., 2004; England et al., 2008). Bacterial two-hybrid (BACTH) experiments have identified that Crl interacts with domain 2 of σ^S (Monteil et al., 2010). The full mechanism of how Crl promotes holoenzyme assembly is not yet understood however it has been suggested to act as a tether or by promoting holoenzyme stability (Banta et al., 2014).

1.2.6 The stringent response

The stringent response is an adaptive response to amino acid starvation observed in a wide range of bacteria and plant chloroplasts. It is characterised by a number of changes in cellular activity, most notably a decrease in stable RNA (rRNA and tRNA) synthesis (Lazzarini and Dahlberg, 1971), which can otherwise account for up to 80% of bacterial total RNA, and an increase in transcription of genes responsible for amino acid biosynthesis (Stephens et al., 1975; Zhou and Jin, 1998).

RelA and (p)ppGpp synthesis

Early research into the stringent response identified a "magic spot" appearing on thin-layer chromatograms that coincided with nutrient limitation (Cashel and Gallant, 1969). Additionally, "magic spot" production was not observed in "relaxed" strains of *E. coli*, unable to mount a stringent response. This "magic spot" was soon identified as two phosphorylated nucleotides, guanosine tetraphosphate (ppGpp or MS I) and guanosine pentaphosphate (pppGpp or MS II), collectively known as (p)ppGpp but often referred to as ppGpp (Cashel and Kalbacher, 1970).

(p)ppGpp synthesis occurs when a deacetylated tRNA, symptomatic of amino acid starvation, enters the acceptor site (A-site) of a ribosome (Haseltine and Block, 1973). RelA, bound to the 50S ribosome subunit, detects the presence of the uncharged tRNA and, through a change in conformation, begins synthesis of ppGpp and pppGpp from GDP and GTP, respectively. In addition to these phosphorylated guanosine substrates, the reaction also requires ATP as the source of two phosphate molecules. Upon synthesis of ppGpp, RelA is dislodged from the ribosome although single-molecule studies have shown that it remains in an active conformation for a short amount of time and synthesises ppGpp in the cytosol (Wendrich et al., 2002; English et al., 2011). This supports the "hopping" model, which suggests that RelA moves between ribosomes and explains how low concentrations of RelA (1 per 200 ribosomes) (Pedersen and Kjeldgaard, 1977) can suitably monitor presence of uncharged tRNA molecules throughout the entire cell.

In addition to RelA, *E. coli* possesses an additional gene with ppGpp synthase activity called SpoT. Unlike RelA, SpoT produces ppGpp in response to diverse signals other than amino acid starvation including fatty acid and iron limitation (Seyfzadeh et al., 1993; Vinella et al., 2005). Compared to RelA, SpoT possesses weaker ppGpp synthase ability and also has ppGpp hydrolase activity, allowing it to degrade ppGpp to GTP or GDP (Xiao et al., 1991) Since the discovery of SpoT, proteins with ppGpp synthase activity are referred to as RelA-SpoT homologue (RSH) enzymes.

Direct regulation of transcription by ppGpp and DksA

Following the discovery of ppGpp, characterisation of its synthesis via RelA in a range of bacterial species and understanding of its effects on stable RNA synthesis, the mechanism of action through which a small molecule exhibited such drastic physiological changes remained unknown for a number of decades. To exhibit such a broad range of changes, regulation of transcription through RNA polymerase was considered the most obvious target (Wagner, 2002). This idea was supported by multiple studies on Relaxed mutants resistant to ppGpp possessing mutations in the *rpoB* gene (Nene and Glass, 1983; Little et al., 1983; Tedin and Bremer, 1992).

In 2001, authors Barker et al. identified ppGpp as a direct transcriptional regulator through a series of *in vitro* and *in vivo* experiments (Barker et al., 2001b, 2001a). *In vivo* studies identified ppGpp-null strains having lower activity from amino acid biosynthetic promoters and higher activity from rRNA promoters. I*n vitro* transcription experiments identified ppGpp inhibits transcription from rRNA promoters at the initiation step by decreasing the half-life of RNAP open complexes (Barker et al., 2001b). Since promoters have a wide variety of basal stability, ppGpp will destabilise and downregulate transcription from some promoters more than others. rRNA promoters form unusually unstable open complexes and are therefore particularly sensitive to ppGpp-mediated destabilisation (Barker et al., 2001b). ppGpp also has the ability to activate transcription of many genes, including those responsible for amino acid biosynthesis. This was initially thought to occur as an indirect effect of negative regulation; not only are these RNAP-promoter open complexes for these genes more stable but the reduction in rRNA synthesis was thought to increase the availability of free RNAP to initiate new complexes (Magnusson et al., 2007). In addition to this model for indirect activation, another mechanism of direct activation was identified following the characterisation of DksA.

DksA is a secondary channel-binding factor that acts synergistically with ppGpp to potentiate the stringent response. It was initially identified as a suppressor of temperature sensitive growth of DnaK deletion mutants when expressed on a multicopy plasmid (DnaK suppressor A) (Kang and Craig, 1990). *dksA* deletion mutants are unable to regulate rRNA in response to changes in amino acid concentration, growth rate or growth phase (Paul et al., 2004). *In vitro* transcription experiments identified that DksA

greatly enhances the effects of ppGpp on inhibition rRNA transcription (Paul et al., 2004). Before discovery of DksA, activation of transcription by ppGpp *in vitro* had not been demonstrated. Subsequently, it was shown that DksA potentiates direct activation of amino acid promoters by ppGpp *in vitro* (Paul et al., 2005).

Following x-ray crystallography studies performed on *T. thermophilus* RNAP, it was initially proposed that ppGpp binds RNAP at the active site (Artsimovitch et al., 2004). Contrary to this, further mutagenesis of proposed binding residues in *E. coli* did not weaken the effect of ppGpp *in vitro* and it was suggested that initial results could have been influenced by an insufficient concentration of ω in RNAP preparations (Vrentas et al., 2005, 2008). Interestingly, it has since been shown that (p)ppGpp does not bind or affect *T. thermophilus* RNAP *in vitro* but acts through other protein targets (Liu et al., 2015).

The requirement of ω for RNAP to respond to ppGpp *in vitro* (Vrentas et al., 2005) was soon better understood. Two independent crosslinking and structural studies on *E. coli* RNAP identified a binding site for ppGpp, at the interface between ω and β' subunits (Ross et al., 2013; Zuo et al., 2013). Site-directed mutagenesis revealed that several predicted binding residues are essential for a response to ppGpp. Considering the ppGpp binding region of RNAP is not directly involved in the enzymatic activity of RNAP, the mechanism of action of ppGpp is considered allosteric. The mechanism is not yet understood although it has been suggested that ppGpp affects the stability of open complexes through binding to the RNAP shelf and core domains involved in open complex formation (Ross et al., 2013; Zuo et al., 2013). Additionally, it also also unknown how this binding site and mechanism relates to the activity of DksA in the secondary channel.

Regulation of GTP biosynthesis by ppGpp

Direct regulation of RNAP by ppGpp has been well studied in *E. coli* however is not conserved in all bacteria and alternative targets for controlling the stringent response have been identified. In both *E. coli* and *B. subtilis*, rRNA promoters have been shown to have characteristically short half-lives dependent on high concentrations of the initiating NTP (iNTP) (Gaal et al., 1997; Krásný and Gourse, 2004). Whilst *E. coli* regulates rRNA transcription through ppGpp binding RNAP, the stringent response in *B. subtilis* works through regulation of concentrations of GTP, the iNTP at all rRNA promoters. ppGpp does

not affect *B. subtilis* RNAP and the binding site is not conserved (Krásný and Gourse, 2004; Ross et al., 2013) but instead inhibits the activity of a number of enzymes in the GTP biosynthesis pathway. Through metabolic profiling of WT and ppGpp^0 mutants upon exposure to amino acid limitation, authors Kriel et al. (2012) were able to identify ppGpp targets. Two enzymes, Gmk, which converts GMP to GDP, and HprT, which converts hypoxanthine to IMP and guanine to GMP, were identified as putative ppGpp targets. Microarray data showed that transcript levels weren't affected but *in vitro* enzyme assays showed ppGpp potently inhibited activity of both purified Gmk and HprT (Liu et al., 2015).

ppGpp and the stringent response in *Streptomyces* spp.

Following its discovery in *E. coli*, ppGpp was identified as present in a number of *Streptomyces* species (Oki et al., 1975; Hamagishi et al., 1981; Ochi, 1987a, 1987b). Following isolation of a number of relaxed mutants, a connection between the production of ppGpp, antibiotic production and morphological differentiation was soon made (Ochi, 1986, 1987b, 1990). Relaxed mutants of *Streptomyces* sp. MA406-A-1, *S. griseus* and *S. coelicolor* that lacked the ability to accumulate ppGpp were unable to produce the antibiotics formycin, streptomycin, and actinorhodin and prodigiosin, respectively. Strauch et al. (1991) were able to elicit the stringent response in *S. coelicolor* by addition of serine hydroxamate or performing a nutrient downshift on exponentially growing cultures. In response to these conditions, an increase in ppGpp was observed and total RNA synthesis decreased. S1 nuclease mapping experiments performed on the *rrnD* gene set showed a decrease in transcription from all four promoters. Interestingly, in these experiments, especially following nutrient downshift, it was observed that an increase in ppGpp coincided with a drastic decrease in cellular GTP concentrations. Deletion of the *relA* gene in *S. coelicolor* created a strain completely unable to produce ppGpp upon entry into stationary phase (Chakraburtty and Bibb, 1997; Kang et al., 1998). Following amino acid depletion, the *relA* deletion mutant showed no decrease in rRNA synthesis and failed to produce both ACT and RED.

In a novel method for studying the stringent response, Hesketh et al. (2001) placed an N-terminal truncation of *relA* under the control of a thiostrepton-inducible promoter. Expression of this fragment allowed production of ppGpp in a thiostrepton-dependent manner, independent of nutrient sufficiency. Microarray experiments performed with

this strain revealed a more global picture of the stringent response in *S. coelicolor* but did not provide an explanation for how ppGpp induces antibiotic production (Hesketh et al., 2007). More recently, it has also been shown that ppGpp inhibits the enzyme polynucleotide phosphorylase (PNPase) in *Streptomyces* spp. (Gatewood and Jones, 2010). PNPase is an enzyme with both exonuclease and RNA 3'-polyribonucleotide polymerase activity, responsible for degradation of RNA and synthesis of heteropolymeric RNA tails. Experiments showed that ppGpp effectively inhibits both activities *in vitro*, as well increasing the half-life of bulk mRNA in *S. coelicolor.* From this it was suggested that an increase in stability of bulk mRNA may contribute to antibiotic biosynthesis as part of the stringent response however this remains experimentally unproven.

Chapter 2:

Materials and methods

2 Materials and methods

2.1 Materials

2.1.1 Chemicals and reagents

- Agarose (Melford)
- Amaranth (Sigma-Aldrich)
- Ammonium Persulfate (APS) (Sigma-Aldrich)
- Ampicillin (Melford Laboratories)
- Antifoam (Sigma-Aldrich)
- Apramycin (Duchefa Biochemie)
- Bromophenol Blue (Amersham Biosciences)
- Casamino Acids (Difco)
- Chloramphenicol (Melford Laboratories)
- Chloroform (Sigma-Aldrich)
- Deoxyribonucleotide phosphates (dNTPs) (New England Biolabs)
- Dimethylsulphoxide (DMSO) (Sigma-Aldrich)
- Dithiothreitol (DTT) (Melford Laboratories)
- Glycerol (Fisher Scientific)
- Isopropyl-β-D-thiogalactopyranoside (IPTG) (Melford Laboratories)
- Kanamycin (Melford Laboratories)
- Malt extract (Oxoid)
- Mannitol (Sigma-Aldrich)
- Nalidixic acid (Duchefa Biochemie)
- N, N-dimethyl-formamide (Fisher Scientific)
- Nutrient agar (Difco)
- Orange G dye (Fisher Scientific)
- Phenol (Fisher Scientific)
- Phenylmethylsulfonyl Fluoride (PMSF) (Sigma-Aldrich)
- 1,4-Piperazinediethanesulfonic acid (PIPES) (Sigma-Aldrich)
- Protease Inhibitor Cocktail tablets (Sigma-Aldrich)

- Protein A magnetic beads (New England Biolabs)
- Protein G magnetic beads (New England Biolabs)
- SequaGel - UreaGel System (National Diagnostics)
- Sigmacote® (Sigma-Aldrich)
- Sodium dodecyl sulphate (SDS) (Fisher Scientific)
- Sodium trichloroacetate (NaTCA) (Sigma-Aldrich)
- Soya flour (Infinity Foods)
- Sucrose (Sigma-Aldrich)
- TES (N-Tris(hydroxymethyl)methyl-2-aminoethane sulfonic acid) (Fisher Scientific)
- Tetramethyl-ethylenediamine (TEMED) (Fisher Scientific)
- Thrombin from bovine plasma (Sigma-Aldrich)
- Sodium tri-isopropylnapthalene sulfate (TPNS) (Acros Organics)
- Trichloroacetic acid (TCA) (Sigma)
- Tris (2-Amino-2-hydroxymethyl-propane-1,3-diol) (Fisher Scientific)
- Tryptone (Difco)
- X-gal (Melford Laboratories)
- Xylene cyanol (Sigma-Aldrich)
- Yeast extract (Difco/Oxoid)

2.1.2 Enzymes

DNA polymerases

- DNA Polymerase I, Large (Klenow) Fragment (New England Biolabs)
- Phusion High-Fidelity DNA Polymerase (New England Biolabs)
- GoTaq DNA Polymerase (qPCR Master Mix) (Promega)

DNA modifying enzymes

- Antarctic Phosphatase (New England Biolabs)
- T4 DNA Kinase (New England Biolabs)
- T4 DNA Ligase (New England Biolabs)

DNA/RNA restriction enzymes

- DNA endonucleases (New England Biolabs)
- S1 Nuclease (Life Technologies)
- RNase A from bovine pancreas (Sigma-Aldrich)
- RQ1 RNase-free DNase (Promega)
- RiboShredder™ RNase Blend (Cambio)

2.1.3 Antibodies

- anti-FLAG M2 monoclonal antibody, from mouse, Sigma F18041MG.
- anti-σ^{HrdB} polyclonal antibody, from rabbit, a gift from P. Doughty.
- anti-σ^{R} polyclonal antibody, from rabbit, a gift from M. Feeney and M. Buttner.
- anti-RNAP β subunit monoclonal antibody, from mouse, Abcam ab12087.

2.1.4 Buffers and solutions

DNA Manipulation

TE buffer (pH 8.0): 10 mM Tris-HCl (pH 8.0), 1 mM EDTA

10X TBE electrophoresis buffer (1 L): 108 g Tris base, 55 g boric acid, 9.3 g Na_2EDTA, dissolved in 1 L dH_2O

5X DNA loading buffer: 40% sucrose, 0.25% orange G dye, dissolved in dH_2O

Alkaline lysis solution I: 50 mM Tris-HCl (pH 8.0), 10 mM EDTA

Alkaline lysis solution II: 0.2 N NaOH, 1% (v/v) SDS,

Alkaline lysis solution III: 3 M potassium acetate (pH 5.5)

2X Kirby mix (100 mL): 2 g TPNS, 12 g sodium 4-aminosalicilate, 5 mL 2M Tris-HCl (pH 8.0), 6 mL equilibrated phenol (pH 8.0), up to 100 mL with dH_2O

Phenol:Chloroform:Isoamyl Alcohol (P:C:IAA) (50 mL): 25 mL equilibrated phenol (pH 8.0), 24 mL chloroform, 1 mL isoamyl alcohol

STE buffer: 10.3% sucrose, 25 mM Tris-HCl pH 8.0, 25 mM EDTA, 2 mg/ml lysozyme from hen egg white

TFBI buffer: 100 mM $RbCl_2$, 50 mM MnCl 2 .4H 2 O, 30 mM potassium acetate, 10 mM $CaCl_2.2H_2O$, 15% (v/v) glycerol

TFBII buffer: 10 mM MOPS, 10 mM $RbCl_2$, 10 mM $CaCl_2.2H_2O$, 15% (v/v) glycerol

S1 Nuclease mapping

S1 hybridisation buffer: 2.53 g PIPES dissolved in 90 mL dH_2O, 1.67 mL 0.5 M EDTA, adjusted to pH 7.0, autoclaved. 93.1 g NaTCA added and volume up to 167.4 mL with dH_2O. Stored as 100 µL aliquots at -70 °C.

5X S1 digestion buffer (100 mL): 1.4 M NaCl, 150 mM sodium acetate (pH 4.4), 22.5 mM zinc acetate, 100 µg/mL partially-cleaved, non-homologous DNA, heated at 97 °C for 5 min, stored at -70 °C.

S1 termination solution: 2.5 M ammonium acetate, 50 mM EDTA.

Formamide loading buffer: 80% (w/v) deionised formamide, 1 mM EDTA, 0.01% (w/v) xylene cyanol, 0.01% (w/v) bromophenol blue. For loading S1 nuclease reactions add 10 mM NaOH.

Protein purification

Ni-NTA charge buffer: 50 mM $NiCl_2.6H_2O$

Ni-NTA binding buffer: 50 mM Tris-HCl pH 7.9, 0.5 M NaCl, 5 mM imidazole

Ni-NTA wash buffer: 50 mM Tris-HCl pH 7.9, 0.5 M NaCl, 20 mM imidazole

Ni-NTA elution buffer: 50 mM Tris-HCl pH 7.9, 0.5 M NaCl, 1 M imidazole

Ni-NTA strip buffer: 50 mM Tris-HCl pH 7.9, 0.5 M NaCl, 100 mM EDTA

Gel Filtration (GF) buffer: 50 mM Tris-HCl pH 7.9, 50 mM NaCl, 5% (w/v) glycerol, 5 mM β-mercaptoethanol

Anion-exchange binding buffer: 50 mM Tris-HCl pH 7.9, 50 mM NaCl, 5% (v/v) glycerol, 5 mM β-mercaptoethanol

Anion-exchange elution buffer: 50 mM Tris-HCl pH 7.9, 1 M NaCl, 5% (v/v) glycerol, 5 mM β-mercaptoethanol

SDS-PAGE Coomassie Brilliant Blue stain: 0.25% (w/v) Coomassie Brilliant Blue (Sigma), 50% (v/v) methanol, 10% (v/v) glacial acetic acid

SDS-PAGE de-stain solution: 50% (v/v) methanol, 10% (v/v) glacial acetic acid

In vitro transcription

RNAP dilution buffer: 10 mM Tris-HCl, pH 8.0, 10 mM KCl, 5% (v/v) glycerol, 0.4 mg/ml BSA, 0.1% Triton X-100 and 10 mM β-mercaptoethanol

2X *in vitro* transcription buffer: 80 mM Tris-HCl pH 7.9, 20 mM $MgCl_2$, 1.2 mM EDTA, 40% glycerol. Autoclave and add 0.1 vol 8 mM KH_2PO_4 pH 7.5. Prior to use add 1.5 mM DTT and 0.25 mg/mL BSA.

15X NTP mix: 1.5 mM ATP, 1.5 mM CTP, 1.5 mM GTP, 0.75 mM UTP, 5 mCi [α-^{32}P] UTP

Formamide loading buffer: 80% (w/v) deionized formamide, 1 x TBE buffer, 10 mM EDTA, 0.08% (w/v) xylene cyanol, 0.08% (w/v) amaranth.

2.1.5 Strains

Streptomyces strains used and created in this study

Name	Genotype	Reference
M145	SCP1$^-$ SCP2$^-$	(Bentley et al., 2002)
J1915	M145 Δ*glkA*	(Kelemen et al., 1995)
S101	J1915 Δ*rbpA*::*aac*(3)IV	(Newell et al., 2006)
S129	M145 Δ*rpoC*::*rpoC*^{HIS} ΔrbpA::aac(3)IV)	(Newell et al., 2006)
M570	M600 Δ*relA*::*hyg*	(Chakraburtty and Bibb, 1997)
	M145 Δ*dksA*	This study, chapter 3
	M145 ΔSCO6164 ΔSCO6165	This study, chapter 3
	M145 Δ*dksA* ΔSCO6164 ΔSCO6165	This study, chapter 3
S200	J1915 Δ*carD::hyg tipAp-carD*	This study, chapter 4
S201	J1915 Δ*carD::hyg pRT802::carD-DAS+4*	This study, chapter 4
S202	J1915 Δ*carD::hyg pRT802::carD-3XFLAG*	This study, chapter 5

Table 2.1 – A table listing the *Streptomyces* strains used in this study

E. coli strains used in this study

Name	Genotype	Reference
DH5α	F– Φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17 (rK–, mK+) *pho*A *sup*E44 λ– *thi*-1 *gyr*A96 *rel*A1	(Angelis, 1986)
BL21 (λDE3/pLysS)	F- ompT gal dcm lon hsdSB(rB- mB-) λ(DE3) pLysS(cmR)	(Miroux and Walker, 1996)
ET12567 (pUZ8002)	dam-13::Tn9, dcm-6, hsdM, Cmr, pUZ8002 is a derivative of RK2 with a mutation in the oriT (aph)	(MacNeil et al., 1992; Paget et al., 1999b)
ET12567 (pR9406)	dam-13::Tn9, dcm-6, hsdM, Cmr, pR is a derivative of RK2 with a mutation in the oriT (amp)	(Jones et al., 2013)

Table 2.2 – A table listing the *E. coli* strains used in this study.

2.1.6 Plasmids

Name	Features	Source/Reference
pBluescript II SK+	*E. coli* sub-cloning vector: *bla, lacZα,* Amp^R	Stratagene,(Alting-Mees and Short, 1989)
pIJ6902	*S. coelicolor* integrative expression plasmid: *aac(3)IV*, *tipAp, oriT, tsr,* ΦC31 *attP,* Apr^R, $Thio^R$	(Huang et al., 2005)
pIJ6650	*S. coelicolor* counter-selectable suicide plasmid: *aac(3)IV, oriT, glkA, lacZα, ori* (pUC18), Apr^R	(Paget et al., 2001b)
pKC1132	*S. coelicolor* suicide plasmid: *aac(3)IV, oriT, lacZα,*Apr^R	(Bierman et al., 1992)
pIJ963	*S. coelicolor* plasmid, possesses *hyg* gene.	(Zalacain et al., 1986)
pSX162	*S. coelicolor* integrative expression plasmid: *aac(3)IV*, *oriT, ermEp*, lacZα,* ΦC31 *attP,* Apr^R.	(Newell et al., 2006)
pRT802	*S. coelicolor* integrative plasmid: *aphII*, *oriT, lacZα,* ΦBT1 *attP,* Kan^R.	(Gregory et al., 2003)
pET15b	*E. coli* T7 based His-tag expression vector, Amp^R	Novagen

Table 2.3 –A table listing the plasmids used in this study.

2.1.7 Oligonucleotides

All oligonucleotides are 5’ to 3’ with restriction sites in bold.

Chapter 3		
Name	Sequence (5' to 3')	
DksA_HindIII_F	CC**AAGCTT**GTGACGTAGTCGGCGTTCAC	*dksA* upstream flank
DksA_BamHI_R	CC**GGATCC**CGTCTTCTTCTTCGCCACCATG	
DksA_BamHI_F	CC**GGATCC**AAGCAGGAGCGCCGGTACTGAG	*dksA* downstream flank
DksA_EcoRI_R	CC**GAATTC**AGCAGGGAGGCGGCGCGGAAG	
61645_up_for	CCC**AAGCTT**CGGCTCCCGGTCGAGGATGA	SCO6164/5 upstream flank
61645_up_rev	CCC**TCTAGA**GCGGGAGGCGTCGAGCGACAC	
61645_down_for	CCC**TCTAGA**GTGCGGGAGGCCGGACGATGA	SCO6164/5 downstream flank
61645_down_rev	CCC**GAATTC**GAGAAGCAGGTGCGGTGCTT	
61645_col_for	CGCGGCATACCCATATGGCC	ΔSCO6164-5 colony PCR
61645_col_rev	AAGGTCACCGTCGGCACGTC	
DksA_O_F	G**GGATCC**CTCCTGCTTCTGCTTGCACTC	*dksA* overexpression
DksA_O_R	G**AAGCTT**GACTGAGCACGAGGGTAC	
F_D164N	GAGGCGGACACAGGCAGCAAGAACATCA	*dksA* ^{D164N} mutagenesis
R_D164N	GTTGTCGCCCGCGCCGTCGCCGGA	
F_D167N	ACAGGCAGCAAGAACATCACGCGCGA	*dksA* ^{D167N} mutagenesis
R_D167N	GTTCGCCTCGTCGTCGCCCGCGCCGT	
F_D1647N	GCGAACACAGGCAGCAAGAACATCACGCGCG A	*dksA* $^{D164/7N}$ mutagenesis
R_D1647N	CTCGTTGTCGCCCGCGCCGTCGCCGGA	

Table 2.4 - A table listing the DNA oligonucleotides used in chapter 3 of this study.

Chapter 4		
Name	Sequence (5' to 3')	
carD_eco_F	C**GAATTC**GTCCACCGGACACGGCGAAG	carD upstream flank
carD_bam_R	C**GGATCC**TCCTTTGATCTGGCGAGTTTC	
carD_bam_F	C**GGATCC**CTCGCCTCCTGATCCGACTG	*carD* downstream flank
carD_xba_R	G**TCTAGA**CCACCTGCACCGCGACGTTC	
carD_D_F	G**CATATG**ACGTTCAAGGTTGGCGAC	*carD* expression from *tipAp*
carD_O_R	G**AAGCTT**AGTCGGATCAGGAGGCGAGCAC	
ext_carD_F	CC**AGATCT**ATACAGGTCGGGATCGGTAC	*carD* full complementation
ext_carD_R	CC**AGATCT**AATCACGGCGGCGGTCCTGG	
carD_3XFLAG_F	CCCC**AAGCTT**TCGCACACGCGGAACCGGAC	*carD*-DAS+4/3XFLAG
carD_3XFLAG_R	CCCC**AAGCTT**GGAGGCGAGCACCTCGTCGA	
atpIp_F	GCCGCAATACCAGACAAGTTGC	*atpIp in vitro* transcription
atpIp_R	GCCGCGGGCACGGCAGCCTG	
rrnDp1-4_F	GTGACGTCGAGGCAGCCGAAC	*rrnDp1-4 in vitro* transcription
rrnDp1-4_R	GTATCAACATATCTGGCGTTG	
rrnDp2_nt	**AGCTT**TTCCCGCAAGAGCCGTTGACACGGAGC GAGCGGGGAGGTAGATTCGAACAGTG	*rrnDp2* promoter
rrnDp2_t	**GATCC**ACTGTTCGAATCTACCTCCCCGCTCGCT CCGTGTCAACGGCTCTTGCGGGAAA	
rrnDp3_nt	**AGCTT**CGGAAACGAAGGCCGGTAAGACCGGCT CGAAAGTTCTGATAAAGTCGGAGCCGCG	*rrnDp3* promoter
rrnDp3_t	**GATCC**GCGGCTCCGACTTTATCAGAACTTTCGA GCCGGTCTTACCGGCCTTCGTTTCCGA	
rrnDp2w_nt	**AGCTT**TTCCCGCAAGAGCCGTAAGACCGGAGC GAGCGGGGAGGTAGATTCGAACAGTG	*rrnDp2w* promoter
rrnDp2w_t	**GATCC**ACTGTTCGAATCTACCTCCCCGCTCGCT CCGGTCTTACGGCTCTTGCGGGAAA	
rrnDp3s_nt	**AGCTT**CGGAAACGAAGGCCGGTTGACACGGCT CGAAAGTTCTGATAAAGTCGGAGCCGCG	*rrnDp3s* promoter
rrnDp3s_t	**GATCC**GCGGCTCCGACTTTATCAGAACTTTCGA GCCGTGTCAACCGGCCTTCGTTTCCGA	
rrnD_S1_F	GTGACGTCGAGGCAGCCGAAC	*rrnDp1-4* S1 mapping probe
rrnD_S1_R	GTATCAACATATCTGGCGTTG	

Table 2.5 - A table listing the DNA oligonucleotides used in chapter 4 of this study

Chapter 5		
Name	Sequence (5' to 3')	
RbpA_3XFLAG_F	CC**ACTAGT**CC**AAGCTT**CGGCTGCGGCTGCGAC	*rbpA*-3XFLAG
RbpA_3XFLAG_R	CC**GGATCC**CTACTTATCGTCATCGTCCT	

Table 2.6 - A table listing the DNA oligonucleotides used in chapter 5 of this study

Chapter 6		
Name	Sequence (5' to 3')	
PSCO6551_F	GCGCTTTACACAGCGTGAG	*SCO6551p* promoter qPCR
PSCO6551_R	GATCCTTCGGTTGTGGTCAG	
PSCO2763_F	CGTACGCGCCTGATCACC	*SCO2763p* promoter qPCR
PSCO2763_R	CCTGCTCACCTGATCTACCG	
PTRXC_EXT	CAGTGCGAGGTGAACTCGTCC	*trxCp in vitro* transcription
PTRXC_INT	CACCACTCCGCCCAGAAGTCG	
SigA_F	GG**CATATG**GCAGCGACCAAAGCAAG	σ^{A} expression
SigA_R	CC**AGATCT**GCGCTCTCAGTCCAGGTAGTCGCG	
SigA-VRA_F	CGCGCCACGCAGCTGATGACCGAGCTTAGCGAG	$\sigma^{A\ (VR)}$ mutagenesis
SigA-VRA_R	GACGCCAGCCTCGATCCGCTTGGCTAGCTCGAC	
SigA-double_F	GCTGGCGTCCGCGCCACGCAGCTGATGACCGAGCTTAGC	$\sigma^{A\ (RTVR)}$ mutagenesis
SigA-double_R	CTCGATCCGGGTGGCTAGGCGGACCTCTTCCTCGGCGTTG	

Table 2.7 – A table listing the DNA oligonucleotides used in chapter 6 of this study

2.2 Growth, selection and storage of bacterial strains

2.2.1 Media

E. coli

Lennox broth (LB): A liquid media used for growth of *E. coli* strains. 10 g Difco Bacto tryptone, 5 g Difco yeast extract, 5 g NaCl, 1 g glucose dissolved in 1 L dH_2O. Typically divided into 100 mL aliquots in screw cap bottles and autoclaved.

Lennox agar (LA): A solid media used for growth of *E. coli* strains. 10 g Difco Bacto tryptone, 5 g Difco yeast extract, 5 g NaCl, 1 g glucose dissolved in 1 L dH_2O. Typically divided into 100 mL aliquots in 250 mL Erlenmeyer flasks with 1.5 g agar, stoppered with a foam bung and foil, and autoclaved.

2 x YT broth: A richer liquid media used for growth of *E. coli* strains. Also used when heat shocking Streptomyces spores in conjugation protocol.

Streptomyces

Mannitol soya flour (MS) agar: A solid media used for general growth of Streptomyces strains, especially when good sporulation is required. 20 g mannitol dissolved in 1 L tap water. Typically divided into 100 mL aliquots in 250 mL Erlenmeyer flasks with 2 g agar and 2 g soya flour, stoppered with a foam bung and foil, and autoclaved twice with shaking in between.

2 x Pre-germination (PG) medium: A liquid media used for pre-germination of *Streptomyces* spores before liquid cultures. 1 g Difco yeast extract, 1 g Difco casamino acids, dissolved in 100 mL dH_2O. Typically divided into 10 mL aliquots in screw cap universal bottles and autoclaved.

Difco Nutrient (DN) agar: A rich, solid media used for rapid growth of Streptomyces strains for testing antibiotic resistance/sensitivity. 2.3 g Difco nutrient agar dissolved in 100 mL dH_2O and autoclaved.

YEME (10% sucrose): A rich, liquid media used for growth of Streptomyces strains. 3 g Difco yeast extract, 5 g Difco bacto-peptone, 3 g Oxoid malt extract, 10 g glucose (or 10 g glycerol for J1915 based strains), 100 g sucrose, dissolved in 1 L dH_2O and autoclaved.

Prior to use add 2 mL/litre 2.5 M $MgCl_2.6H_2O$ (5 mM final) and 1 mL/litre 10% (v/v) antifoam.

Supplemented minimal medium, solid (SMMS): A minimal, solid medium used for production of *S. coelicolor* antibiotics. 2 g Difco casamino acids, 5.73 g TES buffer, dissolved in 1 L dH2O and adjusted to pH 7.2 with 5N NaOH. Typically divided into 200 mL aliquots in 250 mL Erlenmeyer flasks with 3 g agar, stoppered with a foam bung and foil, and autoclaved. Prior to use add 2 mL 50 mM NaH_2PO_4 + K_2HPO_4 (1 mM final), 1 mL 1M $MgSO_4$ (5mM final), 3.6 mL 50% (w/v) glucose (50 mM final) and 0.2 mL SMM trace element solution (see below).

SMM trace element solution: 0.1 g/L each of $ZnSO_4.7H_2O$, FeSO4.7H2O, $MnCl_2.4H_2O$, $CaCl_2.6H_2O$, and NaCl, made fresh over 2-4 weeks, stored at 4 °C.

Minimal media (MM): A minimal, solid medium used for growth of *Streptomyces* strains and screening for 2-deoxyglucose sensitivity. 0.5 g L-asparagine, 0.5 g K2HPO4, 0.2 g $MgSO_4.7H_2O$, 0.01 g $FeSO_4.7H_2O$, 5 g mannitol, dissolved in 1 L dH_2O and adjusted to pH 7.0 - 7.2. Typically divided into 200 mL aliquots in 250 mL Erlenmeyer flasks with 2 g agar, stoppered with a foam bung and foil, and autoclaved.

2.2.2 Antibiotic selection

Name	Stock solution	Working conc. in liquid media	Working conc. on solid medium
Ampicillin (amp)	100 mg/mL, dissolved in dH_2O, filter sterilised	100 µg/mL	100 µg/mL
Apramycin (apr)	50 mg/mL, dissolved in dH_2O, filter sterilised	25 µg/mL	25 µg/mL
Chloramphenicol (cml)	34 mg/ml, dissolved in 100% EtOH	34 µg/mL	25 µg/mL
Hygromycin (hyg)	50 mg/mL, dissolved in dH_2O, filter sterilised	10 µg/mL	10 µg/mL
Kanamycin (kan)	50 mg/mL, dissolved in dH_2O, filter sterilised	20 µg/mL	20 µg/mL
Nalidixic acid (nali)	25 mg/mL, dissolved in 0.15 M NaOH	25 µg/mL	25 µg/mL
Thiostrepton (thio)	50 mg/mL, dissolved in DMSO	15 µg/mL	15 µg/mL

Table 2.8 - A table listing the antibiotics used in this study and stock/working concentrations in liquid and solid media.

2.2.3 Growth and storage

E. coli

General growth conditions: *E. coli* were grown on solid or liquid media at 37 °C for up to 24 h. For liquid culture, *E. coli* were grown in LB (with selection) in an orbital shaker at 250 rpm. For solid culture, *E. coli* were grown on LA (with selection) and once colonies formed plates were stored at 4 °C for up for 2 weeks.

Making glycerol stocks: For long term storage of *E. coli* strains, 0.5 mL 40% glycerol was added to 0.5 mL overnight culture and stored at -80 °C.

Streptomyces

General growth conditions: Streptomyces were grown on solid media at 30 °C for 5-7 days until colony formation and sporulation.

Growing *S. coelicolor* liquid cultures: Spore stock concentrations were measured and the required volume to give a starting OD_{450} of 0.05 was added 1 mL TES buffer. The mixture was briefly centrifuged, the supernatant was removed and the spore pellet was resuspended in 1 mL of TES buffer. The spore suspension was heat shocked at 50 °C for 10 min, mixed with 1 mL 2 X PG media and 2 µL 5 M $CaCl_2$, and incubated in a universal tube at 37 °C with vigorous shaking for 3-5 h. Following pre-germination of spores, the culture was briefly centrifuged and resuspended in 1 mL YEME media. The suspension was used to inoculate the *S. coelicolor* liquid culture which was grown at 30 °C in an orbital shaker at 200 rpm. *S. coelicolor* were grown in Erlenmeyer flasks with a stainless steel spring loop in the base of the flask to prevent formation of mycelium clumps and improve aeration.

Making spore stocks: A single colony was streaked to confluency on MS agar and incubated at 30 °C for 5 days until the mycelium turned dark grey, indicating production of mature spores. To harvest spores, 9 mL sterile water was added to the plate and scraped with a metal loop suspending spores in the water. The crude suspension was filtered through a "spore tube" (2 cm of non-absorbent cotton wool in a 10 mL syringe, autoclaved to sterilise) into a 15 mL tube and centrifuged at 1,000 x g for 10 min. The supernatant was discarded and the pellet was resuspended in 1-2 mL sterile 20% glycerol.

Spore stocks were stored at -20 °C and remain viable for years, even if repeatedly thawed and re-frozen.

Making mycelial preps: For strains defective in sporulation, a single colony was streaked to confluency on MS agar and incubated at 30 °C for 5 days. 10 mL sterile water was added to the plate and vigorously scraped with a metal loop suspending the mycelium in the water. 5 mL of the suspension was used to inoculate 50 mL YEME (with appropriate selection) and grown in a sprung flask for 1-2 days at 30 °C in a 250 rpm shaking incubator. Once sufficiently grown, the culture was transferred to a 50 mL centrifuge tube and centrifuged at 2,000 x g for 10 min. The pellet was washed with 25 mL 10.3% sucrose solution, centrifuged at 2,000 x g for 10 min and resuspended in 5 mL 20% glycerol before dividing into 200 µL aliquots and frozen at -70 °C.

Replica plating: An MS agar plate containing ~300 sporulating colonies was pressed onto a sterile velvet firmly enough to transfer spores only. The velvet was then "replica plated" onto a maximum of four DN agar plates with selection for different antibiotic resistance markers. Typically, the last plate was non-selective to confirm absence of growth on any previous plates was due to antibiotic selection and not a problem transferring cells. Plates were incubated overnight at 30 °C to show antibiotic sensitive and resistant colonies on the original plate.

2.3 DNA Manipulation

2.3.1 DNA digest

DNA digests were typically performed with ~1 µg plasmid DNA, dissolved in 1X appropriate restriction buffer and ~10U restriction enzyme. Digests were typically performed at 37 °C at least 3 h.

2.3.2 Polymerase Chain Reaction (PCR)

Polymerase chain reactions were performed in 0.2 mL thin-walled reaction tubes using the following components:

Component	Volume	Final Concentration
5X Phusion HF or GC Buffer	10 µL	1X
10 mM dNTPs	1 µL	200 µM
10 µM Forward Primer	2.5 µL	0.5 µM
10 µM Reverse Primer	2.5 µL	0.5 µM
Template DNA	variable	< 250 ng
DMSO (optional)	(1.5 µL)	3%
Phusion DNA Polymerase	0.5 µL	1.0 units/50 µL PCR
Nuclease-free water	to 50 µL	

Reaction conditions were varied according to primer melting temperature and required PCR product size as follows:

Step	Temperature	Time
Initial Denaturation	98 °C	30 s
25-35 Cycles	98 °C 45–72 °C 72 °C	5 - 10 s 10 - 30 s 15 - 30 s per kb
Final Extension	72 °C	5 - 10 min
Hold	4–10 °C	

2.3.3 *S. coelicolor* colony PCR

A colony was picked from an MS agar plate, homogenised in 50 µL dH_2O in a 1.5 mL microcentrifuge tube, and heated in a boiling water bath for 5 min. The cell debris was

removed by centrifugation at 16,000 x g for 2 min and 2 µL of the supernatant was used as template DNA in standard PCR reaction conditions.

2.3.4 Inverse PCR for site-directed mutagenesis

Prior to PCR, 200 pmol of each primer was phosphorylated in 1X T4 kinase buffer with 2 mM ATP and 0.005U T4 DNA kinase (New England Biolabs). The reaction mix was incubated at 37 °C for 30 min before heat inactivation at 90 °C for 5 min. Phosphorylated primers were used for PCR following standard conditions. Following PCR, the product was self-ligated following standard conditions, treated with ~10U DpnI enzyme to remove plasmid template and transformed into competent *E. coli*.

2.3.5 Gel electrophoresis

Agarose gel electrophoresis was used for analysis of DNA size and quality. 0.8% and 1% agarose gels were used for undigested plasmids and DNA fragments respectively, run at 2 V/cm. For visualisation, gels were post-stained with GelRed™ Nucleic Acid Gel Stain (Biotium) according to manufacturer's instructions. The stained gel was visualised using a UV transilluminator.

2.3.6 Gel purification

To purify DNA fragments following gel electrophoresis, agarose gel fragments were excised with a scalpel and purified with the Wizard® SV Gel and PCR Clean-Up System (Promega) according to manufacturer's instructions.

2.3.7 DNA dephosphorylation

DNA was dephosphorylated in 1X Antarctic Phosphatase buffer with 5 units Antarctic Phosphatase (New England Biolabs) and incubated at 37 °C for 15 min (60 min for 3' extensions) before heat inactivation at 70 °C for 5 min.

2.3.8 DNA ligation

DNA ligations were typically performed in 10 µL 1X T4 ligase buffer with an insert to vector ratio of 3:1 and 0.2U T4 DNA ligase (New England Biolabs). Ligations were typically performed at 16 °C for at least 3 h.

2.3.9 Oligonucleotide annealing for construction of dsDNA fragments

To anneal single stranded oligonucleotides for construction of double stranded DNA fragments, 10 µL of each oligo (100 µM) were added to 25 µL nuclease-free water and 5 µL NEBuffer 3 restriction digest buffer. The mixture was heated at 95 °C for 5 min before allowing to cool to room temperature. The mixture was diluted 100-fold with nuclease-free water and 1 µL was added to a standard ligation mix before transformation of *E. coli*.

2.3.10 Determining DNA/RNA concentration

For rough estimations, DNA and RNA concentrations were determined using a NanoDrop 1000 (Thermo Fisher Scientific). 1 µL DNA was loaded onto the pedestal and measured using the appropriate sample settings. For more accurate calculations typically prior to *in vitro* assays, DNA concentrations were determined using a Qubit High Sensitivity DNA Assay Kit and Qubit 2.0 Fluorometer according to manufacturer's instructions.

2.4 Nucleic acid extraction and purification

2.4.1 Small scale plasmid isolation from *E. coli* (Wizard Miniprep)

To purify plasmid DNA from *E. coli,* the Wizard® Plus SV Miniprep DNA Purification System was used according to manufacturer's instructions.

2.4.2 Small scale plasmid isolation from *E. coli* (Alkaline Lysis method)

To purify plasmid DNA from *E. coli*, cells were harvested from 3 mL overnight culture at 16,000 x g for 5 min and the cell pellet was resuspended in 200 µL alkaline lysis solution I. 400 µL alkaline lysis solution II was added, mixed by inverting 5 times before addition of 300 µL alkaline lysis solution III + 1 µL 10 mg/mL RNase A and incubation for 10 min at room temperature. The mixture was centrifuged at 16,000 x g for 10 min and the

supernatant was transferred to a new tube. 150 µL P:C:IAA was added, vortexed for 2 min, centrifuged at 16,000 x g for 10 min and the upper phase was transferred to a new tube. 600 µL isopropanol was added, chilled on ice for 10 min and centrifuged at 16,000 x g for 10 min. The supernatant was removed and the pellet was washed with 200 µL 70% ethanol before centrifugation at 16,000 x g for 1 minute. All supernatant was removed and the pellet was resuspended in 50 µL dH_2O.

2.4.3 Large scale plasmid isolation from *E. coli* (Qiagen Midiprep)

To purify plasmid DNA from *E. coli* on a large scale, the QIAGEN Plasmid Midi Kit (Qiagen) was used according to manufacturer's instructions.

2.4.4 Chromosomal DNA isolation from *S. coelicolor*

To isolate chromosomal DNA from *S. coelicolor*, cells were harvested from 1.5 mL of overnight culture at 16,000 x g for 5 min and the cell pellet was washed with 1 mL 10.3 % sucrose. The pellet was resuspended in 250 µL STE buffer and incubated at 37 °C for 30 min. 330 µL 2X Kirby mix was added and vortexed for 30 s. 670 µL of P:C:IAA was added and vortexed for 30 s. The mixture was centrifuged at 16,000 x g for 5 min and the upper phase was transferred to a new tube. 250 µL of P:C:IAA was added and vortexed for 30 s. The mixture was centrifuged at 16,000 x g for 5 min and the upper phase was transferred to a new tube. 0.1 volume 3M sodium acetate (pH 6) and 1 volume isopropanol was added, mixed and centrifuged at 16,000 x g for 5 min. The pellet was resuspended in 200 µL TE buffer containing 10 µg/mL RNase A and incubated at 37 °C for 30 min. As before, 0.1 volume 3M sodium acetate (pH 6) and 1 volume isopropanol was added, mixed and centrifuged at 16,000 x g for 5 min. The pellet was washed with 70% ethanol and air-dried for 15 min. Finally, the DNA pellet was resuspended in 100 µL TE buffer.

2.4.5 RNA isolation from *S. coelicolor*

To isolate RNA from *S. coelicolor*, cells were harvested from 10 mL of culture at 6,000 x g for 1 minute, the supernatant removed and the cell pellet resuspended in 1 mL 2X Kirby mix. The cells were disrupted by sonication (3 x 3s @ 35% ampl.), 500 µL P:C:IAA was added, mixed and stored at -20 °C until ready to process.

The mixture was thawed and centrifuged at 16,000 x g for 5 min. The upper phase was transferred to a new tube, 600 µL of P:C:IAA was added, vortexed for 2 min and centrifuged at 16,000 x g for 5 min. The upper phase was transferred to a new tube, 0.1 volume 3M sodium acetate (pH 6) and 1 volume isopropanol was added, mixed and incubated at -20 °C overnight. The mixture was centrifuged at 16,000 x g for 10 min, the supernatant discarded and the pellet was washed with 70% ethanol. The pellet was resuspended in 200 µL 1X DNase buffer + 0.5 µL DNase and incubated at 37 °C for 30 min. 200 µL of RNase-free water and 200 µL P:C:IAA was added and mixed. The mixture was centrifuged at 16,000 x g for 10 min and the upper phase was transferred to a new tube. 0.1 volume 3M sodium acetate (pH 6) and 1 volume isopropanol was added, mixed and incubated at -20 °C for 1 h. The mixture was centrifuged at 16,000 x g for 5 min, the pellet washed with 100 µL 70% ethanol and the RNA pellet was resuspended in 50 µL RNase-free H_2O.

2.5 Introduction of DNA into *E. coli*

2.5.1 Preparation of chemically competent *E. coli* ($CaCl_2$ method)

An overnight culture was used to inoculate 50 mL LB with selection (1:100 dilution) and grown at 37 °C to an OD600 of 0.4 - 0.6. Cells were harvested by centrifugation at 2,000 x g for 10 min, resuspended in 20 mL ice-cold 0.1 M $CaCl_2$, and incubated on wet ice for 30 min. Following incubation, cells were pelleted by centrifugation at 2,000 x g for 10 min, gently resuspended in 2 mL 0.1 M $CaCl_2$ + 15% glycerol, and snap frozen in 100 µL aliquots in a dry ice-ethanol bath. Competent cell aliquots were stored at -80 °C for up to 6 months.

2.5.2 Preparation of chemically competent *E. coli* ($RbCl_2$ method)

An overnight culture was used to inoculate 50 mL LB with selection (1:100 dilution) and grown at 37 °C to an OD600 of 0.4 - 0.6. Cells were harvested by centrifugation at 2,000 x g for 10 min, resuspended in 6 mL TFBI buffer, and incubated on wet ice for 30 min. Following incubation, cells were pelleted by centrifugation at 2,000 x g for 10 min, gently resuspended 1 mL TFBII buffer and snap frozen in 100 µL aliquots in a dry ice-ethanol bath. Competent cell aliquots were stored at -80 °C for up to 6 months.

2.5.3 Transformation of chemically competent *E. coli*

E. coli transformations were typically performed with 1 - 10 ng of DNA (plasmid mini-prep or ligation mix) with 100 µL of competent cells. Cells were thawed on wet ice, DNA was mixed gently and incubated on ice for 30 min. The mixture was heat shocked in a water bath at 42 °C for 45 s, returned to ice for 2 min, 900 µL LB added and incubated at 37 °C. The transformation mix was plated onto LA with the appropriate selection and incubated overnight at 37 °C.

2.6 Introduction of DNA into *S. coelicolor*

2.6.1 Conjugation from *E. coli*

To introduce DNA into *S. coelicolor* via conjugation, the plasmid was first introduced into the donor strain *E. coli* ET12567 (pUZ8002) (cml^R kan^R) via transformation. For conjugation of kan^R plasmids, the donor strain *E. coli* ET12567 (pR9406) (cml^R amp^R) was used. A single colony was used to inoculate a 5 mL LB (with cml, kan and plasmid selection) overnight culture. The following day the overnight culture was diluted 1:100 to inoculate 50 mL LB (with kan and plasmid selection) and grown at 37 °C to an OD600 of 0.4 - 0.6. The cells were washed twice with 50 mL LB and resuspended in 0.1 volume of LB. While washing the *E. coli,* approximately 10^8 *Streptomyces* spores were added to 0.5 mL 2 x YT broth, heat shocked in a water bath at 50 °C for 10 min before allowing to cool to room temperature. To conjugate, mix 0.5 mL *E. coli* with 0.5 mL heat shocked *Streptomyces* spores. Mix, centrifuge briefly for 15 s and pour off most of the supernatant (leaving ~100 µL). The pellet was resuspended in the residual liquid, plated out onto MS agar + 10 mM $MgCl_2$ and incubated at 30 °C for 16 - 20 h. To select for transconjugants, overlay the plate with 1 mL dH_2O containing 0.5 mg nalidixic acid and the appropriate antibiotic for selection of the transferred plasmid. Plates were incubated at 30 °C for 5 - 7 days and transconjugants were picked and streaked to single colonies on MS agar plates containing 25 ug/mL nalidixic acid and the appropriate antibiotic for selection of the transferred plasmid.

2.7 Analysis of nucleic acids

2.7.1 S1 nuclease protection assay

To create the probe, the reverse primer (downstream from the 5' end) was first labelled with ^{32}P as follows:

Component	Volume
Reverse primer (10 pmol/µL)	3 µL
10X T4 Polynucleotide kinase buffer	4 µL
ATP [γ-^{32}P]	5 µL
Polynucleotide kinase (10 U/µL)	1 µL
dH_2O	Up to 40 µL

Incubate for 30 min at 37 °C then add 4 µL 3M sodium acetate (pH 6) and 80 µL 100% ethanol. Vortex to mix and leave at -80 °C overnight. The following day the mixture was centrifuged at 16,000 x g for 30 min at 4 °C and the radioactive supernatant was removed. The pellet was washed with 100 µL 75% ethanol, the supernatant removed and the pellet air-dried for 20 min. The remaining PCR reaction was mixed, added to the labelled reverse primer and amplified using standard conditions.

Following PCR, the probe was purified using a Qiagen PCR purification kit, eluting with 50 µL RNase-free water. To confirm successful amplification, 2 µL of the PCR product was run on an agarose gel. A reading of >200 cpm/µL generally indicated a good probe.

To perform S1 mapping, 30-40 µg RNA was mixed with 5-10 ng of probe and lyophilised with a Speedivac. The dried pellet was resuspended in 20 µL 1X S1 hybridisation buffer, incubated in a waterbath at 70 °C for 10 min before allowing the temperature to fall to 42 °C overnight. 300 µL S1 digestion mix was forcibly added to the mixture, thoroughly mixed and placed in ice water until all samples were ready. The digest mix was incubated in a waterbath at 37 °C for 45 min until the reaction was stopped by addition of 75 µL S1 stop buffer. 1 µL glycogen and 400 µL isopropanol was added, mixed and incubated at -20°C for 1 h. The mixture was centrifuged at 16,000 x g at 4°C for 15 min and all supernatant was removed. The pellet was washed with 70% ethanol and air-dried for 10 min. The pellet was resuspended in 6 µL S1 formamide loading buffer, heated at 95 °C

for 2 min and 3 µL was run on an 8% UREA-PAGE sequencing gel at 600V for 1 h 20 min. The gel was dried onto Whatman paper and exposed overnight in a phosphorimaging cassette before visualisation the next day with a Typhoon phosphoimager.

2.7.2 qPCR

qPCR was performed in a 96-well PCR plate using a StepOnePlus Real-Time PCR system (Applied Biosystems). Each reaction contained 12.5 µL GoTaq® qPCR Master Mix, 2 µL sample DNA, up and downstream primers (0.1–1 µM each) and dH_2O up to a final volume of 25 µL. Data was collected and analysed using Applied Biosystems® Real-Time PCR Software.

2.8 *In vitro* transcription

To perform *in vitro* transcription, 50 nM RNAP, 100 nM σ and appropriate concentrations of additional proteins were combined in 1X *in vitro* transcription buffer and incubated on ice for 10 min. 5 nM DNA template was added and incubated at 30 °C for 10 min. The reaction was started by addition of radioactive NTP mix. After incubation at 30 °C for 10 min, 1 µL 5 mg/mL heparin was added and incubated at 30 °C for 5 min. 1 µL 10 mM UTP was added and incubated at 30 °C for 10 min. The reaction was stopped by addition of formamide loading buffer. Samples were run on an 8% UREA-PAGE sequencing gel at 600V for 1 h 20 minutes. The gel was dried onto Whatman paper and exposed overnight in a phosphorimaging cassette before visualisation the next day with a Typhoon phosphoimager.

2.9 Protein purification

2.9.1 Ni-NTA sepharose affinity chromatography

For preparation of a Ni-NTA sepharose column, iminodiacetic acid (IDA) sepharose fast-flow resin was added to a syringe stopped with glass wool to give a column volume of 1-4 mL. Resin was washed with 3 column volumes (CV) dH_2O followed by 5 CV Ni-NTA charge buffer and 5 CV Ni-NTA binding buffer. Cleared cell lysate was loaded onto the column before washing with 10 CV Ni-NTA binding buffer, 5 CV Ni-NTA wash buffer and

1 CV Ni-NTA elution buffer. The column was washed with 5 CV Ni-NTA strip buffer, 5 CV dH_2O and stored in 20% ethanol.

2.9.2 Anion-exchange chromatography

Anion-exchange chromatography was performed using a Mono Q 5/50 GL column (GE Healthcare) and ÄKTA FPLC (Amersham PLC). The complete system was washed with 5 CV dH_2O followed by equilibration with 5 CV anion-exchange binding buffer. Protein was loaded onto the column and eluted using a linear gradient of anion-exchange binding buffer. 1 mL fractions were collected, analysed by SDS-PAGE and pooled as appropriate. Following elution, the column was washed with 5 CV dH_2O and 5 CV 20% ethanol before storage.

2.9.3 Size exclusion chromatography

Size-exclusion chromatography was performed using a HiLoad TM 16/60 Superdex 200 pg column (GE Healthcare) and ÄKTA FPLC (Amersham PLC). The complete system was washed with 1 CV dH_2O followed by equilibration with 1 CV gel filtration buffer. Protein was loaded onto the column and 1 mL fractions were collected, analysed by SDS-PAGE and pooled as appropriate. Following elution, the column was washed with 1 CV dH_2O and 1 CV 20% ethanol before storage.

2.9.4 Protein sample analysis by SDS-PAGE

For analysis of purified protein samples by SDS-PAGE, precast NuPAGE® 4-12% Bis-Tris polyacrylamide gels (Invitrogen) were used. Samples were dissolved in 1X NuPAGE® LDS sample buffer and run according to manufacturer's instructions.

2.9.5 Determining protein concentration

Samples were analysed for protein concentration using a Qubit Protein Assay Kit and Qubit 2.0 Fluorometer according to manufacturer's instructions.

Chapter 3:

Results I: Investigating the role of DksA in *S. coelicolor*

3 Investigating the role of DksA in *S. coelicolor*

3.1 Overview

ppGpp and the stringent response has long been implicated in the production of antibiotics. This chapter sets out to identify how ppGpp and DksA regulate transcription in *S. coelicolor*. This chapter shows that the ppGpp binding site on RNAP recently identified in *E. coli* is not conserved in *S. coelicolor*. This suggests that ppGpp regulates transcription through a different binding site or by binding to alternative targets. Although the mechanism by which ppGpp regulates transcription in *S. coelicolor* is different to that in *E. coli*, *S. coelicolor* possesses three DksA-like proteins. Whilst deletion of these proteins appears to have no effect on the growth of *S. coelicolor* under any tested conditions, overexpression of DksA stimulates antibiotic production dependent on conserved aspartic acid residues at the predicted coiled-coil tip.

3.2 The proposed *E. coli ppGpp* binding site is not conserved in *S. coelicolor*

Structural studies on *E. coli* RNAP identified the ppGpp-binding site at the interface between ω and β' subunits (Ross et al., 2013; Zuo et al., 2013). Following elucidation of the ppGpp binding site, Ross et al. (2013) performed multiple sequence alignments of the ppGpp-binding regions on β' and ω subunits with the corresponding subunits from *B. subtilis* and *T. thermophilus.* In these two distantly related bacterial species, RNAP did not crosslink ppGpp and was unaffected by ppGpp *in vitro* (Krásný and Gourse, 2004; Vrentas et al., 2008). Sequence alignments revealed that residues in β' and ω, identified as essential for ppGpp-binding, are not conserved in these species. Either this suggested the existence of an alternative ppGpp-binding site on RNAP, or that ppGpp does not affect RNAP but has alternative targets in these species. Indeed, later studies have shown that ppGpp has the ability to regulate enzymes in the GTP biosynthesis pathway including guanylate kinase (GMK) that converts GMP to GDP, which mediates the stringent response (Liu et al., 2015).

To investigate whether ppGpp binds to RNAP in *S. coelicolor*, multiple sequence alignments were performed to investigate the conservation of the proposed ppGpp binding site (Figure 3.1). Amino acid sequences for β' and ω from *E. coli*, *R. sphaeroides*,

B. pertussis, *M. xanthus*, *B. subtilis* and *S. coelicolor* were obtained from NCBI and clustalW alignments of the three ppGpp-binding regions (β' 600-657, β' 340-399 and ω 1-59) were performed (Figure 3.1).

Organisms closely related to the γ-proteobacterium *E. coli*, such as α-proteobacterium *R. sphaeroides*, β-proteobacterium *B. pertussis,* and δ-proteobacterium *M. xanthus,* show high conservation of the ppGpp binding site. In ppGpp-binding region 1 on the β' subunit three of the four binding residues were identical or physicochemically similar in all proteobacteria. The fourth residue, Y626, was not conserved in any of these species suggesting it is less important for ppGpp binding. In *B. subtilis* and *T. thermophilus,* where ppGpp does not affect RNAP *in vitro*, residues K615 and I619 were not conserved and physicochemically dissimilar. D622 was conserved in all species and Y626 was only conserved in *T. thermophilus.* In *S. coelicolor*, D622 was conserved whilst K615, I619 and Y626 were all exchanged for physicochemically dissimilar residues.

In ppGpp-binding region 2 on the β' subunit, R362 was conserved or exchanged for a similarly charged lysine in all aligned sequences including *B. subtilis* and *T. thermophilus* where ppGpp does not affect RNAP *in vitro*. In *S. coelicolor,* a lysine residue was present at the aligned position.

In ppGpp-binding region 3 on the ω subunit, the ARVT motif was conserved in all aligned proteobacterial sequences. In *B. subtilis* and *T. thermophilus* sequences, the motif was poorly conserved with only one of four residues conserved in each. In *S. coelicolor*, none of four ppGpp-binding residues on the ω subunit were conserved.

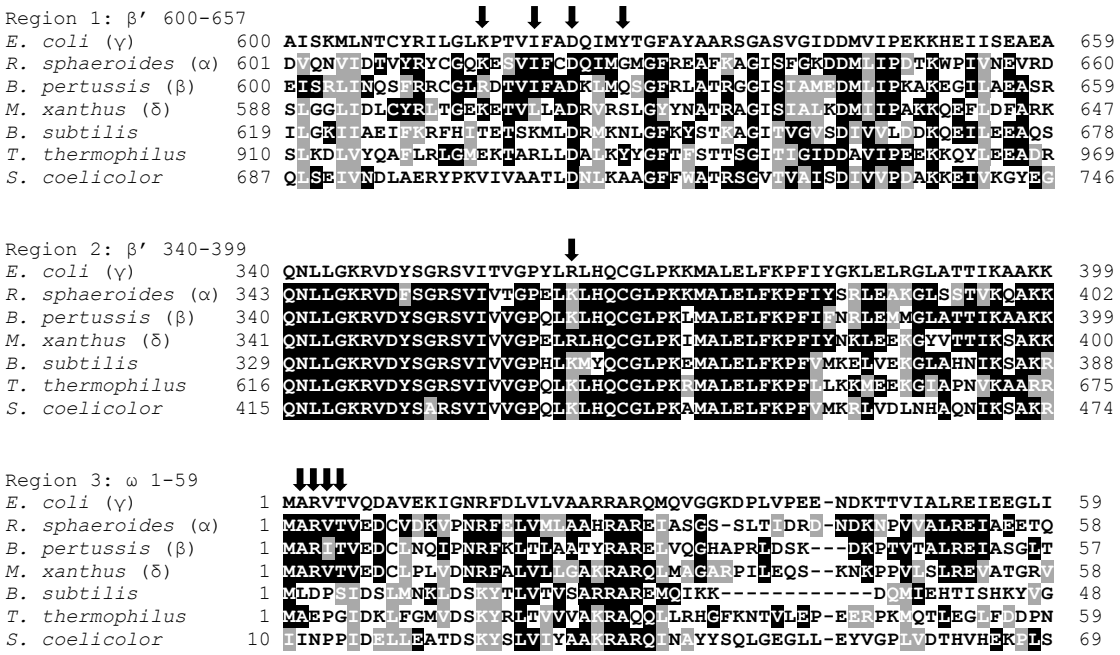


Figure 3.1 - The proposed ppGpp binding site is not conserved in *S. coelicolor*. A multiple sequence alignment of the amino acid sequence of three regions β' 600-657, β' 340-399 and ω 1-59, that form the proposed ppGpp binding site. Performed using CLUSTALW multiple sequence alignment tool (version 2.1). Box shading analysis was performed using BOXSHADE 3.21, using "consensus to a single sequence" function against *E. coli* sequences. Identical residues are shaded black, similar residues are shaded grey and residues predicted to contact ppGpp are indicated with arrows. For β' (*rpoC*) *Escherichia coli* K12-W3110, gene ECK3979; *Rhodobacter sphaeroides* 2.4.1, gene RSP_1712; *Bordetella pertussis* Tohama I, gene BP0016; *Myxococcus xanthus* DK 1622, gene MXAN_307; *Bacillus subtilis subtilis* 168, gene BSU01080; *Thermus thermophilus* HB8, gene TTHA1812; *Streptomyces coelicolor* A3(2), gene SCO4655. For ω' (*rpoZ*), *Escherichia coli* K12-W3110, gene ECK3639; *Rhodobacter sphaeroides* 2.4.1, gene RSP_1669; *Bordetella pertussis* Tohama I, gene BP1577; *Myxococcus xanthus* DK 1622, gene MXAN_4890; *Bacillus subtilis subtilis* 168, gene BSU15690; *Thermus thermophilus* HB8, gene TTHA1561; *Streptomyces coelicolor* A3(2), gene SCO1478.

3.3 Identification and bioinformatic analysis of DksA in *S. coelicolor*

Although ppGpp might not bind RNAP, it is still conceivable that DksA might bind RNAP and elicit a stringent response in *S. coelicolor*. To identify homologues of DksA in *S. coelicolor*, a BLASTP search was performed using the Department of Energy (DOE) Joint Genome Institute's (JGI) Integrated Microbial Genomes (IMG) service (http://img.jgi.doe.gov/). BLASTP searches with the *E. coli* DksA sequence against the *S. coelicolor* genome returned the hypothetical protein SCO2075 with an E value of 2×10^{-6}. A clustalW alignment of the two proteins showed that SCO2075 possesses an additional 81 amino acid, lysine-rich N-terminal extension (Figure 3.2A). A BLASTP search performed with this region revealed a histone H1-like region, similar to Hc2 in *Chlamydia trachomatis* (Pedersen et al., 1996). The alignment also revealed that the two highly conserved essential aspartic acid residues (D71 and D74) present at the coiled-coil tip of *E. coli* DksA are present in SCO2075, as is the conserved four cysteine zinc-finger motif (Tehranchi et al., 2010; Henard et al., 2014).

A structural prediction performed with Phyre2 modelled 130 residues (G109-R238) of SCO2075 using the DksA structure as the highest scoring template (Perederina et al., 2004; Kelley et al., 2015). The structure was highly similar to DksA and the two conserved aspartic acid residues (D164 and D167) were present at the tip of a coiled-coil structure (Figure 3.2B).

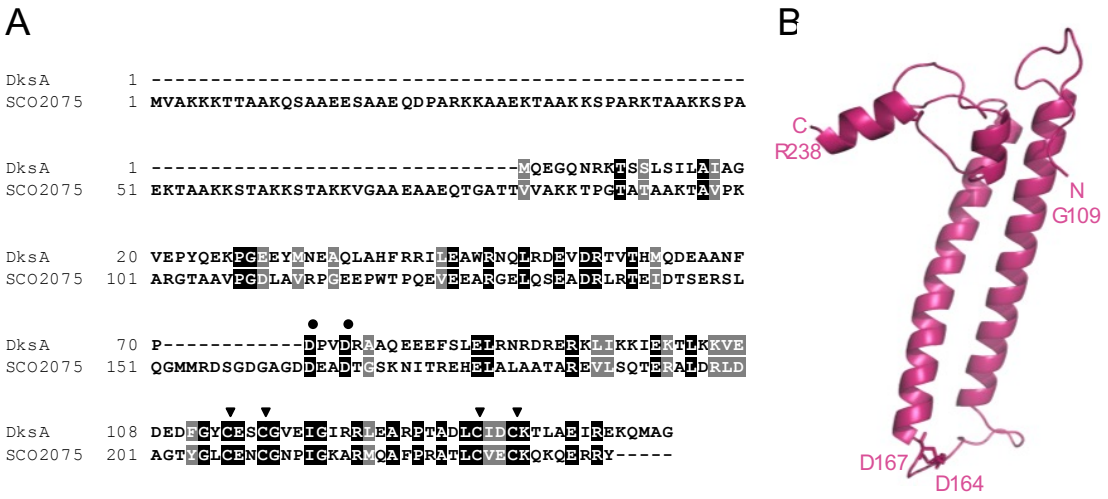


Figure 3.2 - **SCO2075 is a DksA homologue.** (A) A multiple sequence alignment of *E. coli* DksA (gene JW0141) and *S. coelicolor* SCO2075 performed using CLUSTALW multiple sequence alignment tool (version 2.1). Box shading analysis was performed using BOXSHADE 3.21. Identical amino acid residues are shaded black, similar residues are shaded grey, conserved aspartic acid residues are indicated with a black circle and conserved cysteine residues that form a zinc-finger motif are indicated with a black triangle. (B) Structural prediction of SCO2075 residues 108-239, prediction performed with Phyre2 and visualised with PyMOL (version 1.3). Conserved aspartic acid residues D164 and D167 side chains are labelled and shown as stick models.

Whilst searches with the *E. coli* DksA sequence only returned SCO2075, performing a BLASTP search with the SCO2075 amino acid sequence against the *S. coelicolor* genome returned two additional proteins, SCO6164 and SCO6165 with E values of 9×10^{-7} and 5×10^{-9}, respectively. They are encoded by neighbouring genes, separated by 70 bp so are potentially expressed in an operon. Multiple sequence alignments revealed that SCO6164 and SCO6165 lack the N-terminal histone-like region that SCO2075 possesses and are more similar in size to *E. coli* DksA (Figure 3.3). Phyre2 structural predictions modelled 104 and 115 residues of SCO6164 and SCO6165, respectively, with *E. coli* DksA as the top-scoring template. Similar to SCO2075, the predicted structures possessed a coiled-coil motif (data not shown). Unlike SCO2075, SCO6165 only had one conserved aspartic acid residue (D50) whilst SCO6164 had neither.

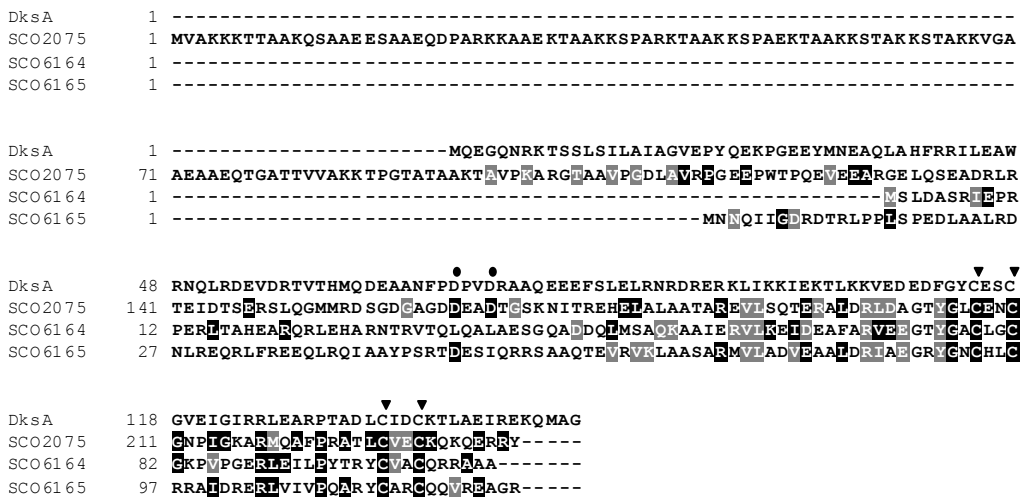


Figure 3.3 - ***S. coelicolor*** **possesses two additional DksA paralogues.** A multiple sequence alignment of the amino acid sequence of *E. coli* DksA (gene JW0141) and *S. coelicolor* SCO2075, SCO6164 and SCO6165. Performed using CLUSTALW multiple sequence alignment tool (version 2.1). Box shading analysis was performed using BOXSHADE 3.21, using "consensus to a single sequence" function against *E. coli* DksA sequence. Identical residues are shaded black, similar residues are shaded grey, conserved aspartic acid residues are indicated with a black circle and conserved cysteine residues that form a zinc-finger motif are indicated with a black triangle.

3.4 Deletion of DksA has no effect on growth

3.4.1 A Δ*dksA* (ΔSCO2075) mutant has no observable phenotype

One of the most commonly used methods for gene deletion in prokaryotes is deletion via homologous recombination of a mutant allele. In this method, large sections of DNA, homologous to the regions upstream and downstream of the target gene, are adjoined, creating a "mutant allele". Two crossovers can then occur, either simultaneously or sequentially, replacing the wild-type copy of the gene with a stable mutant allele. To allow positive selection of a mutant allele, a selectable marker such as an antibiotic resistance gene, can be placed between the two homologous regions.

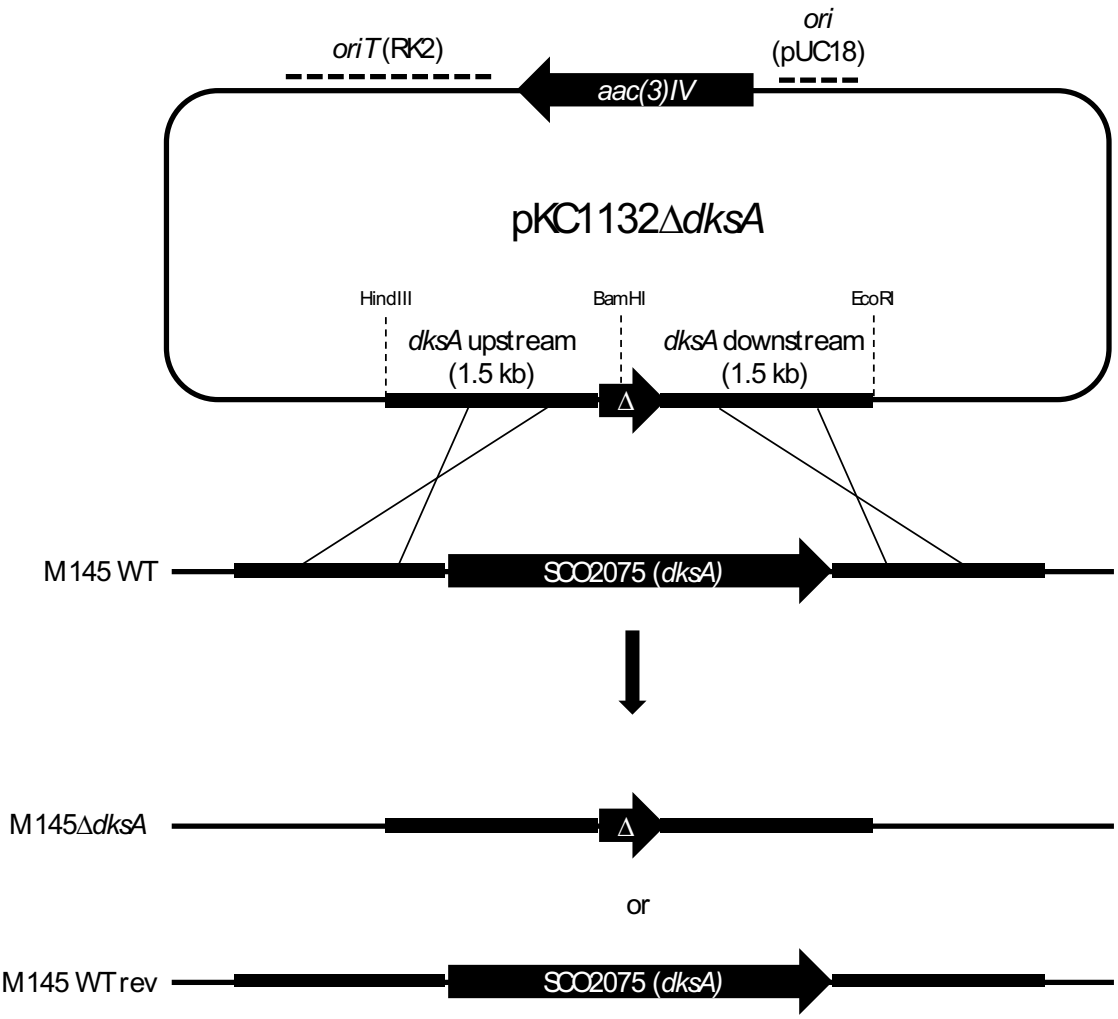


Figure 3.4 - **Construction of a Δ*dksA* deletion mutant**. The diagram shows the plasmid pKC1132Δ*dksA,* the homologous regions and the two possible stable double crossover strains, M145Δ*dksA* and M145 wild type revertant (M145 WT rev).

The strategy for creation of a Δ*dksA* (ΔSCO2075) mutant is outlined in Figure 3.4. To create a Δ*dksA* mutant allele, 1.5 kb upstream and downstream flanking regions were first PCR amplified from gDNA with the primers DksA_HindIII_F/DksA_BamHI_R and DksA_BamHI_F/DksA_EcoRI_R, respectively. The PCR products were cloned into EcoRV-cut pBlueScript SKII+ and sequenced to confirm successful amplification with no mutations. The flanks were combined by sub-cloning the upstream flank as a HindIII/BamHI fragment into pKC1132, followed by the downstream flank as a BamHI/EcoRI fragment, creating an in-frame markerless mutant allele in pKC1132, a conjugative, non-replicating plasmid. The mutant allele construct, pKC1132Δ*dksA* was confirmed by restriction analysis before transformation of *E. coli* ET12567/pUZ8002 and conjugation into *S. coelicolor* M145. Single crossover transconjugants were selected by

plating onto MS agar + apramycin. To allow a second crossover recombination event to occur, single crossover recombinants were passed through a round of non-selective growth on MS agar, spores were harvested and approximately 400 cfu were plated out onto MS agar. To screen for Δ*dksA* mutants, colonies were replica plated onto DN agar and DN agar containing apramycin. Apramycin-sensitive colonies were assumed to represent either Δ*dksA* mutants or wild type revertants. Δ*dksA* mutants were identified through colony PCR by amplification with the primer set dksA_O_F/dksA_O_R that bind either side of the *dksA* ORF.

A Δ*dksA* mutant strain was selected and grown on a variety of media including MS agar (a medium suitable for normal growth and sporulation) (Figure 3.5A), SMMS agar (a minimal media optimal for observing antibiotic production (Kieser et al., 2000)) (Figure 3.5B) and L agar (Figure 3.5C). Under all tested conditions, no phenotypic difference to wild type M145 was observed. It had previously been reported that when grown in the presence of 0.25 M KCl, a Δ*dksA* mutant overproduced actinorhodin. In this study, M145Δ*dksA* had no observed difference from M145 when grown on SMMS agar containing 0.25 M KCl (Figure 3.5D).

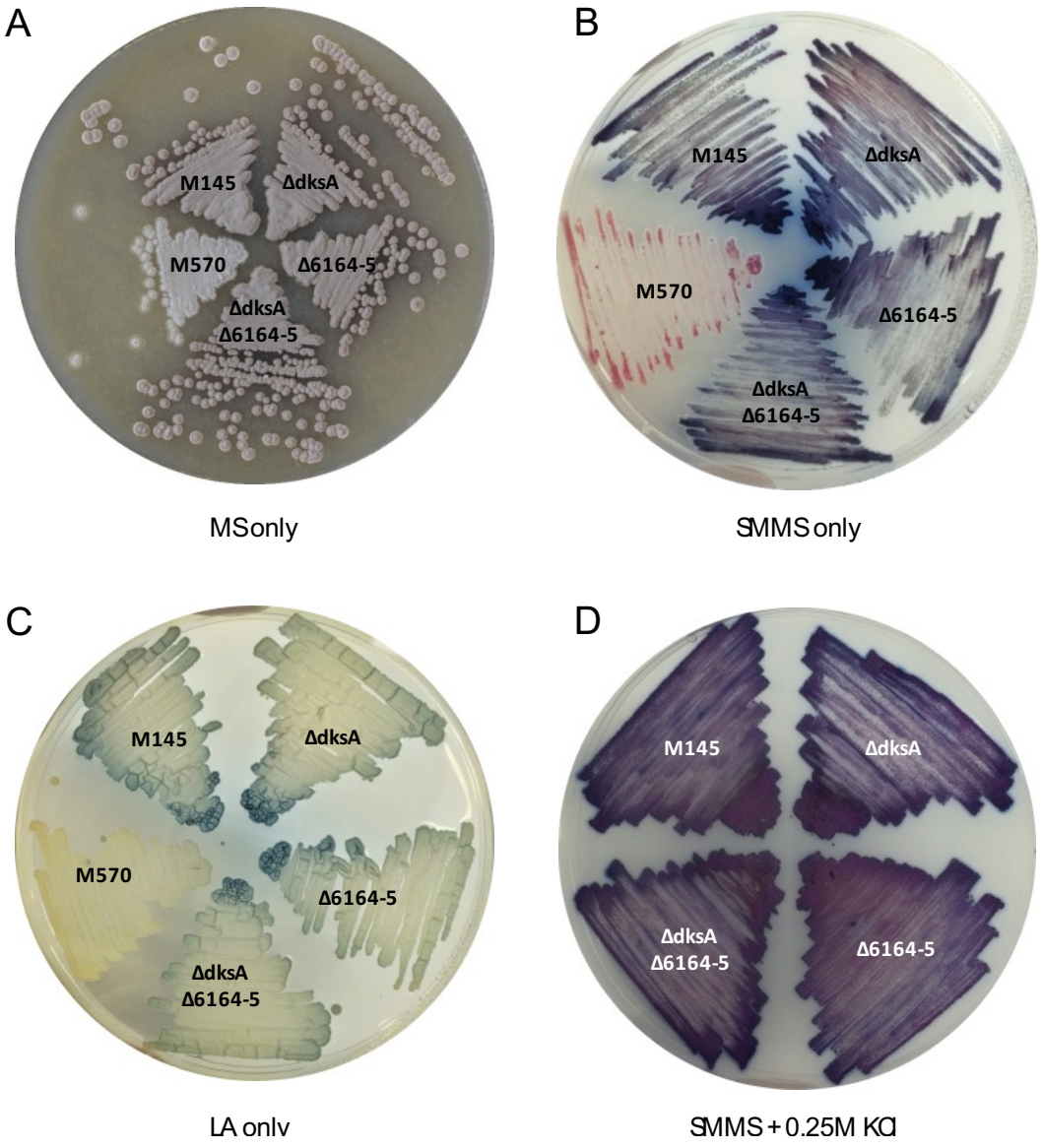


Figure 3.5 - **A Δ*dksA* mutant has no observable phenotype.** M145, M145Δ*dksA,* M145ΔSCO6164-5, M145Δ*dksA* ΔSCO6164-5 and M570 streaked onto (A) MS agar (B) SMMS agar (C) L agar and incubated for 5 days at 30 °C. (D) M145, M145Δ*dksA,* M145ΔSCO6164-5, and M145Δ*dksA* ΔSCO6164-5 streaked onto MS agar + 0.25M KCl and incubated for 5 days at 30 °C.

3.4.2 A Δ*dksA* ΔSCO6164 ΔSCO6165 triple mutant has no observable phenotype

As no obvious phenotype could be seen in a Δ*dksA* deletion mutant it was considered that there might be a functional redundancy between *dksA,* SCO6164 and SCO6165. To investigate this, ΔSCO6164 ΔSCO6165 (ΔSCO6164-5) double and Δ*dksA* ΔSCO6164-5 triple deletion mutant strains were made. As SCO6164 and SCO6165 are neighbouring genes, it was possible to delete both genes simultaneously. To create a ΔSCO6164-5 mutant allele, 1.5 kb upstream and downstream flanking regions were PCR amplified from gDNA with the primer sets 61645_up_for/61645_up_rev and 61645_down_for/61645_down_rev respectively. The PCR products were cloned into EcoRV-cut pBlueScript SKII+ and sequenced to confirm successful amplification with no mutations. The flanks were combined by sub-cloning both fragments sequentially into pKC1132, a conjugative, non-replicating plasmid, creating the plasmid pKC1132ΔSCO6164-5. The mutant allele construct was confirmed by restriction analysis before transformation of *E. coli* ET12567/pUZ8002 and conjugation into *S. coelicolor* M145 and M145Δ*dksA*. Single crossover transconjugants were selected by plating onto MS agar containing apramycin. As with creating the Δ*dksA* mutant, to allow a second crossover recombination event to occur single crossover recombinants were passaged through a round of non-selective growth on MS agar, spores were harvested and approximately 400 cfu were plated out onto MS agar. To screen for ΔSCO6164-5 mutants, colonies were replica plated onto DN agar and DN agar + apramycin. Apramycin-sensitive colonies were assumed to represent either ΔSCO6164-5 mutants or wild type revertants. ΔSCO6164-5 mutants were identified through colony PCR by amplification with the primer set 61645_col_for/61645_col_rev that bind either side of the SCO6164-5 region.

On MS agar, wildtype M145, M145Δ*dksA,* the double mutant M145ΔSCO6164-5 and triple mutant, M145Δ*dksA* ΔSCO6164-5 were identical in colony size, morphology and all other notable features (Figure 3.5A). For comparison, M570, a Δ*relA* strain unable to produce ppGpp and consequently elicit the stringent response, was also streaked out onto MS agar. Consistent with a published phenotype, M570 produced large, white colonies due to production of fewer mature spores and an inability to cease outgrowth in response to amino acid limitation. On SMMS agar, M145, M145Δ*dksA,* the double

mutant M145ΔSCO6164-5 and triple mutant, M145Δ*dksA* ΔSCO6164-5 were identical in production of actinorhodin (Figure 3.5B). In contrast and consistent with published observations, M570 was unable to produce actinorhodin under these conditions (Chakraburtty and Bibb, 1997). Similar to previous observations with M145Δ*dksA*, deletion of SCO6164-5 had no effect on growth on SMMS agar containing 250 mM KCl (Figure 3.5D).

3.5 Overexpression of DksA stimulates antibiotic production

3.5.1 Cloning *dksA* downstream of *ermEp** promoter stimulates ACT production

To further investigate the role of DksA, we aimed to overexpress the protein in *S. coelicolor* by placing it under the control of the strong, constitutive *ermEp** promoter on the integrative expression plasmid pSX162 (Bibb et al., 1986; Newell et al., 2006).

The *dksA* ORF was PCR amplified from gDNA using the primers dksA_O_F/dksA_O_R. The forward and reverse primers were designed to introduce a BamHI site 30 bp upstream of the start codon to ensure inclusion of the *dksA* ribosome binding site and a HindIII site downstream of the stop codon, respectively. The PCR product was cloned into EcoRV-cut pBlueScript SKII+, sequenced to confirm amplification with no mutations and sub-cloned into pSX162 as a BamHI/BamHI fragment (using a BamHI site downstream of the primer HindIII site from pBlueScript SKII+), creating the plasmid pSX162::*dksA*. The recombinant plasmid was confirmed by restriction digest analysis to ensure correct orientation of the BamHI/BamHI *dksA* fragment, used to transform *E. coli* ET12567/pUZ8002 and transferred via conjugation into *S. coelicolor* M145. As a negative control, the empty plasmid pSX162 was also transferred via conjugation into *S. coelicolor* M145. Transconjugants were selected by overlaying with apramycin.

When grown on SMMS agar, the strain overexpressing *dksA* from the constitutive *ermEp** promoter on pSX162 appeared to produce significantly more pigmented blue actinorhodin than the control strain harbouring the empty plasmid pSX162 (Figure 3.6A).

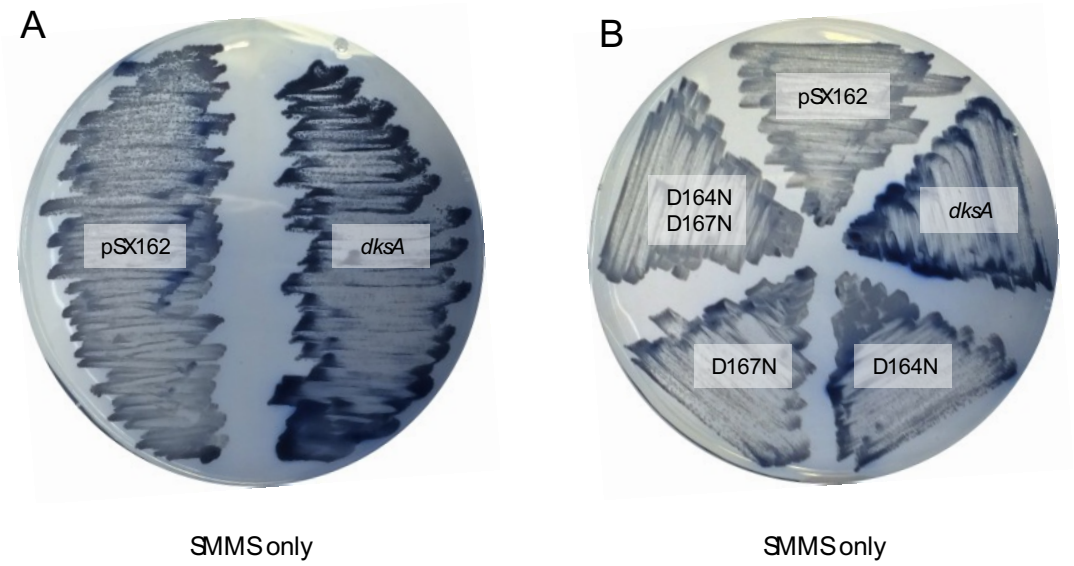


Figure 3.6 - **Overexpression of DksA stimulates actinorhodin production**. (A) M145/pSX162 and M145/pSX162::*dksA* streaked onto SMMS agar and incubated for 5 days at 30 °C. (B) M145/pSX162, M145/pSX162::*dksA,* M145/pSX162::*dksA*^{D164N}, M145/pSX162::*dksA*^{D167N} and M145/pSX162::*dksA*$^{D164N/D167N}$ streaked onto SMMS agar and incubated for 5 days at 30 °C.

3.5.2 Effect of *dksA* overexpression requires aspartic acid residues at coiled-coil tip

In *E. coli,* it has been shown that the two aspartic acid residues at the tip of the DksA are essential for function in transcription initiation (Perederina et al., 2004; Tehranchi et al., 2010). To study if the *dksA* overexpression phenotype was specific to the role of the protein in initiation, investigations were made into the importance of the conserved aspartic acid residues at the coiled-coil tip in overexpressed DksA. Site-directed mutagenesis by inverse PCR was performed on the template pBlueScript::*dksA* using the primers F_D164N/R_D164N, F_D167N/R_D167N and F_D1647N/R_D1647N, creating the alleles *dksA*^{D164N}, *dksA*^{D167N} and *dksA*$^{D164N/D167N}$, respectively. Sequencing was performed on transformants to confirm successful creation of these alleles with no additional mutations. As with creation of pSX162::*dksA*, each allele was sub-cloned into pSX162 as a BamHI/BamHI fragment and the correct orientation of the fragment was confirmed by restriction analysis. Recombinant plasmids pSX162::*dksA*^{D164N}, pSX162::*dksA*^{D167N} and pSX162::*dksA*$^{D164N/D167N}$ were used to transform *E. coli* ET12567/pUZ8002 and transferred via conjugation into *S. coelicolor* M145. Transconjugants were selected by overlaying with apramycin.

When grown on SMMS agar, as seen previously, the strain overexpressing *dksA* from the constitutive *ermEp** promoter on pSX162 appeared to produce significantly more pigmented blue actinorhodin than the negative control strain harbouring the empty plasmid pSX162 (Figure 3.6B). In strains overexpressing *dksA*^{D164N} or *dksA*^{D167N} alleles where just one of the aspartic acid residues had been mutated to an alanine residue, the amount of actinorhodin produced appeared to be more than the negative control however less than the strain overexpressing the wild type *dksA*. The strain overexpressing the *dksA*$^{D164N/D167N}$ allele, where both aspartic acid residues have been mutated to alanine residues produced no more actinorhodin than the negative control strain harbouring the empty plasmid pSX162.

3.6 Discussion

3.6.1 The proposed *E. coli* ppGpp binding site is not conserved in *S. coelicolor*

Following crystallography studies on *T. thermophilus* RNAP, it was initially proposed that ppGpp orchestrates the stringent response through binding RNAP at the active site (Artsimovitch et al., 2004). This binding site was later disproved following mutagenesis studies on the proposed ppGpp-binding residues (Vrentas et al., 2008). During the course of the current study, a new binding site for ppGpp at the interface between ω and β' subunits (Ross et al., 2013; Zuo et al., 2013). However, amino acid alignment of the proposed ppGpp-binding regions from diverse bacterial species indicated that key interacting residues are not conserved in all organisms. Furthermore, *in vitro* transcription reactions performed with RNAP from organisms that lack the key ppGpp-binding residues (such as *B. subtilis* and *T. thermophilus*) revealed that ppGpp has no effect on transcription (Krásný and Gourse, 2004; Vrentas et al., 2008). Similar alignments performed in this study have identified that the proposed ppGpp-binding site is not conserved in *S. coelicolor.* Of the nine residues identified in *E. coli*, only two are conserved or similar in *S. coelicolor.* This suggests that if ppGpp is able to bind to RNAP in *S. coelicolor* it must be through an alternative mechanism or binding site altogether.

3.6.2 How does ppGpp exert the stringent response in *S. coelicolor*?

In *B. subtilis*, it was proposed that that ppGpp exerts the stringent response through modulation of intracellular GTP levels, although the underlying mechanism was initially unclear (Krásný and Gourse, 2004). The recent discovery that ppGpp directly inhibits the activity of enzymes involved in biosynthesis of GTP, the initiating NTP for rRNA promoters, appears to have answered this question (Liu et al., 2015). Experiments performed by Liu et al. (2015) identified the mechanism through which ppGpp competitively inhibits GMK activity. This activity was demonstrated on purified GMK enzymes from a number of bacterial species including *B. subtilis*, *T. thermophilus* and notably *S. coelicolor,* but not on the enzymes from most tested proteobacterial species including *E. coli*.

With the exception of two alphaproteobacterial species, *A. tumefaciens* and *S. meliloti*, the two mechanisms for regulation of transcription by ppGpp on RNAP and GMK appear to be distinctly conserved, with organisms exclusively using one mechanism or the other

(Liu et al., 2015). This suggests that ppGpp regulates the stringent response in *S. coelicolor* through regulation of intracellular GTP concentrations.

For effective regulation of transcription by intracellular GTP concentrations, the promoters would be expected to initiate transcription with a GTP at the +1 position. In *B. subtilis*, all of the *rrnB* and *rrnO* promoters initiate with GTP and are stringently controlled (Krásný and Gourse, 2004).

S1 nuclease mapping experiments on the *S. coelicolor, rrnD* region identified four promoters (Baylis and Bibb, 1988) each of are stringently controlled (Strauch et al., 1991) (Figure 3.7). Whilst transcription was predicted to initiate from each promoter with CTP or ATP, the authors note that due to experimental limitations this could vary by 1 bp. Notably, transcription 1 bp downstream from each of the predicted transcription start sites would initiate with GTP. Given that purine nucleotides are highly preferred for transcription initiation in bacteria (Basu et al., 2014), it seems likely that initiation at these promoters initiates with GTP.

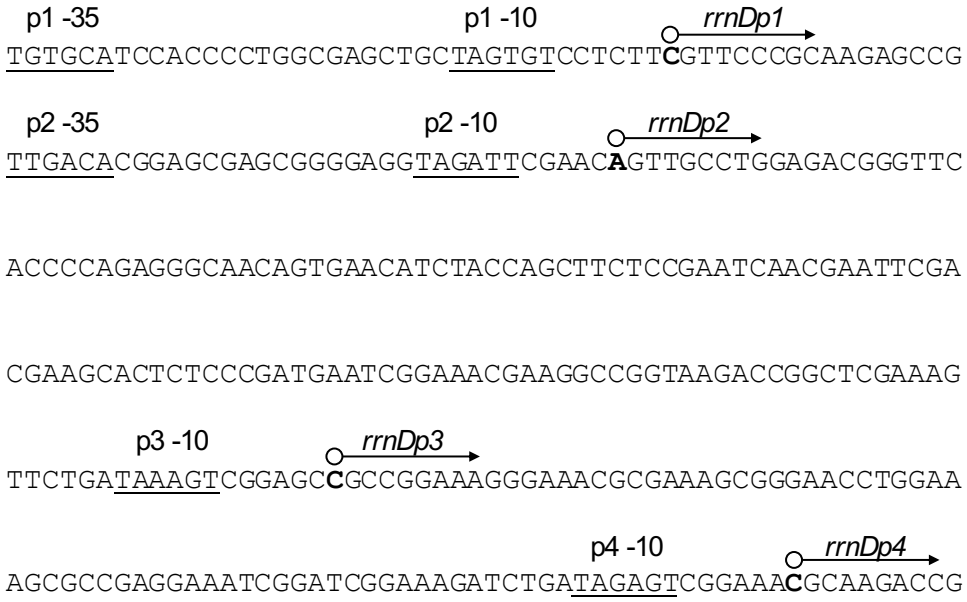


Figure 3.7 - ***rrnDp1-4* with predicted promoter regions and transcription start sites**. DNA sequence of the *S. coelicolor rrnD* promoter region. Open circles indicate the transcription start sites predicted by Baylis and Bibb (1988), predicted σ^{HrdB} -35 and -10 promoter regions are underlined.

3.6.3 *S. coelicolor* possesses a DksA homologue although it is not important for growth

In *E. coli*, ppGpp and DksA directly regulate transcription and the stringent response through binding RNAP. Liu et al. (2015) reported that presence of DksA homologue was restricted to organisms where ppGpp binds RNAP and directly regulates transcription. It was also suggested that DksA was absent from organisms where GMK activity is directly inhibited by ppGpp, including *S. coelicolor*. In contrast to this finding, BLASTP searches performed in this study with the *E. coli* DksA sequence identified a DksA homologue, SCO2075, and two putative paralogues, SCO6164 and SC06165, in *S. coelicolor.* Previous searches performed by Liu et al. (2015) used sequence homology and presence of a DxxDxA motif at the coiled-coil tip to identify DksA. As *S. coelicolor* SCO2075 has a DxxDxG motif at the coiled-coil tip this could have prevented previous identification.

This chapter details the creation of *S. coelicolor* Δ*dksA* single deletion mutant, ΔSCO6164-6165 double mutant and Δ*dksA* ΔSCO6164-6165 triple mutant. Under all studied conditions, deletion of any combination of these genes had no observed effect on growth. Diverse roles for DksA have been suggested outside of transcription initiation in concert with ppGpp, including resolving conflict between transcription and replication machinery (Tehranchi et al., 2010). If indeed ppGpp does not act directly on transcription initiation in *S. coelicolor,* it is possible that the function of *S. coelicolor* DksA is exclusively in regulation of transcription elongation. Due to the reported functional overlap of the Gre factors and DksA (Zenkin and Yuzenkova, 2015), it is possible that presence of *S. coelicolor* Gre factors are complementing the role of DksA in the strains created in this study. To further clarify the role of *S. coelicolor* DksA it may be necessary to investigate the function of Gre factors.

3.6.4 Overexpression of DksA induces antibiotic production in *S. coelicolor*

This chapter also details the overexpression of DksA in *S. coelicolor*. When placed under the control of the strong, constitutive *ermEp** promoter, overexpression of DksA led to overproduction of the blue pigmented antibiotic actinorhodin. Experiments performed by Aldridge et al. (2013) showed that overexpression of *S. coelicolor* DksA from a *tipAp* promoter had no effect on antibiotic production in an M145 background. Whilst this

result appears to contradict the findings of this study, it is possible that expression from *tipAp* is significantly lower than from *ermEp** promoter and was insufficient to cause the observed phenotype in this study. Aldridge et al. (2013) showed that overexpression of DksA from a *tipAp* promoter restored production of both actinorhodin and prodigines in a *relA* mutant background. This demonstrates that DksA is able to work in a ppGpp-independent manner.

Site-directed mutagenesis was performed on the aspartic acid residues at the coiled-coil tip, creating alleles with one or both residues changed to an asparagine residue. Overexpression of alleles without both aspartic acid residues intact had no effect on actinorhodin production. This suggests that the DksA overexpression phenotype is a consequence of DksA binding to RNAP and inserting its coiled-coli into the secondary channel. Separation-of-function studies have shown that the conserved aspartic acid residues are required for the role of DksA in transcription initiation but are not required for the role of DksA on elongation preventing replication arrest (Perederina et al., 2004; Tehranchi et al., 2010). This suggests that despite the fact ppGpp does not appear to directly affect RNAP in *S. coelicolor,* the overexpression phenotype observed is due to effects of DksA on transcription initiation.

Chapter 4:

Results II: Investigating the role of CarD in *S. coelicolor*

4 Investigating the role of CarD in *S. coelicolor*

4.1 Overview

CarD was identified as an essential transcriptional activator in *M. smegmatis* and *M. tuberculosis* (Stallings et al., 2009). It was initially thought to be a negative regulator of rRNA, was able to rescue an *E. coli* Δ*dksA* null mutant and was therefore thought to have a role in the stringent response. The stringent response is closely linked to the production of antibiotics and other useful secondary metabolites in *S. coelicolor* and therefore, with the rising need for new compounds effective against multidrug resistance infections, it is important to understand the regulation and synthesis of such products. This chapter sets out to identify the role of CarD in *S. coelicolor*, initially by confirming its importance for growth *in vivo* and its role as a transcriptional regulator through *in vitro* transcription assays. This chapter also details the purification and analysis of recombinant CarD and RNAP from *E. coli* and *S. coelicolor*, respectively.

4.2 Identification and bioinformatic analysis of CarD in *S. coelicolor*

4.2.1 Conservation of CarD in *Streptomyces* species

CarD was initially identified as being highly conserved in all sequenced mycobacteria with *M. smegmatis* CarD sharing 98.1% and 95.7% homology with CarD proteins found in *M. tuberculosis* and *M. leprae*, respectively (Stallings et al., 2009). A BLASTP search identified SCO4232 as a putative CarD orthologue in *S. coelicolor,* with the NCBI accession number GI:7242753. The *S. coelicolor carD* gene is 483 bp in length and encodes a 160 amino acid protein. A number of other bacterial species including the fellow *Actinobacteria Corynebacterium glutamicum* and diverse non-actinobacterial species such as *Bacillus subtilis* and gram-negative species such as *Caulobacter crescentus* and *Thermus thermophilus* were also found to possess a CarD homologue.

To identify homologues of CarD found in other *Streptomyces* species, a BLASTP search was performed using the Integrated Microbial Genomes (IMG) service (http://img.jgi.doe.gov/). The SCO4232 amino acid sequence, obtained from the *Streptomyces* Annotation Server StrepDB (http://strepdb.streptomyces.org.uk), was used to search the database of all 25 finished *Streptomyces* genomes. A CarD homologue was

found in every annotated finished genome with the conservation ranging from 100% to 98.1% (GI:477544984, *S. albus* J1074).

A TBLASTN search performed on CarD against a number of *Streptomyces* genomes identified that the nucleotide sequence of the ORF was very highly conserved. More interestingly, a CLUSTALW alignment showed that the region directly upstream of the *carD* ORF was very highly conserved in *Streptomyces* spp. for approximately 300 bp (Figure 4.1). This raises the possibility of an upstream regulatory element such as a transcription factor binding site, riboswitch or even an unannotated gene or RNA, potentially relevant to the role of CarD.

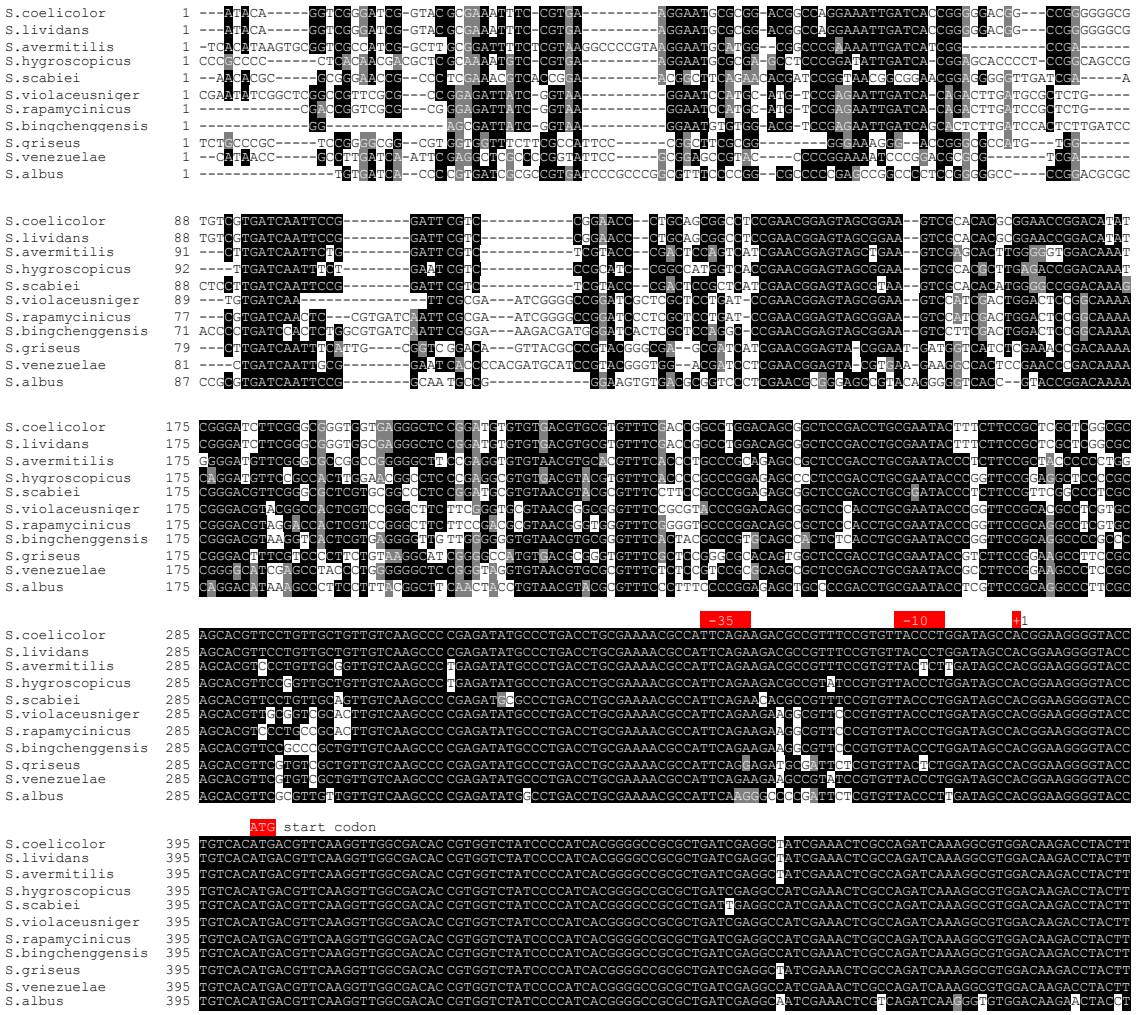


Figure 4.1 - **Multple sequence alignment of the *carD* upstream region.** A multiple sequence alignment of the first 107 bp of the *carD* ORF and 400 bp of upstream DNA from a range of *Streptomyces* spp., performed using CLUSTALW multiple sequence alignment tool (version 2.1). Box shading analysis was performed using BOXSHADE 3.21. Identical residues are shaded black, similar residues are shaded grey.

4.2.2 Conservation of CarD in the *Actinobacteria*

To look more broadly at the conservation and distribution of CarD, a BLASTP search was performed against diverse representatives of the *Actinobacteria* order (Gao and Gupta, 2012). CarD was present and well conserved in each of these representatives with percentage identities compared to *S. coelicolor* CarD ranging from 87% in *Frankia alni* to 58% in *Bifidobacterium longum* (Figure 4.2). Homologues were all similar in length with the exception of CarD from *B. longum* which had a 35 amino acid C-terminal extension and contributed to its poor identity score.

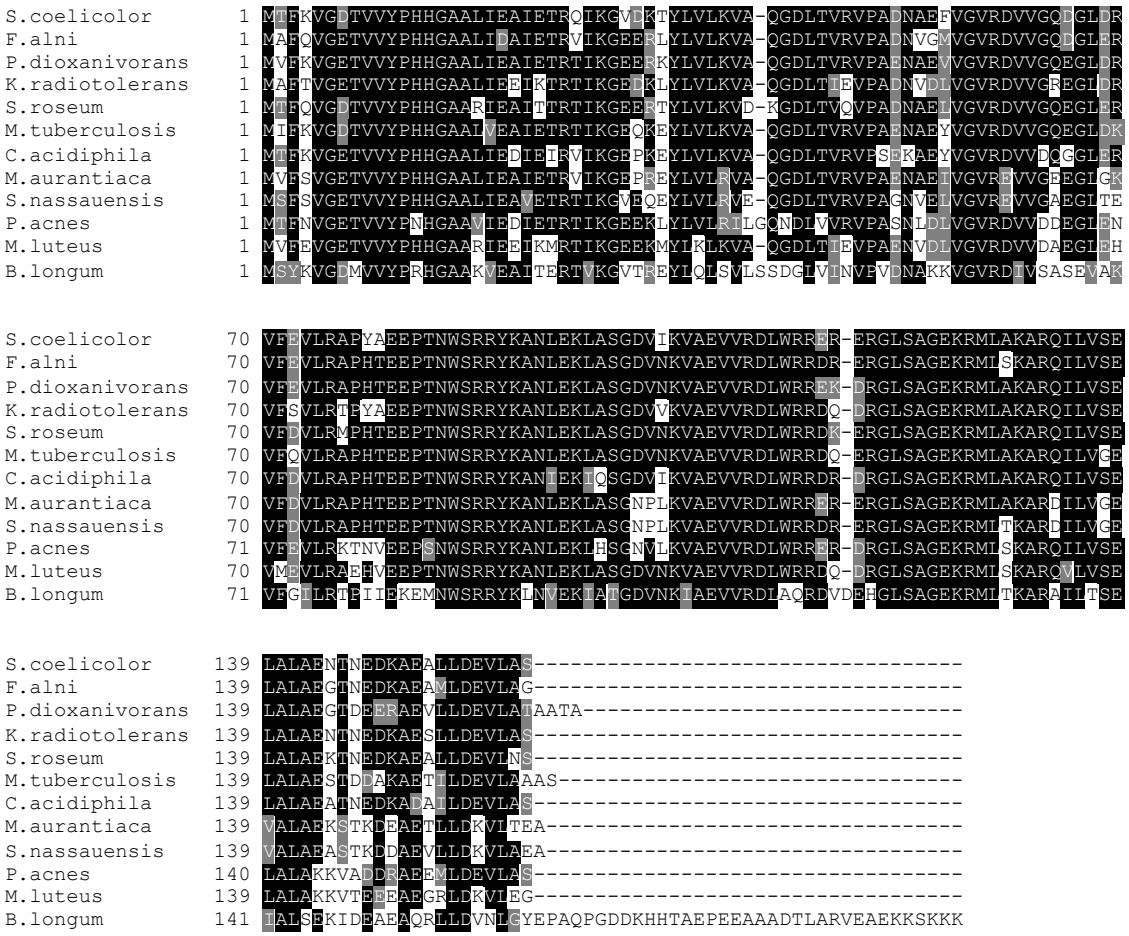


Figure 4.2 - **CarD distribution in the *Actinobacteria*.** A multiple sequence alignment of CarD homologues found in diverse representatives of the *Actinobacteria* order, performed using CLUSTALW multiple sequence alignment tool (version 2.1). Box shading analysis was performed using BOXSHADE 3.21. Identical amino acid residues are shaded black, similar residues are shaded grey.

Strain	Gene name	Accession	Length	Identity (%)
Streptomyces coelicolor	SCO4232	Q9L0Q9	160	100
Frankia alni	FRAAL6525	Q0RBN6	160	86.88
Pseudonocardia dioxanivorans	Psed_6102	F4CLS3	164	86.88
Kineococcus radiotolerans	Krad_0898	A6W6E7	160	86.25
Streptosporangium roseum	Sros_0963	D2B9G1	160	86.25
Mycobacterium tuberculosis	Rv3583c	P9WJG3	162	85
Catenulispora acidiphila	Caci_8145	C7QIR8	160	84.38
Micromonospora aurantiaca	Micau_5835	D9T0G9	161	78.75
Stackebrandtia nassauensis	Snas_0142	D3Q1L4	161	76.88
Propionibacterium acnes	PPA0357	Q6AAV4	161	74.38
Micrococcus luteus	Mlut_03770	C5C8X3	160	72.5
Bifidobacterium longum	BLD_0751	B3DSS8	197	58.12

Table 4.1 - **CarD distribution in the *Actinobacteria*.** A table showing the gene name, accession, length and percentage sequence identity of CarD homologues in diverse representatives of the *Actinobacteria* order, performed using CLUSTALW multiple sequence alignment tool (version 2.1)

4.2.3 RNA-seq analysis of *carD* in exponentially growing *S. coelicolor*

With the decreasing cost of high-throughput next generation sequencing experiments, an increasing number of datasets are becoming publicly available, allowing further analysis beyond initial publication. Romero et al. (2014) performed RNA-seq on *S. coelicolor* M145 liquid cultures in late exponential phase, enriching for mRNA by depleting 16S and 23S rRNA. The study identified a number of RNA processing and degradation sites as well as providing a single-nucleotide resolution map of the transcriptional landscape. The RNA-seq data for this study was deposited in the GEO archive under the accession number GSM1126846 in the form of aligned BEDGRAPH files representing the forward and reverse strands. Using Integrated Genome Browser (IGB) (Version 8.3.4) aligned BEDGRAPH files were visualised against the *S. coelicolor* (A3)2 genome sequence and annotation downloaded from the NCBI genome database (Figure 4.3).

Enrichment at the *carD* locus shows gene expression in late exponential phase. The lack of enrichment on the positive strand downstream from the ORF suggests that *carD* is monocistronic. Interestingly, the putative "cutoRNA" named *scr4233* is located directly downstream of *carD* on the opposite strand, defined as a convergent untranslated overlapping RNA (Moody et al., 2013). Presence or function of this RNA has yet to be determined experimentally.

The *carD* transcription start site could be identified at position 4636723, 21 nucleotides upstream from the ORF start codon. Further upstream, σ^{HrdB} promoter-like sequences can be identified at positions -10 and -35 (TACCCT and TTCAGA, respectively, separated by 17 bp) (Figure 4.3B).

In contrast to *rbpA* which is regulated by a σ^{R}-dependent promoter and is induced in response to diamide, microarray experiments performed by Kallifidas et al. (2010) identified *carD* as a gene downregulated in response to diamide in both wild-type and Δ*sigR* backgrounds. A similar downregulation response is also seen in a large number of genes involved in protein and RNA synthesis such as ribosomal proteins and RNA polymerase subunits. The specific mechanism for this downregulation is unknown.

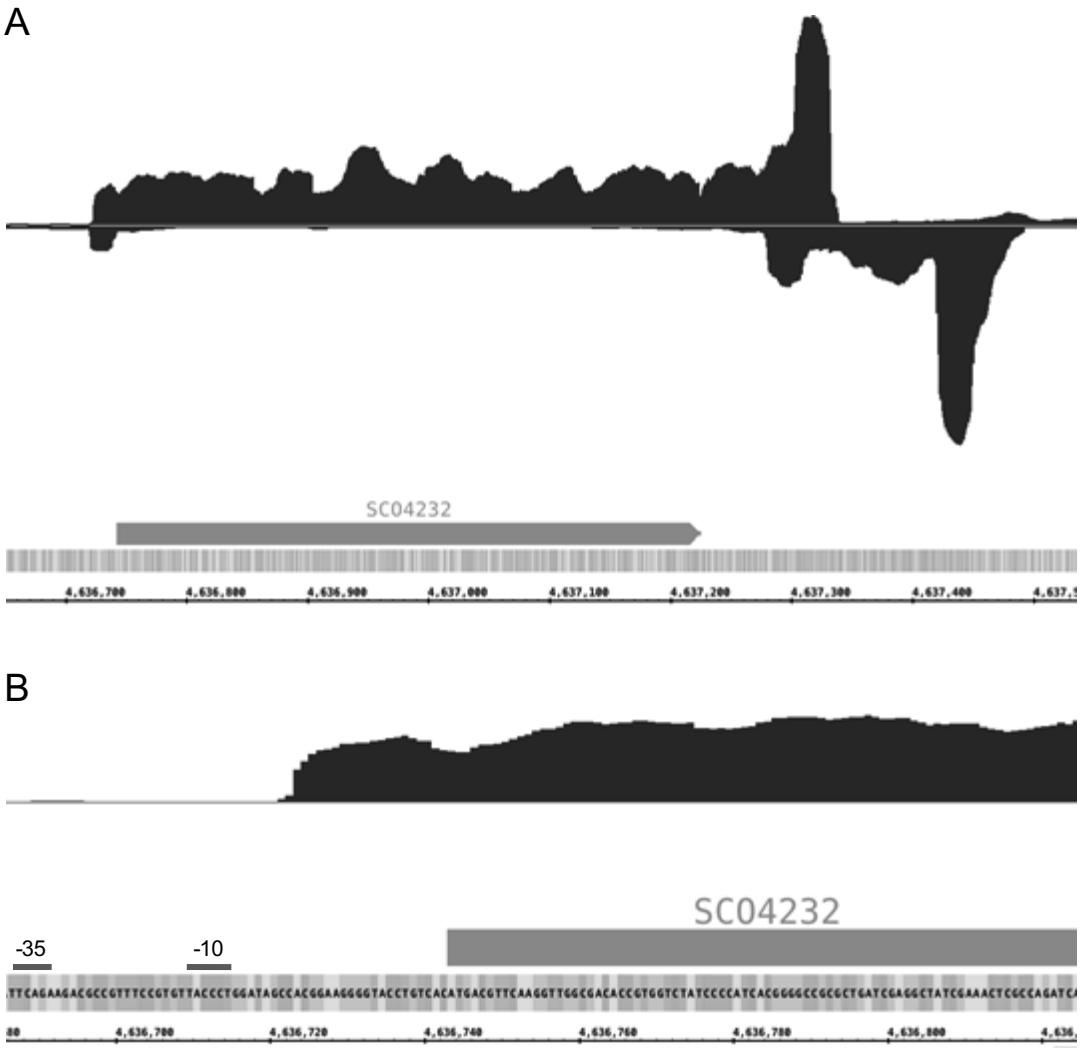


Figure 4.3 - **Analysis of *carD* mRNA transcript through RNA-seq** (A) RNA-seq analysis of the *carD* gene region and downstream sRNA *scr4233* (B) RNA-seq analysis of the *carD* TSS and promoter region. RNA-seq data was downloaded from the GEO archive under the accession number GSM1126846 in the form of aligned BEDGRAPH files representing the forward and reverse strands. Data was visualised with Integrated Genome Browser (IGB) (Version 8.3.4)

4.3 *carD* is an essential gene in *S. coelicolor*

CarD has been identified as essential for viability in both *M. smegmatis* and *M. tuberculosis* (Stallings et al., 2009). Attempts to delete the gene were unsuccessful, although it was possible to create a conditional mutant strain by expressing *carD* under the control of a tetracycline-inducible promoter and knocking out the wild-type copy of the gene. This experiment was performed in *M. smegmatis* and *M. tuberculosis* and both strains only grew in the presence of anhydrotetracycline. This demonstrated that CarD was essential for viability. To investigate if this was also the case for *S. coelicolor,* we aimed to make a *carD* deletion mutant and investigate the phenotype.

4.3.1 Construction of a *carD* deletion mutant

The strategy for creation of a ΔcarD mutant is outlined in Figure 4.4. To create a Δ*carD*::*hyg* mutant allele, 1.5 kb upstream and downstream flanking regions were PCR amplified from *S. coelicolor* gDNA using the primer pairs carD_eco_F/carD_bam_R and carD_bam_F/carD_xba_R, respectively. The PCR products were cloned into EcoRV-cut pBlueScript SKII+ and sequenced to confirm successful amplification with no mutations. They were combined in pBlueScript SKII+ by subcloning the downstream flank as a BamHI/XbaI-cut fragment, creating an in-frame markerless mutant allele in which 130 of 161 codons had been deleted. To allow positive selection of this mutant allele, a hygromycin (*hyg*) resistance cassette was subcloned from pIJ963 to between the two flanking regions in the *carD* ORF as a BglII-cut fragment, creating a Δ*carD::hyg* mutant allele. The Δ*carD::hyg* mutant allele was digested with XbaI/HindIII, blunt-ended and sub-cloned into EcoRV-cut pIJ6650, a conjugative, non-replicating plasmid. The mutant allele construct was confirmed by restriction analysis before transformation of *E. coli* ET12567/pUZ8002 and conjugation into *S. coelicolor* J1915, an M145Δ*glkA* (glucose kinase) deletion mutant. Regardless of additional carbon sources, strains possessing GlkA are sensitive to the glucose analogue 2-deoxyglucose (2-DOG). This allows selection of strains lacking *glkA* by growth in the presence of 2-DOG; by using a mutant strain lacking wild-type *glkA* (such as J1915), it can therefore be used as a counter-selectable marker allowing for selection of loss of *glkA* present on a plasmid backbone (Van Wezel and Bibb, 1996). Single crossover transconjugants were selected by plating onto MS agar containing apramycin. To select for the loss of the *glkA* present on the pIJ6650 backbone, single

crossover transconjugants were plated onto minimal media agar (with 0.5% mannitol as a carbon source) containing 2-DOG and hygromycin. Despite screening numerous transconjugants, no 2-DOG^R Hyg^R colonies could be isolated suggesting that *carD* is an essential gene in *S. coelicolor*.

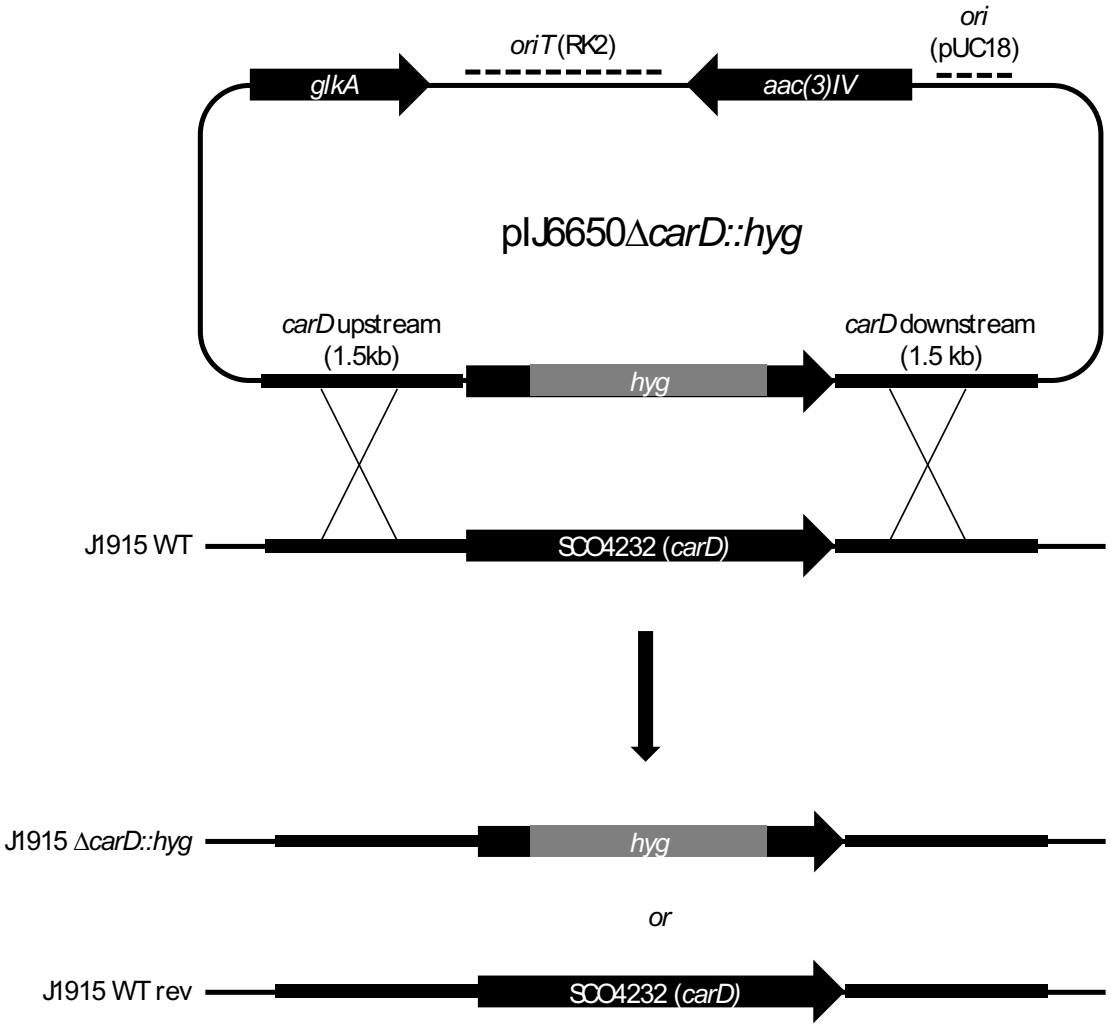


Figure 4.4 - **Construction of a Δ*carD* deletion mutant**. The diagram shows the plasmid pIJ6650Δ*carD::hyg,* the homologous regions and the two possible stable double crossover strains, J1915Δ*carD::hyg* and J1915 wild type revertant (J1915 WT rev).

As it was not possible to obtain a double crossover deletion mutant, we instead aimed to make a conditional *carD* deletion mutant (Figure 4.5).

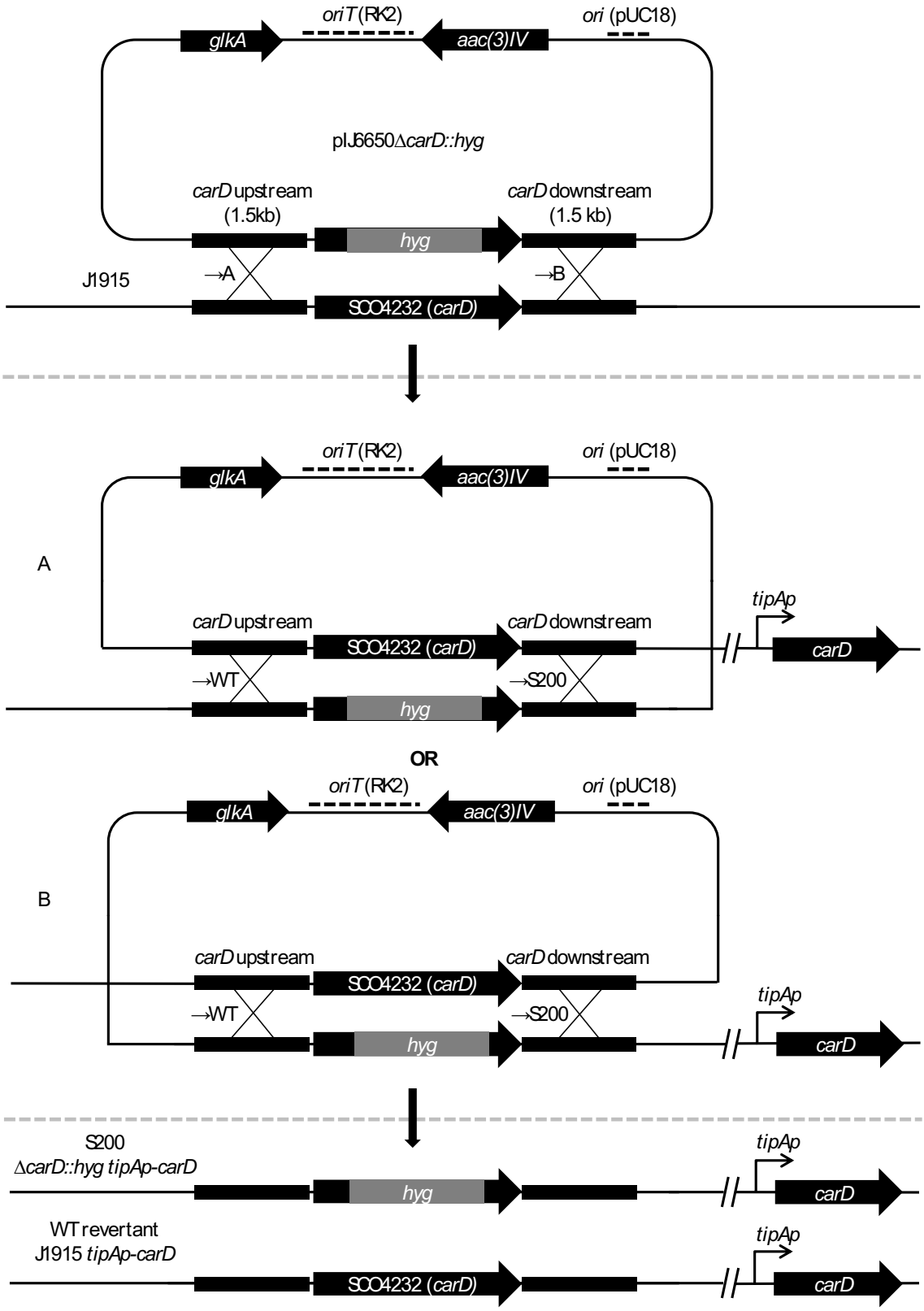


Figure 4.5 - **Construction of a conditional *carD* deletion mutant** The diagram shows the plasmid pIJ6650Δ*carD::hyg,* the homologous regions and the two initial single crossover mutants, "A" and "B". Following expression of carD from the *tipAp* promoter, it was possible to obtain two stable double crossover strains, J1915Δ*carD::hyg/tipAp-carD* (S200) and J1915/*tipAp-carD* wild type revertant (WT rev).

carD was placed under the control of the thiostrepton-inducible *tipAp* promoter in pIJ6902. The *carD* ORF was PCR amplified from gDNA using the primer set carD_D_F/carD_O_R, designed to include an NdeI site overlapping the ORF start codon and a BamHI site downstream from the stop codon, respectively. The PCR product was cloned into EcoRV-cut pBlueScript SKII+, sequenced to confirm amplification with no mutations and sub-cloned into pIJ6902 as an NdeI/BamHI fragment, creating the plasmid pIJ6902::*carD*. The recombinant plasmid was confirmed by restriction digest analysis, used to transform *E. coli* ET12567/pUZ8002 and transferred via conjugation into a J1915/pIJ6650Δ*carD::hyg* single crossover mutant, selecting for transconjugants with thiostrepton.

Transconjugants were plated out onto MM containing 2-DOG and thiostrepton to force the loss of *glkA* and "healthy" colonies were picked onto MS containing thiostrepton. It was assumed that 2-DOG^R Thio^R colonies isolated at this stage were either Δ*carD* mutants or wild-type revertants. To screen for Δ*carD* mutants, colonies were replica plated onto DN agar, DN agar containing thiostrepton and DN agar containing thiostrepton and hygromycin; colonies that required thiostrepton for growth and were resistant to hygromycin would represent Δ*carD::hyg* mutants expressing the essential *carD* from the *tipAp* promoter. Three independent clones were isolated with the expected phenotype representing the strain J1915Δ*carD::hyg/tipAp-carD*. To confirm these phenotypes, these independent clones were streaked onto a number of selective plates and compared to a wild-type revertant (J1915/*tipAp-carD*) and the parent strain J1915 (Figure 4.6). J1915Δ*carD::hyg/tipAp-carD* #1, #2 and #3 grew very slowly on MM (Figure 4.6A), but were all partially complemented on MM + thio (Figure 4.6B). All three strains were viable in the presence of hyg and 2-DOG (Figure 4.6C–F) but still grew poorly in the absence of thiostrepton. J1915 was only viable on MM and in the presence of 2-DOG, whilst the wild-type revertant J1915/*tipAp-carD* was only viable in the presence of 2-DOG and thio (Figure 4.6). Notably, the three J1915Δ*carD::hyg/tipAp-carD* mutants were resistant to hygromycin and 2-DOG but dependent on thiostrepton. J1915Δ*carD::hyg/tipAp-carD* independent clone #1 was chosen and named S200.

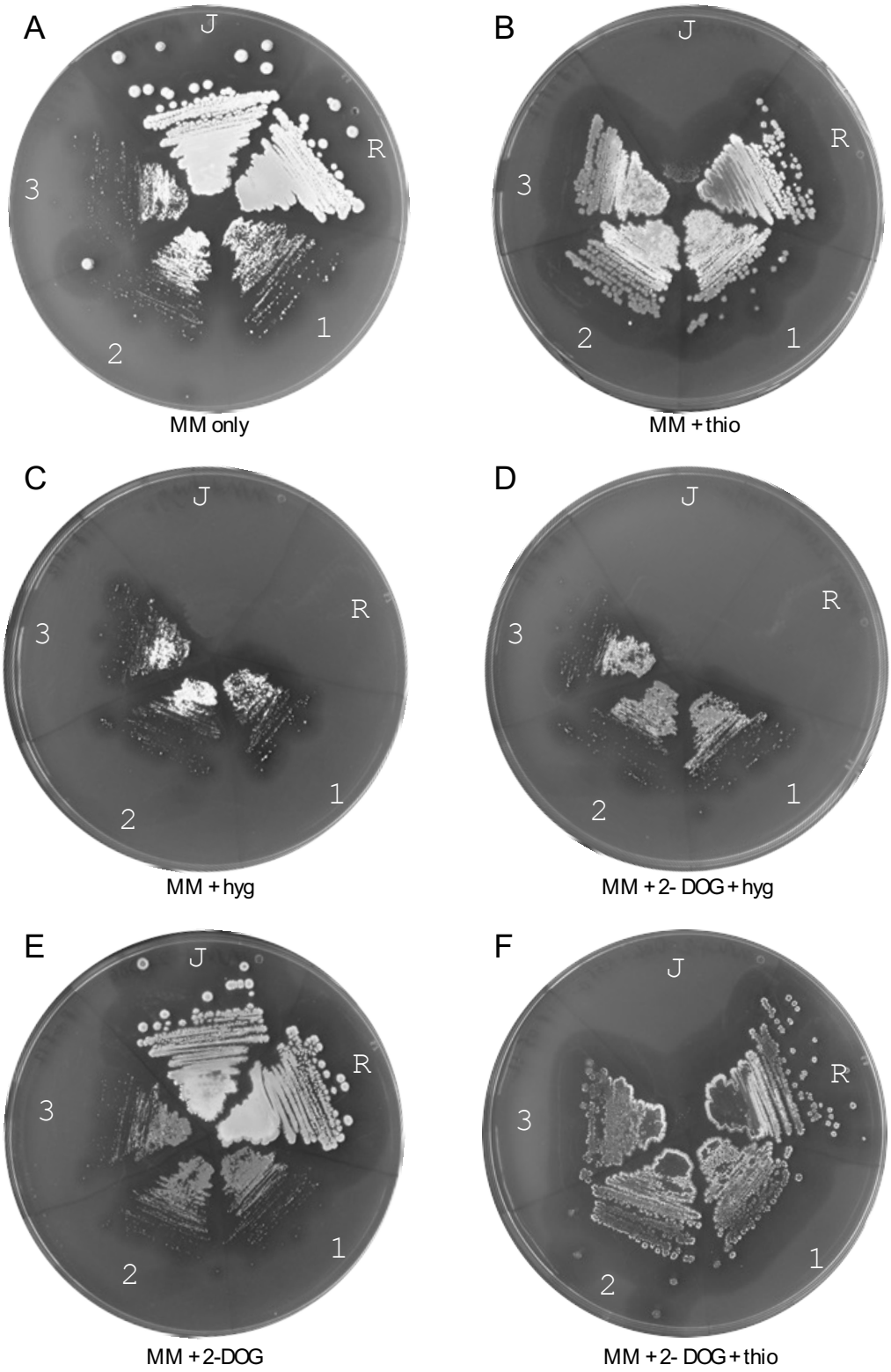


Figure 4.6 - **Phenotypic analysis of three conditional *carD* deletion mutants.** Strains streaked out onto (A) MM only; (B) MM + thio; (C) MM + hyg; (D) MM + 2-DOG + hyg; (E) MM + 2-DOG; (F) MM + 2-DOG + thio and incubated for 5 days at 30 °C. J: J1915; R: J1915/tipAp-*carD* revertant; 1, 2, 3: J1915Δ*carD::hyg*/tipAp-*carD* independent mutants.

4.3.2 Phenotypic analysis of Δ*carD* mutant

S200 (J1915Δ*carD::hyg/tipAp-carD)* and J1915/*tipAp-carD* were streaked out onto MS only and MS containing a range of thiostrepton concentrations (5 µg/mL, 15 µg/mL and 25 µg/mL) and incubated for 5 days at 30 °C. In the absence of thiostrepton, S200 was barely viable (Figure 4.7A). Some growth was visible where concentrated spores were plated and very small colonies could be observed throughout the streak. At a range of thiostrepton concentrations, S200 growth was restored. Colonies were smaller than wild-type at 5 µg/mL thiostrepton but restored to wild-type size at 12 and 25 µg/mL thiostrepton (Figure 4.7B–D). Despite the ability of thiostrepton to complement growth, at all concentrations colonies were lighter in colour, reminiscent of a *whi* phenotype for mutants blocked in late stages of sporulation (Flärdh and Buttner, 2009). Growth of wild-type revertant J1915/*tipAp-carD* was consistent with a wild-type phenotype in absence and presence of thiostrepton at all concentrations.

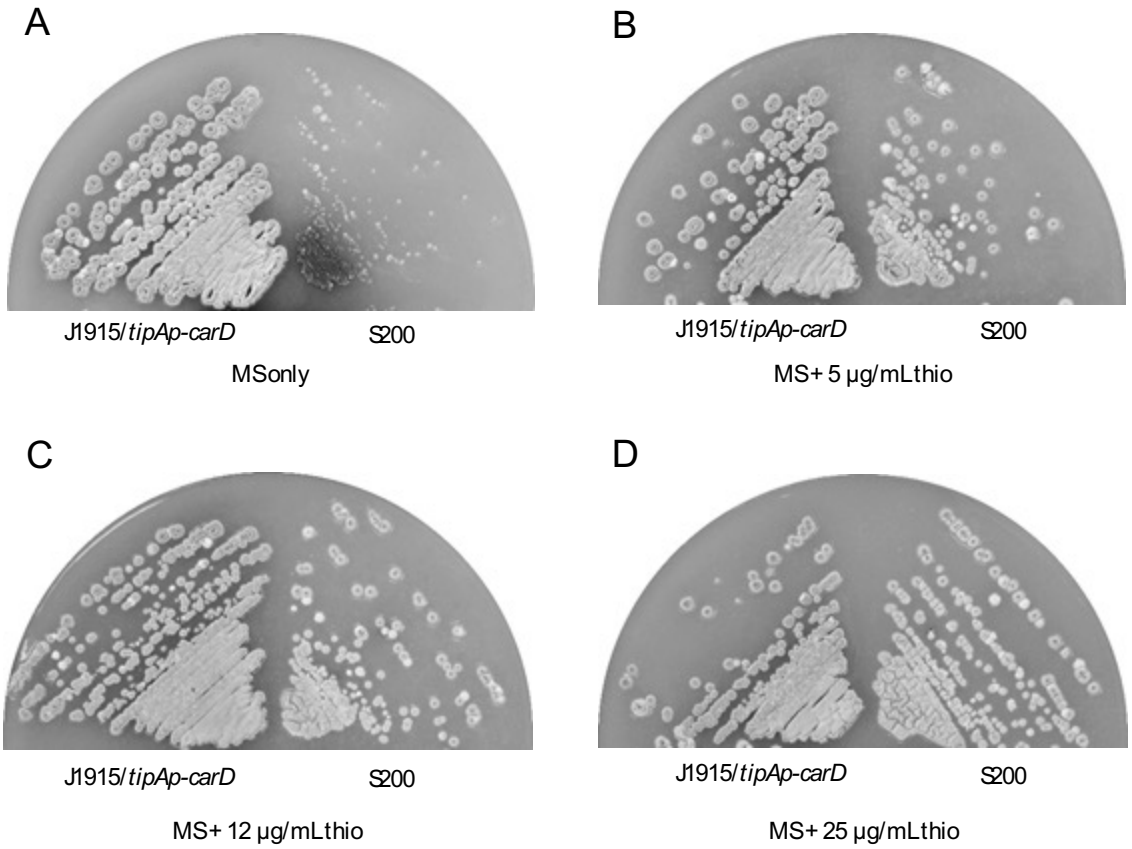


Figure 4.7 - **CarD is required for growth in *S. coelicolor*.** J1915/tipAp-*carD* revertant and S200 (J1915Δ*carD::hyg/tipAp-carD)* streaked out onto (A) MS only (B) MS + thiostrepton (5 µg/mL) (C) MS + thiostrepton (12 µg/mL) (D) MS + thiostrepton (25 µg/mL) and incubated for 5 days at 30 °C.

4.3.3 Full complementation of S200 with an additional copy of *carD*

To confirm that the mutant phenotype observed when S200 was grown in the absence (or indeed presence) of thiostrepton were both due to a lack of CarD and not unwanted polar effects caused during creation of the strain, it was deemed necessary to complement the strain with an additional copy of CarD under the control of its own promoter. The *carD* gene with 400 bp of upstream and downstream DNA was PCR amplified from M145 genomic DNA using the primer set ext_carD_F/ext_carD_R. The PCR product was cloned into EcoRV-cut pBlueScript SKII+, sequenced to confirm amplification with no mutations and subcloned into pRT802 as a BglII/BglII fragment, producing the plasmid pRT802::*carD*. The recombinant plasmid was used to transform *E. coli* ET12567/pR9406 and transferred via conjugation into S200.

As an additional control, a frameshift mutation was introduced into the *carD* ORF to confirm that the complementation of S200 with pRT802::*carD* was exclusively due to the *carD* ORF and not any other elements in the fragment. The plasmid pBlueScript-*carD* was digested with BstBI, "filled-in" using Klenow fragment and religated together. This procedure introduced a 2 bp insertion 212 bp into the ORF causing a frameshift mutation, mistranslation of the C-terminus of the protein and a premature stop codon. The insert was re-sequenced to confirm the presence of the frameshift mutation, subcloned into pRT802 as a BglII/BglII fragment, producing the plasmid pRT802::*carD*^{FS}. This plasmid was transferred via conjugation into S200.

S200/pRT802, S200/pRT802::*carD* and S200/pRT802::*carD*^{FS} were streaked out onto MS only and MS containing 15 µg/mL thiostrepton and incubated for 5 days at 30 °C. In both the presence and absence of thiostrepton, S200/pRT802::*carD* was fully complemented, including the ability to produce dark grey spores (Figure 4.8A–B). S200/pRT802 and S200/pRT802::*carD*^{FS} appeared to be unaffected by the presence of the empty pRT802 and plasmid containing frameshifted *carD*, respectively, with both strains resembling S200.

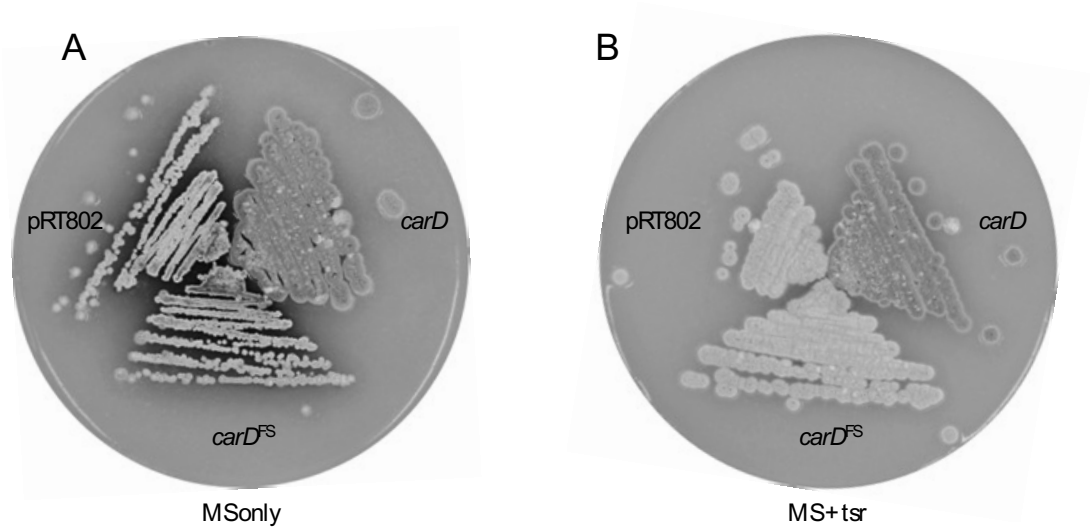


Figure 4.8 - ***carD* fully complements a conditional *carD* deletion mutant.** S200/pRT802, S200/pRT802::*carD* and S200/pRT802::*carD*^{FS} on (A) MS agar and (B) MS agar + 15 µg/mL thiostrepton and incubated for 5 days at 30 °C.

4.3.4 Depletion of *rbpA* and *carD* in an *sspB*-dependent manner

As an alternative method for investigating the essential nature of CarD, a method for targeted proteolysis previously used in a variety of species including *E. coli* (McGinness et al., 2006), *B. subtilis* (Griffith and Grossman, 2008) and *M. tuberculosis* (Kim et al., 2011) was adapted for use in *S. coelicolor*. Specific protein degradation can be achieved through a number of different mechanisms including C- or N-terminal tags that are specifically recognised by proteolytic enzymes. Trans-translation is a natural system that exists for degradation of incomplete polypeptides that result from the stalling of ribosomes during translation (Keiler, 2008). tmRNA (or SsrA) is a small RNA with both mRNA and tRNA properties. It initiates the trans-translation process by binding to the unoccupied A-site of a stalled ribosome, causing the ribosome to translocate to the mRNA portion of the tmRNA leading to the addition of a short peptide tag. This both rescues the stalled ribosome complex and targets the nascent polypeptide for degradation (Gottesman et al., 1998; Choy et al., 2007; Keiler, 2008). The degradation tag is then recognised directly by a protease or protease adapter for proteolysis. The adaptability of this system for use as a tool comes from the variability of each tmRNA system between bacterial species. SsrA-tagged proteins are degraded by a number of proteases including ClpX, which is specifically enhanced by the adapter protein SspB (Dougan et al., 2003). By modifying the last 3 residues of the SsrA-tag, non-SspB-specific degradation can be abolished making ClpX-mediated degradation fully dependent on the adapter protein

SspB (McGinness et al., 2006; Griffith and Grossman, 2008). In *M. tuberculosis* and *M. smegmatis,* this was achieved by changing the *E. coli* SsrA tag with the amino acid sequence "AANDENYALAA" to "AANDENYSENYADAS" (named DAS+4) (Kim et al., 2011). The SspB protein was placed under control of a TetR-regulated promoter allowing the degradation of the DAS+4 tagged protein to be induced by addition of anhydrotetracycline.

To initially develop the system for *S. coelicolor*, it was decided to target RbpA, which is required for rapid growth. An *rbpA*-DAS+4 fragment was synthesised (Eurofins Genomics, see appendix) and cloned into the plasmid pRT802 as a SpeI/BamHI fragment, creating the plasmid pRT802::*rbpA*-DAS+4. The synthesised fragment was designed with HindIII restriction sites upstream of the *rbpA* promoter and at the intersection between the *rbpA* ORF and the in-frame C-terminal DAS+4 tag. To ensure that the DAS+4 tag had no effect on the physiological role of the protein in the absence of SspB, pRT802::*rbpA*-DAS+4 was transferred via conjugation into the Δ*rbpA* mutant S101 and the ability of the tagged protein to complement the mutant was compared to the positive and negative controls pRT802::*rbpA* and pRT802, respectively.

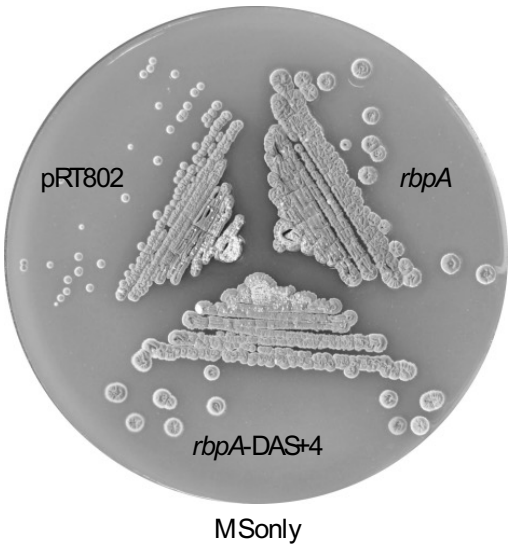


Figure 4.9 - ***rbpA*-DAS+4 complements an Δ*rbpA* deletion mutant.** S101/pRT802, S101/pRT802::*rbpA* and S101/pRT802::*rbpA*-DAS+4 streaked onto MS agar and incubated for 5 days at 30 °C.

The plasmid pRT802::*rbpA*-DAS+4 fully complemented the small colony mutant phenotype (observed in negative control S101/pRT802) and was indistinguishable from the strain complemented with wild-type *rbpA* (S101/pRT802::*rbpA*). This confirms that RbpA-DAS+4 was fully functional and suitable for use in depletion experiments.

To ensure efficient expression and to remove some common restriction sites, *sspB* was synthesised using the *S. coelicolor* codon usage table (Eurofins Genomics, see appendix). The gene was subcloned into the plasmid pIJ6902 as an NdeI/BamHI fragment, placing the gene under the control of the thiostrepton-inducible *tipAp* promoter and creating the plasmid pIJ6902::*sspB*. Once confirmed by restriction analysis, pIJ6902::*sspB* and the empty vector pIJ6902 were transformed into ET12567/pUZ8002 and transferred via conjugation into S101/pRT802::*rbpA*-DAS+4 and the negative control strain S101/pRT802::*rbpA*.

The four strains were streaked to single colonies on MS agar and MS agar containing a range of thiostrepton concentrations. In the absence of thiostrepton, all four strains showed no comparable difference and resembled a wild type phenotype. This shows that in the absence of thiostrepton neither the DAS+4 tag nor plasmids pIJ6902 and pIJ6902::sspB had any effect on growth. At thiostrepton concentrations of 5 µg/mL to 25 µg/mL all three control strains were unaffected. S101/pRT802::*rbpA*/pIJ6902::*sspB* shows that expression of *sspB* from the strong *tipAp* promoter has no effect on growth in an otherwise wildtype background. S101/pRT802::*rbpA*-DAS+4/pIJ6902, shows that the DAS+4 tag fully complements the Δ*rbpA* null strain, however isn't effected by thiostrepton in the absence of SspB protein. The third control strain S101/pRT802::*rbpA*/pIJ6902 shows that increasing concentrations of thiostrepton have no effect on growth in the presence of the plasmid pIJ6902 which possesses the *tsr* thiostrepton resistance gene.

For the depletion strain S101/pRT802::*rbpA*-DAS+4/pIJ6902, an Δ*rbpA* mutant small colony phenotype can be seen at concentrations as low as 5 µg/mL. As the concentration is increased to 12 µg/mL and 25 µg/mL the effect is further pronounced. The ability to replicate a clear mutant phenotype in an inducible manner suggests that proteins can be actively depleted by proteolysis using the ClpXP system in *S. coelicolor.*

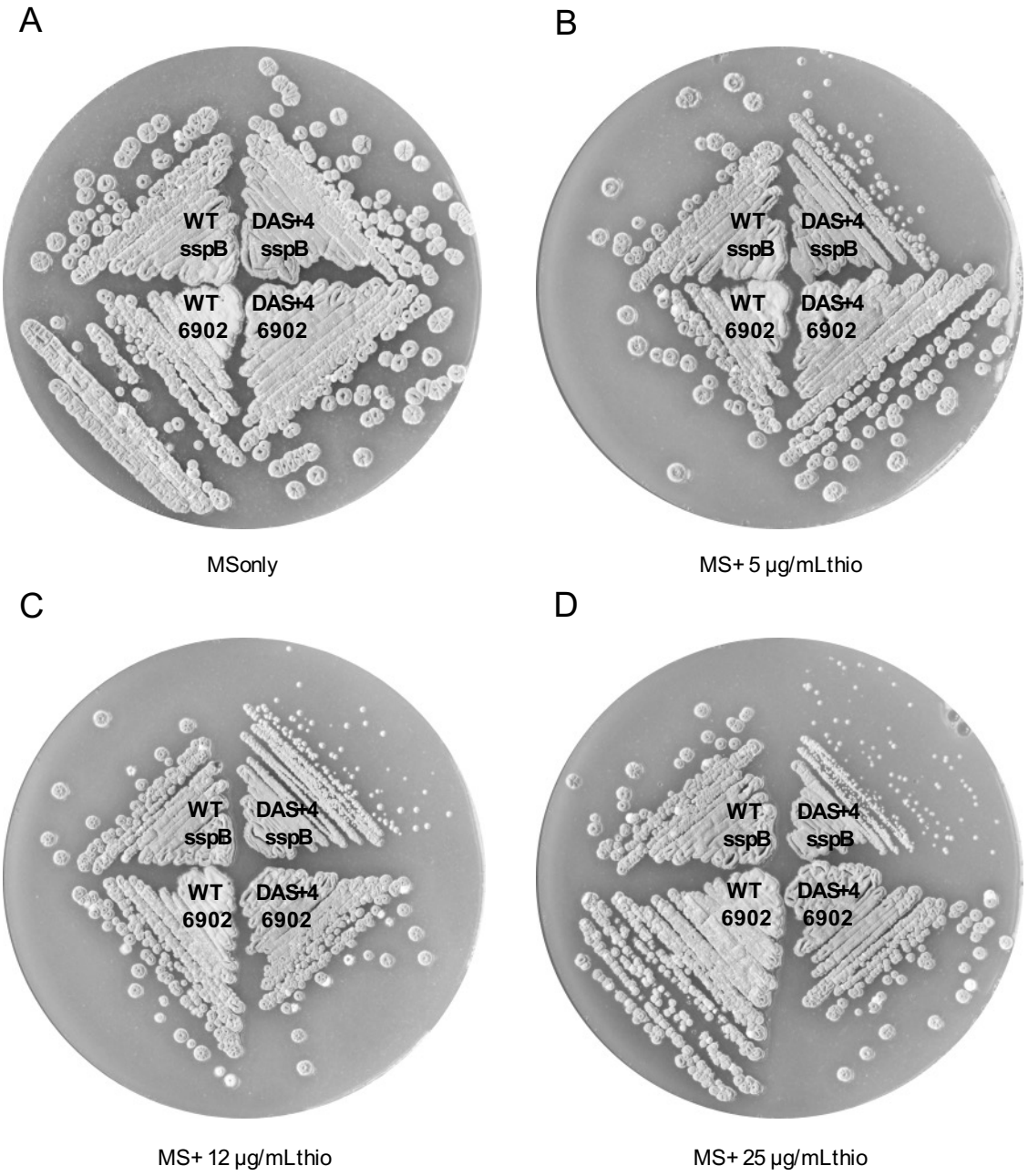


Figure 4.10 - **SspB-mediated depletion of RbpA-DAS+4.** S101/pRT802::*rbpA*/pIJ6902::*sspB* (WT sspB), S101/pRT802::*rbpA*-DAS+4/pIJ6902::*sspB* (DAS+4 sspB), S101/pRT802::*rbpA*/pIJ6902 (WT 6902) and S101/pRT802::*rbpA*-DAS+4/pIJ6902 (DAS+4 6902) streaked out onto (A) MS agar, (B) MS agar + 5 µg/mL thiostrepton, (C) MS agar + 12 µg/mL thiostrepton and (D) MS agar + 25 µg/mL thiostrepton and incubated for 5 days at 30 °C.

Following the successful depletion of RbpA, the system was used to deplete CarD. To achieve this, *carD* (with its own promoter and no stop codon) was PCR amplified from genomic DNA using the primers carD_3XFLAG_F/carD_3XFLAG_R, which introduced flanking HindIII sites, then subcloned into HindIII-cut pRT802::*rbpA*-DAS+4, replacing the *rbpA* gene with *carD* to create pRT802::*carD*-DAS+4.

To ensure that in the absence of SspB the DAS+4 tag had no effect on the physiological role of CarD, the plasmid pRT802::*carD*-DAS+4 was transferred via conjugation into the conditional Δ*carD* mutant S200 and the ability of the tagged protein to complement the mutant was compared to the positive and negative controls pRT802::*carD* and pRT802, respectively.

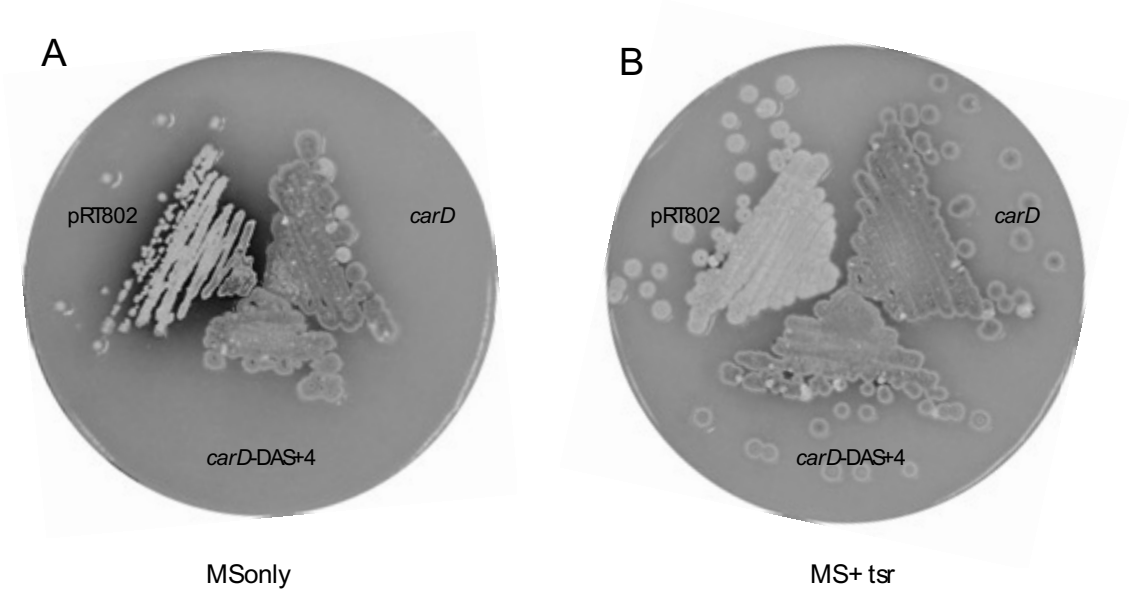


Figure 4.11 - ***carD*-DAS+4 complements S200, a Δ*carD* deletion mutant.** S200/pRT802, S200/pRT802::*carD* and S200/pRT802::*carD*-DAS+4 streaked onto (A) MS agar and (B) MS agar + 15 µg/mL thiostrepton and incubated for 5 days at 30 °C.

The plasmid pRT802::*carD*-DAS+4 fully complemented the mutant phenotype both in the absence (Figure 4.11A) and presence of thiostrepton (Figure 4.11B) and was indistinguishable from the strain complemented with wild-type *carD.* This confirms that CarD-DAS+4 was fully functional and suitable for use in depletion experiments.

To deplete CarD in an SspB-dependent manner, a strain possessing only one chromosomal copy of *carD* with a DAS+4 tag was required. Due to the essential nature of CarD it was not possible to first create a Δ*carD* mutant and complement this with

carD-DAS+4. Instead, the plasmid pRT802::*carD*-DAS+4 was transferred via conjugation into a single crossover strain J1915/pIJ6650*ΔcarD::hyg*. As the strain now possessed an additional functional copy of *carD* (*carD*-DAS+4 on pRT802 integrated at the phage ϕBT1 site), following a round of non-selective growth it was now possible to obtain Apr^S Hyg^R double crossover deletion mutants representing the strain J1915Δ*carD::hyg/pRT802::carD*-DAS+4. This strain was named S201. pIJ6902::*sspB* was transferred via conjugation into this strain and J1915 (as a negative control with untagged, native CarD) using apramycin resistance as selection.

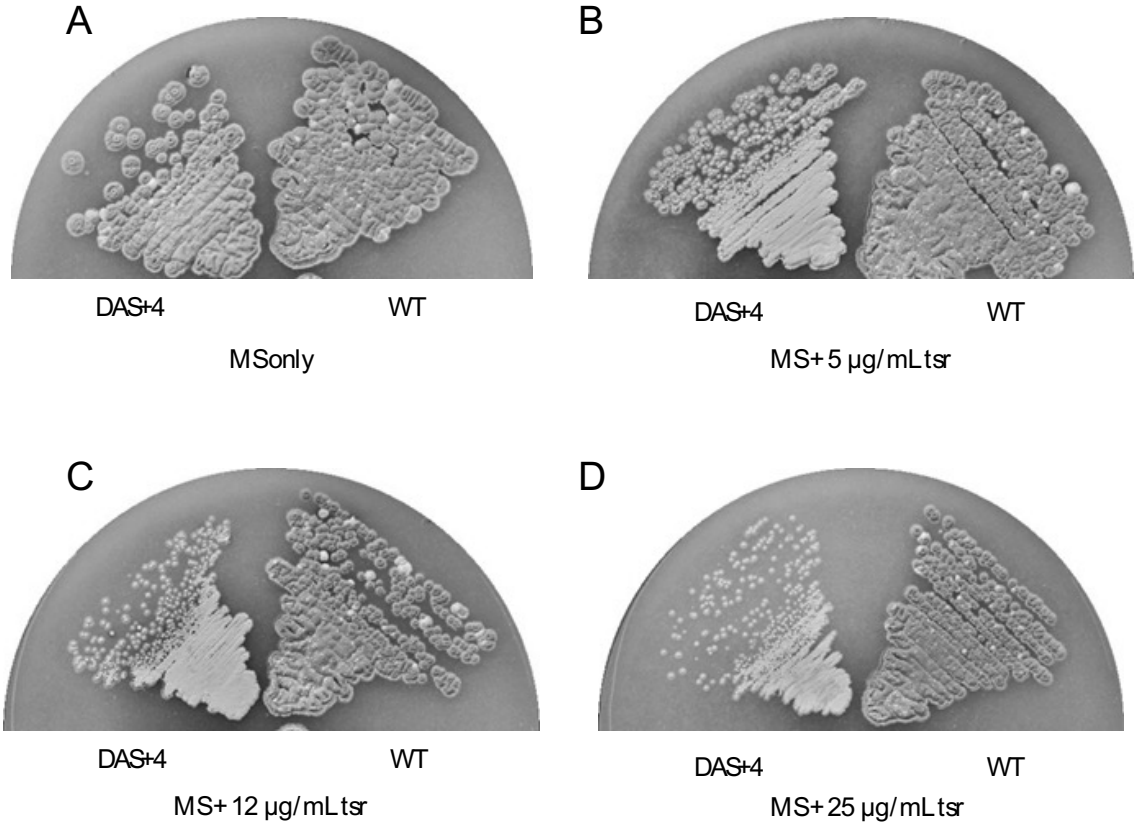


Figure 4.12 - **SspB-mediated depletion of CarD.** S201/pIJ6902::*sspB* and J1915/pIJ6902::*sspB* streaked out onto (A) MS agar, (B) MS agar + 5 µg/mL thiostrepton, (C) MS agar + 12 µg/mL thiostrepton and (D) MS agar + 25 µg/mL thiostrepton and incubated for 5 days at 30 °C.

In the absence of thiostrepton on MS agar, S201/pIJ6902::*sspB* and the control strain J1915/pIJ6902::*sspB* showed no comparable difference (Figure 4.12A). At thiostrepton concentrations of 5 µg/mL to 25 µg/mL, the control strain J1915/pIJ6902::*sspB* was unchanged indicating that presence of thiostrepton and expression of *sspB* have no effect

on growth. At concentrations of 5 µg/mL, below the typical working concentration for thiostrepton in *S. coelicolor* (Huang et al., 2005), a partial CarD null phenotype can be observed (Figure 4.12B). The colonies are smaller and lighter in colour, reminiscent of a Whi phenotype for mutants blocked in late stages of sporulation (Flärdh and Buttner, 2009). This phenotype is further pronounced at a thiostrepton concentrations of 12 µg/mL to 25 µg/mL (Figure 4.12C and Figure 4.12D).

The depletion phenotype observed in this experiment and strain S200 grown in the absence of thiostrepton are similar, further proving that CarD is a protein critical for growth in *S. coelicolor*. These data also demonstrate that SspB-mediated depletion can be effectively used for observing mutant phenotypes in *S. coelicolor*.

4.4 Purification of CarD

To further understand the biochemical role of CarD, it was necessary to perform *in vitro* experiments with purified CarD protein. This section details the purification of CarD heterologously expressed in *E. coli,* chosen as the most suitable host for obtaining sufficient protein yields and purity.

4.4.1 Overexpression of His-CarD protein in *E. coli*

To overexpress and purify CarD protein from *E. coli,* the ORF was amplified from M145 genomic DNA by PCR using primers CarD_D_F/CarD_O_R, designed to introduce an NdeI site overlapping the start codon and a BamHI site downstream from the stop codon, respectively. The PCR product was cloned into EcoRV-cut pBlueScript SKII+, sequenced to confirm amplification with no mutations, and subcloned into pET15b as an NdeI/BamHI fragment, producing the plasmid pET15b-*carD*. The recombinant plasmid was used to transform *E. coli* BL21 (pLysS) and a single colony was used to inoculate 250 mL LB. The culture was grown in an orbital shaker at 37 °C to OD_{600} 0.5–0.6, then the flask was submerged in ice-water for 10 min before addition of 1 mM IPTG and continued incubation in the orbital shaker for 3 h at 30 °C. Cells were harvested by centrifugation, resuspended in 15 mL Ni-NTA binding buffer (containing 1.5 mL protease inhibitor cocktail and 25 µg/mL PMSF) and disrupted by sonication (6 x 10s @ 35% ampl.) (whole cell lysate, WCL).

4.4.2 Purification of His-CarD

The cleared cell lysate (CCL) was separated from the cell debris by centrifugation. To purify His-CarD, the CCL was first loaded onto a Ni-NTA sepharose column and bound His-CarD was eluted with Ni-NTA elution buffer. Eluted protein was filtered through a 0.22 µM filter and further FPLC purified by passing through a size exclusion column (HiLoad™ 16/60 Superdex™ 200 column) with gel filtration (GF) buffer (Figure 4.13B). Elutions were collected as 1 mL fractions and those containing His-CarD (fractions G8–H5) were combined, concentrated to 15 µM by centrifugal filtration (VivaSpin 6, 3000 MWCO, Sartorius), flash frozen in liquid nitrogen and stored at -80 °C. SDS-PAGE analysis of the purification and elution fractions are shown in Figure 4.13A. Purification via this method achieved a yield of approximately 1.04 mg of protein from 250 mL culture.

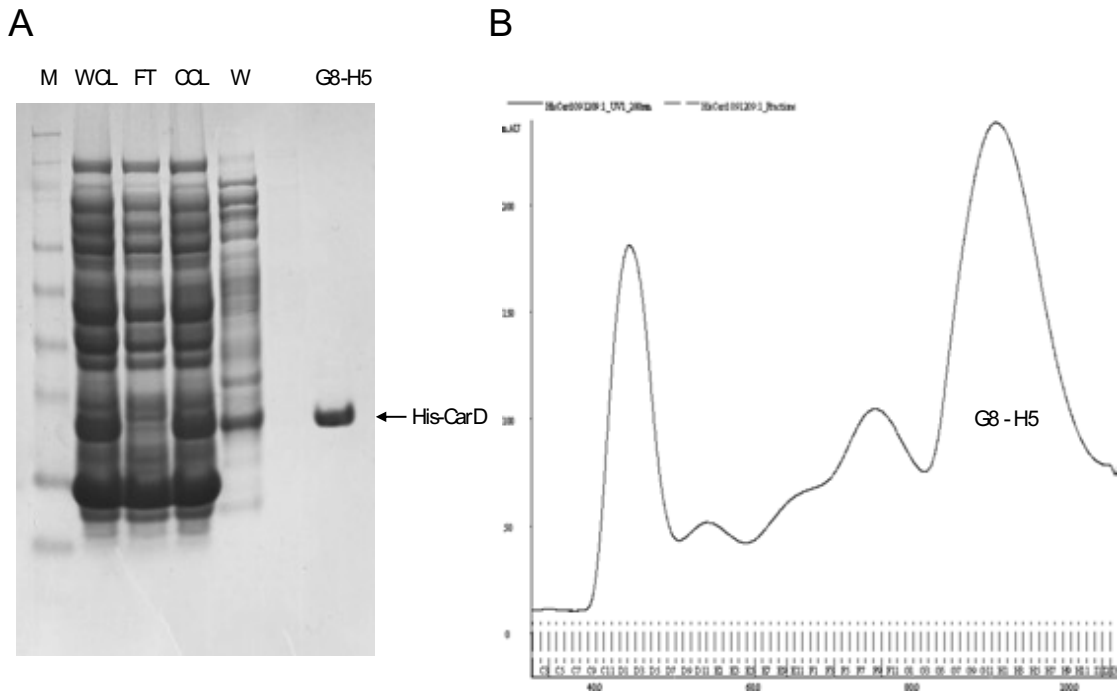


Figure 4.13 - **Purification of His-CarD from *E. coli*** (A) SDS-PAGE analysis of His-CarD purification. Lane 1 (M): SeeBlue Plus2 marker; lane 2 (WCL): whole cell lysate; lane 3 (FT): flow through; lane 4 (CCL): cleared cell lysate, lane 5 (W): wash, lane 6 (G8–H5): size exclusion elution combined fractions G8–H5. (B) Size exclusion purification of His-CarD. HiLoad™ 16/60 Superdex™ 200 column elution trace for purification of His-CarD

4.5 CarD activates transcription from σ^{HrdB}-dependent promoters *in vitro*

Initial published data suggested *M. tuberculosis* CarD as an important transcriptional regulator. It was shown to be an essential protein, its depletion effected rRNA transcription and revealingly, possessed an N-terminal RNAP binding domain homologous to the transcription repair coupling factor (TRCF) that binds to β (Stallings et al., 2009). The role of *M. tuberculosis* CarD in transcription activation was later demonstrated through *in vitro* transcription experiments. The following section details the purification of *S. coelicolor* RNAP and subsequent *in vitro* transcription experiments to investigate the role of *S. coelicolor* CarD in transcription.

4.5.1 Purification of S129 RNA polymerase holoenzyme and core

To purify RNA polymerase holoenzyme in the absence of RbpA, the strain S129 (M145 Δ*rpoC*::*rpoC*^{HIS} Δ*rbpA*::*aac*(3)IV) (Newell et al., 2006) was used. Four fresh spore plates were combined and used to inoculate 6 x 500 mL YEME cultures in 2 L baffled flasks. Cultures were grown at 30 °C to an OD_{450} 1.5–2 before collecting the mycelium by centrifugation at 6,000 rpm for 2 min at 4 °C. Pellets were washed twice with cell wash buffer and centrifuged again at 6,000 rpm for 2 min at 4 °C before snap-freezing in liquid nitrogen and storage at -80 °C until further purification. The following method was developed to improve the yield and purity of *S. coelicolor* RNA polymerase and combines previous published methodology with minor modifications. The initial precipitation of DNA along with associated RNA polymerase by PEG is based on the method described by Gross et al. (1976) with minor adjustments (Fong et al., 2010). The subsequent Ni-affinity purification precedes purification of core RNAP by MonoQ ion exchange chromatography, essentially as described by Hahn et al. (2003).

Pellets were cryogenically ground for 2 x 90 s, resuspended in 20 mL cell wash buffer (containing 1.5 mL protease inhibitor cocktail and 25 µg/mL PMSF) and the cells were disrupted by sonication for 6 x 10 s (35% ampl.) on ice. The cell lysate was centrifuged at 16,000 x g for 30 min at 4 °C and the cleared cell lysate (or supernatant) was transferred to a new tube. 12 mL of PEG-PPT buffer was added and mixed by rotation for 10 min at 4 °C, allowing precipitation of DNA (and bound RNAP) to occur. The mixture was centrifuged at 7,000 x g for 10 min at 4 °C and the supernatant discarded (S1). The pellet

was resuspended in 2 mL PEG-HS buffer to dissociate RNAP, transferred to a 2 mL microcentrifuge tube and centrifuged at 10,000 x g for 10 min at 4 °C. The supernatant (S2) was collected and 6 mL of PEG-Ni binding buffer was added – for six separate pellets this totals 48 mL that was combined for Ni-NTA purification. A column was constructed with a 4 mL IDA-sepharose bed volume and following washes with 4 column volumes of PEG-Ni binding buffer and 3 column volumes of wash buffer the protein was eluted in 6 mL 0.5 M NaCl elution buffer (Figure 4.14A). The protein was diluted with 6 mL dilution buffer (to lower the NaCl concentration to 0.2 M), centrifuged at 7,000 x g for 10 min before loading onto a Mono Q 5/50 GL column. This final anion-exchange purification step was designed to separate RNAP holoenzyme from RNAP core. The elution gradient was set 0% to 40% for 40 mL followed by 40% to 100% for 5 mL and UV trace peaks were analysed by western blot (Figure 4.14B and Figure 4.14C). By western blotting for σ^{HrdB} were we able to identify the fractions without the sigma factor present. Similar fractions were combined and, following final confirmation by western blot (Figure 4.14D), were designated holoenzyme (fractions D6–D9), "depleted" (fractions D11–D12) and core (fractions E1–E2).

To confirm that functional core RNAP was purified, *in vitro* transcription reactions were performed in the presence and absence of additional sigma factors. σ^{HrdB} and σ^{R} purified from *E. coli* (gifts from P. Doughty) were added to *in vitro* transcription reactions performed on *atplp* and *trxCp,* respectively. Transcripts from both promoters were only observed in the presence of the appropriate sigma factor indicating that functional core RNAP had been purified and could be combined with different sigma factors for a functional holoenzyme (Figure 4.14E).

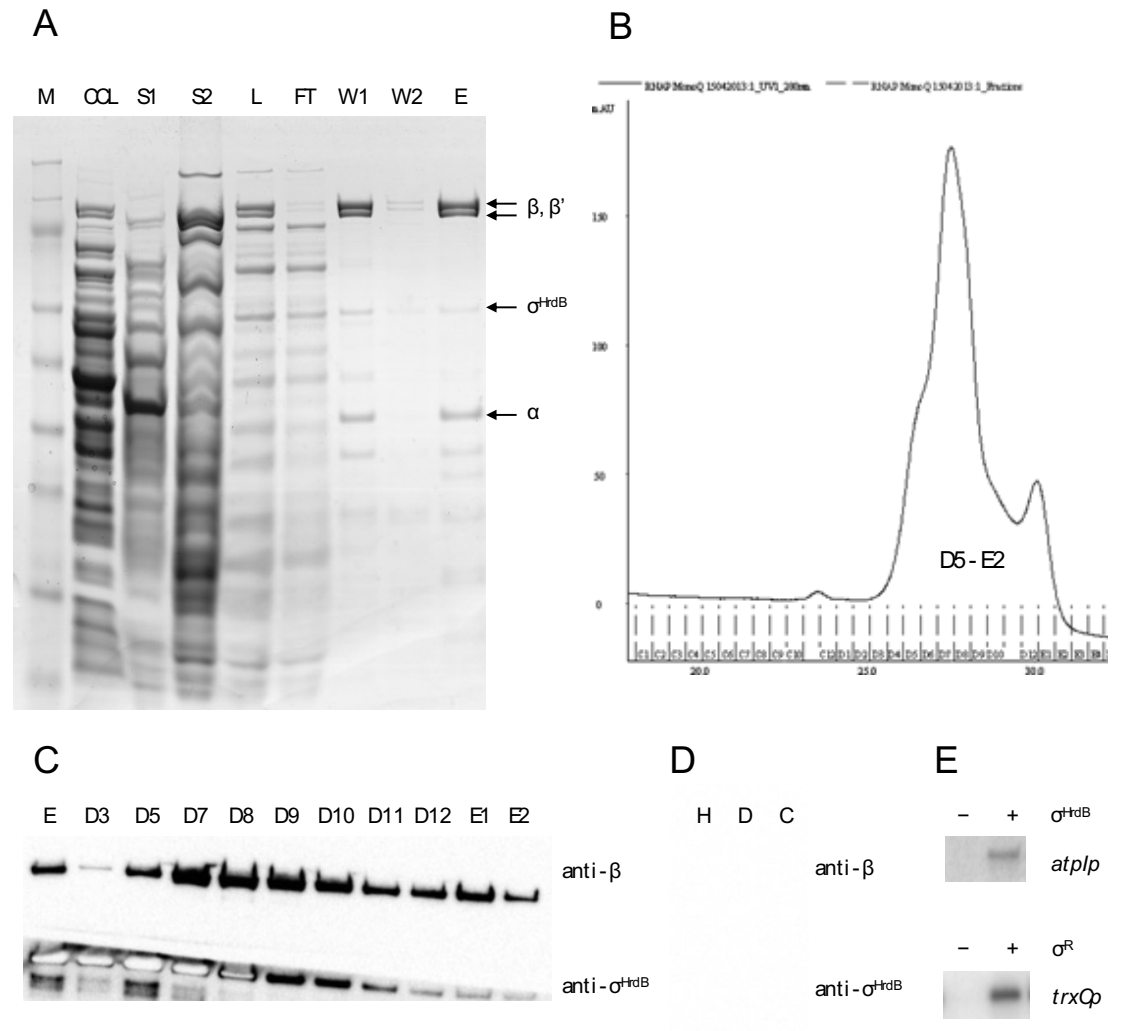


Figure 4.14 - **Purification of S129 RNA polymerase holoenzyme and core.** (A) SDS-PAGE analysis of S129 RNAP purification. Lane 1 (M): SeeBlue Plus2 marker; lane 2 (CCL): cleared cell lysate; lane 3 (S1): PEG-PPT supernatant; lane 4 (S2): PEG-HS supernatant ; lane 5 (L): Ni-NTA load; lane 6 (FT): Ni-NTA loading flowthrough; lane 7 (W1): Ni-NTA wash 1; lane 8 (W2): Ni-NTA wash; lane 9 (E): Ni-NTA elution (500 mM imidazole). (B) Anion-exchange purification of RNAP. Mono Q 5/50 GL column elution trace for purification of RNAP (C) Western blot analysis of anion-exchange purification of S129 RNAP. Lane 1 (E): Ni-NTA elution; lanes 2–10: Mono Q 5/50 GL elution fractions (D) Western blot analysis of final RNAP stocks. Lane 1 (H): RNAP holoenzyme (fractions D6–D9); lane 2 (D): "depleted" RNAP (fractions D11–D12); lane 3 (C): RNAP core (fractions E1–E2). (E) *In vitro* transcription testing of core RNAP. Reactions performed on *atpIp* and *trxCp* promoters with core RNAP and additional σ^{HrdB} and σ^{R}, respectively.

4.5.2 CarD activates transcription from σ^{HrdB}-dependent promoters *in vitro*

M. tuberculosis CarD has been shown to activate transcription from σ^{A}-dependent promoters *in vitro* (Srivastava et al., 2013). To investigate if *S. coelicolor* CarD had a similar effect, multi-round *in vitro* transcription reactions were performed with purified RNAP core and σ^{HrdB} on the σ^{HrdB}-dependent *atpIp* promoter. An *atpIp* fragment was amplified from M145 gDNA using the primers atpIp_F/atpIp_R and purified using the Promega Wizard® SV Gel and PCR Clean-Up System.

The effect of CarD was observed by titrating increasing concentrations into *in vitro* transcription reactions. Addition of CarD activated transcription from the *atpIp* promoter at CarD:RNAP ratios of 2:1 (100 nM) to 10:1 (500 nM) (Figure 4.15A). At a CarD:RNAP ratio of 20:1 (1 µM), a slight decrease in transcription was observed, possibly due to non-specific effects. Consequently for all further experiments a CarD:RNAP ratio of 10:1 was used (500 nM).

Additional *in vitro* transcription reactions were performed on a selection of linear templates containing σ^{HrdB}-dependent promoters. The *actII-ORF4p* promoter was amplified with the primers actIIORF4p_F/actIIORF4p_R, the *rplJp* promoter was amplified with the primers rplJp_F/rplJp_R and the *rrnDp3* and *rrnDp4* promoters were amplified with the primers rrnDp1-4_F/rrnDp1-4_R. In each case, addition of excess CarD (500 nM) stimulated transcription (Figure 4.15B).

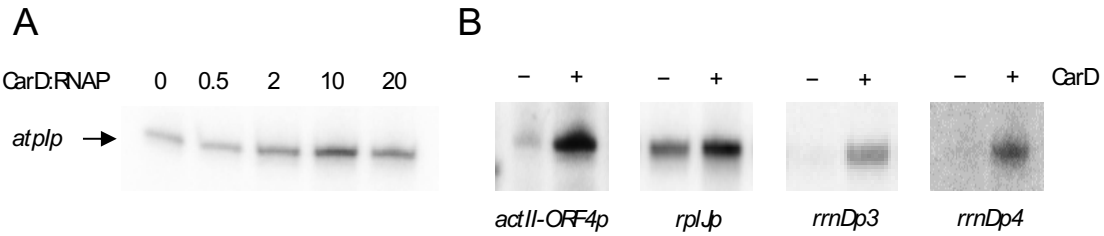


Figure 4.15 - **CarD activates transcription from σ^{HrdB}-dependent promoters.** (A) Multi-round i*n vitro* transcription reactions performed on the *atpIp* promoter template (5 nM) with core RNAP (50 nM), σ^{HrdB} (250 nM) and CarD at CarD:RNAP ratios of 0, 0.5 (25 nM), 2 (100 nM), 10 (500 nM) and 20 (1 µM). (B) Multi-round i*n vitro* transcription reactions performed on the *actII-ORF4p*, *rplJp*, *rrnDp3* and *rrnDp4* promoter templates (5 nM) with core RNAP (50 nM) and σ^{HrdB} (250 nM) in the presence and absence of excess CarD (500 nM).

4.6 Activation of transcription by CarD and RbpA is dependent on the absence of a conserved σ^{HrdB} -35 promoter region

4.6.1 CarD and RbpA selectively activate σ^{HrdB}-dependent *rrnD* promoters

The results showing CarD is a general activator of transcription from σ^{HrdB}-dependent promoters *in vitro* are analogous to the findings for RbpA in *S. coelicolor* (Newell et al., 2006; Tabib-Salazar et al., 2013). It was therefore considered that RbpA and CarD might have overlapping roles in transcription initiation.

To investigate the role of CarD and RbpA in transcription of rRNA, *in vitro* transcription experiments were performed with the full *rrnDp1–4* promoter region, PCR amplified from M145 gDNA using the primers rrnDp1-4_F/rrnDp1-4_R. Four promoters and their transcription start sites were previously identified through high resolution S1 nuclease protection mapping the region upstream of the *rrnD* gene set (Baylis and Bibb, 1988). Using this data we were able to predict four transcription products from our *in vitro* transcription template 115 nt, 195 nt, 341 nt and 392 nt in size for the *p4, p3, p2* and *p1* promoters, respectively.

Under all tested conditions no transcription was observed from the *rrnDp1* promoter. The *rrnDp2* promoter was active in the absence of both RbpA and CarD and interestingly, unlike the σ^{HrdB} promoters tested previously, addition of either CarD or RbpA had no effect on transcriptional output (Figure 4.16A). *rrnDp3* and *rrnDp4* showed very little activity in the absence of CarD and RbpA but both were activated by addition of either protein.

To confirm these data and to ensure that the *rrnD* promoters did not influence each other, the *p2* and *p3* promoters were further investigated in isolation. To construct synthetic promoters, complementary oligonucleotides up to 60 bp in size representing the template and non-template strands and designed to incorporate BamHI and HindIII sites at the 5' and 3' ends respectively, were annealed and cloned into BamHI/HindIII-cut pBlueScript SKII+. Using these plasmids as a template, a linear *in vitro* transcription template could be produced by PCR amplification using M13 universal primers. Following this method, *rrnDp2* and *p3* promoter fragments were constructed using the oligonucleotide pairs rrnDp2_nt/rrnDp2_t and rrnDp3_nt/rrnDp3_t, respectively.

The *in vitro* transcription results were as observed from the single *rrnDp1–4* fragment; the *rrnDp2* promoter was active in the absence of both RbpA and CarD and addition of either or both proteins had little effect on transcriptional output (Figure 4.16B). The *rrnDp3* promoter had little activity in the absence of RbpA and CarD but transcription occurred upon addition of either protein. Addition of both CarD and RbpA at saturating concentrations (RbpA/CarD:RNAP molar ratio of 10:1) had an additive effect on transcriptional output.

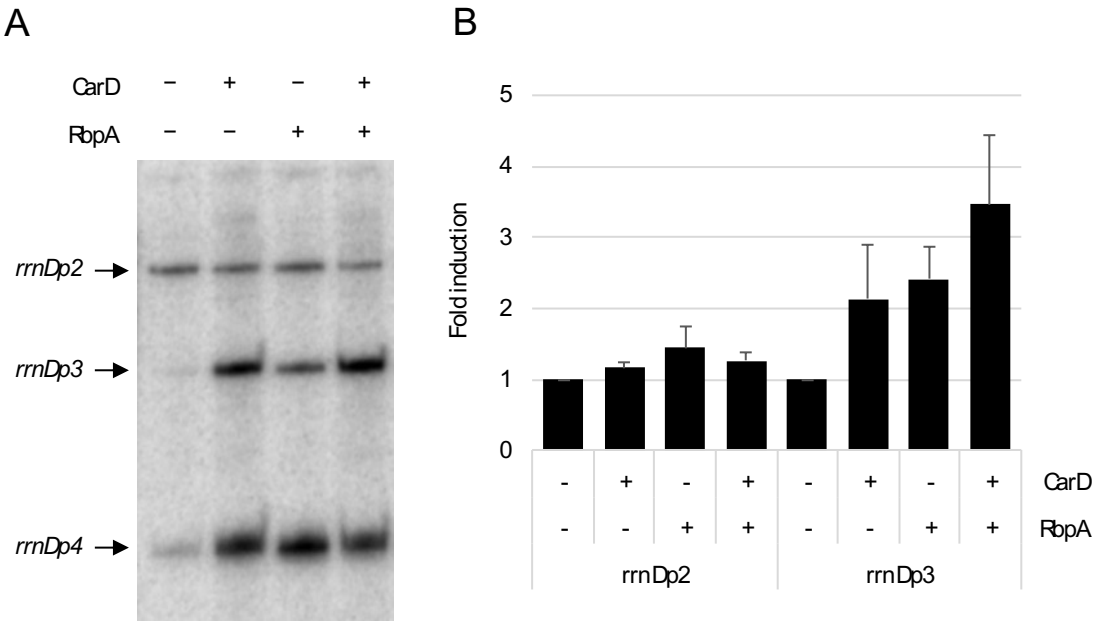


Figure 4.16 - ***In vitro* transcription analysis of *rrnDp1–4* promoter region** (A) Multi-round i*n vitro* transcription reactions performed on the whole *rrnDp1–4* promoter region (5 nM) with core RNAP (50 nM), σ^{HrdB} (250 nM) in the presence and absence of excess CarD (500 nM) and RbpA (100 nM). (B) Multi-round i*n vitro* transcription reactions performed on the *rrnDp2* and *rrnDp3* promoter templates (5 nM) with core RNAP (50 nM), σ^{HrdB} (250 nM) in the presence and absence of excess CarD (500 nM) and RbpA (500 nM). Data are presented as fold-difference relative to reactions lacking both CarD and RbpA. Transcript levels were quantified by phosphorimaging from triplicate data, and standard deviation is indicated.

4.6.2 Changes to the σ^{HrdB} -35 DNA element influence promoter sensitivity to CarD and RbpA

Hahn and Roe (2007) previously aligned the *rrnDp1–4* promoter set, and identified -35 DNA elements that closely matched the consensus for the p1 and p2 promoters, but failed to identify -35 elements for the p3 and p4 promoters (Figure 4.17). The spacing and -10 region was largely similar in all four promoters.

rrnD p1	TGA**TGTGCA**TCCACCCCTGCGAGCTGC**TAGTGT**CCTCTT**C**
rrnD p2	GCCG**TTGACA**CGGAGCGAGCGGGGAGG**TAGATT**CGAAC**A**G
rrnD p3	CCGGTAAGACCGGCTCGAAAGTTCTGA**TAAAGT**CGGAGC**C**
rrnD p4	CGAGGAAATCGGATCGGAAAGATCTGA**TAGAGT**CGGAAA**C**
σ^{HrdB} consensus	**TTGaCA** (N17-18) **TAgaaT**

Figure 4.17 - **Comparison of the *rrnDp1–4* promoter sequences.** Putative -35, -10 sequences and transcription start sites are shown in bold. Identical sequences between the *rrnDp3* and *rrnDp4* promoters are underlined. (Adapted from Hahn and Roe, 2007).

To investigate if the -35 region was related to the differential promoter sensitivity to CarD and RbpA, synthetic promoters were designed initially swapping the conserved -35 region from the *rrnDp2* promoter with the poorly conserved -35 region from the *rrnDp3* promoter (Figure 4.18). Using the same method as used previously for producing the isolated *rrnDp2* and *p3* promoters, complementary oligonucleotides were used and the *in vitro* transcription templates were PCR amplified using M13 universal primers. The newly created promoters named *rrnDp3s* ("*rrnDp3*strong") and *rrnDp2w* ("*rrnDp2*weak") *were* created by annealing the oligonucleotide pairs rrnDp3str-35_nt/rrnDp3str-35_t and rrnDp2wk-35_nt/rrnDp2wk-35_t, respectively.

rrnD p2 GCCG**TTGACA**CGGAGCGAGCGGGGAGG**TAGATT**CGAAC**A**G

rrnD p2w GCCGTAAGACCGGAGCGAGCGGGGAGG**TAGATT**CGAAC**A**G

rrnD p3 CCGGTAAGACCGGCTCGAAAGTTCTGA**TAAAGT**CGGAGC**C**

rrnD p3s CCGG**TTGACA**CGGCTCGAAAGTTCTGA**TAAAGT**CGGAGC**C**

σ^{HrdB} consensus **TTGaCA** (N17-18) **TAgaaT**

Figure 4.18 - **Comparison of *rrnDp2, rrnDp3, rrnDp2w* and *rrnDp3s* promoter sequences.** Putative -35, -10 sequences and transcription start sites are shown in bold. Nucleotides conserved between the *rrnDp3* and *rrnDp4* promoters are underlined.

rrnDp2w, the p2 promoter without the conserved -35 region, was no longer active in the absence of CarD and RbpA (Figure 4.19). Addition of CarD and RbpA separately now activated transcription and addition of both proteins at saturating concentrations had an additive effect on transcriptional output. *rrnDp3s,* the p3 promoter with the conserved -35 region from the p2 promoter, was now active in the absence of CarD and RbpA. Addition of CarD or RbpA, two proteins previously identified as transcriptional activators, inhibited transcription and the addition of both proteins at saturating concentrations had a further inhibitory effect on transcription.

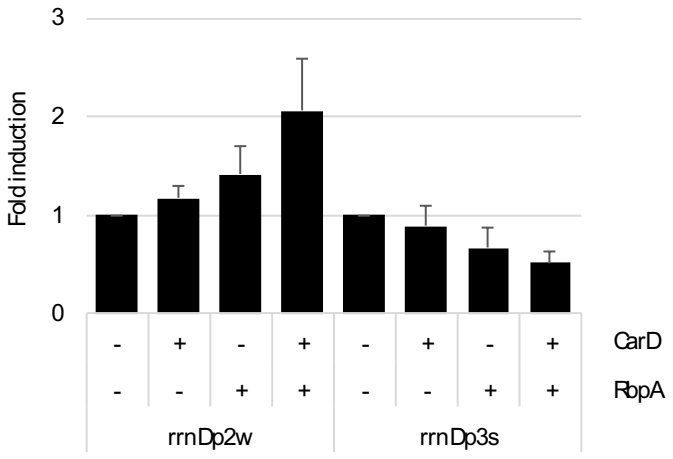


Figure 4.19 - **CarD and RbpA activate transcription from *rrnDp2w* but inhibit transcription from *rrnDp3s*** Multi-round i*n vitro* transcription reactions performed on the *rrnDp2w* and *rrnDp3s* promoter template (5 nM) with core RNAP (50 nM), σ^{HrdB} (250 nM) in the presence and absence of excess CarD (500 nM) and RbpA (500 nM). Data are presented as fold-difference relative to reactions lacking both CarD and RbpA. Transcript levels were quantified by phosphorimaging from triplicate data, and standard deviation is indicated.

4.7 Depletion of CarD or RbpA has no effect on *rrnDp1–4* transcript levels

In this study, two approaches have been used to study the growth of *S. coelicolor* in the absence of CarD and RbpA. Initially the strain S200 was created, a strain which possesses a single copy of *carD* under the control of a thiostrepton-inducible promoter. Growth of this strain in the absence of thiostrepton represents a Δ*carD* null mutant. We also implemented a ClpX-SspB system for targeted degradation of *carD* and *rbpA* tagged with a DAS+4 tag.

Having identified that CarD and RbpA are required for transcription of promoters such as *rrnDp3 in vitro*, we wanted to investigate whether this was the case *in vivo*. The ClpX-SspB system was decided most suitable for depletion studies on CarD and RbpA. As the system works through active targeted degradation, it was assumed that changes in protein concentration would occur more rapidly than exploiting an inducible promoter system which relies on preventing further gene transcription and protein turnover for any observed change in protein concentration. Additionally, the process is experimentally simpler; with the ClpX-SspB system, depletion is induced by addition of thiostrepton, rather than removal of the compound from culture medium.

To perform the depletions, cultures of S201/pIJ6902::*sspB*, S101/pRT802::*rbpA*-DAS+4/pIJ6902::*sspB* and negative controls S201/pIJ6902 and S101/pRT802::*rbpA*-DAS+4/pIJ6902 were grown in YEME-10 to an OD_{450} of 1. A 10 mL sample was taken from each (marked as timepoint 0) and thiostrepton was added to all cultures to a final concentration of 20 µg/mL. 10 mL samples were taken subsequently at 30, 60 and 120 minute intervals and total RNA was extracted from each sample. To analyse the transcriptional output from the *rrnDp1–4* promoters in response to depletion of CarD and RbpA, S1 nuclease protection mapping was performed on the RNA samples. The primers rrnD_S1_F/rrnD_S1_R were used to create a probe suitable for mapping this operon, with the reverse primer rrnD_S1_R labelled with ^{32}P radiolabel on the 5' end. Prior to S1 nuclease protection mapping, 30 µg RNA was hybridised to the *rrnDp1–4* probe following digestion, samples were run on an 8% denaturing urea polyacrylamide gel and visualised by phosphorimaging.

Following depletion of either CarD or RbpA no significant change in transcript level was observed by S1 nuclease protection mapping (Figure 4.20). The same was observed in both control strains lacking the SspB adapter protein required for depletion. The most distinct band was the *rrnDp3* transcript product. Despite the finding that the rrnDp3 promoter is dependent on CarD and RbpA *in vitro*, the transcript level was unaffected following depletion of either protein.

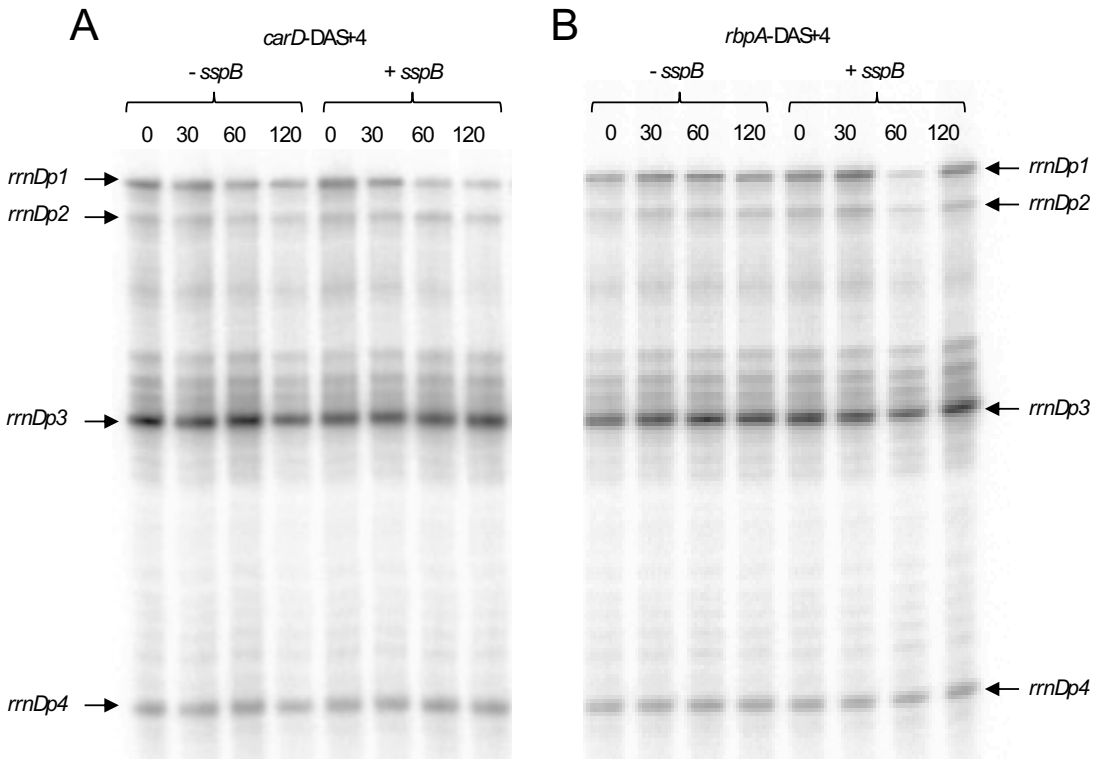


Figure 4.20 - **S1 nuclease mapping of *rrnDp1–4* following the depletion of CarD and RbpA.** (A) S1 nuclease protection analysis of S201/pIJ6902 (- *sspB*) and S201/pIJ6902::*sspB* (*+ sspB*) strains. Cultures were grown in YEME-10 to OD_{450} of 1 and RNA was isolated 30, 60 and 120 min following addition of thiostrepton (20 µg/mL final concentration). 30 µg RNA was hybridised to the *rrnDp1–4* probe and following S1 nuclease digestion, samples were run on an 8% denaturing urea polyacrylamide gel and visualised by phosphorimaging. (B) S1 nuclease protection analysis of S101/pRT802::*rbpA*-DAS+4/pIJ6902(- *sspB*) and S101/pRT802::*rbpA*-DAS+4/pIJ6902::*sspB* (*+ sspB*) strains. Experiment as described in Figure 4.20A.

4.8 Discussion

4.8.1 CarD is required for growth in *S. coelicolor*

Prior to this study, CarD had been identified as an essential gene in *M. tuberculosis* and *M. smegmatis* (Stallings et al., 2009). Attempts to delete the chromosomal *carD* allele were unsuccessful however it was possible to confirm the essentiality of CarD by creation of a strain possessing a single *carD* gene under the control of an inducible promoter. Both *M. tuberculosis* and *M. smegmatis* strains were not viable in the absence of the inducer.

In this study, initial attempts to make a CarD deletion mutant in *S. coelicolor* were also unsuccessful. By placing an additional copy of *carD* under the control of the thiostrepton-inducible promoter it was possible to create a strain with a single inducible copy of *carD*. In the absence of thiostrepton the strain grew very poorly, demonstrating that CarD is required for normal growth in *S. coelicolor*. Whilst it is difficult to strictly define a gene as "essential" it was assumed that small amounts of growth on solid media were attributable to a low level of *carD* expression from the *tipAp* promoter, previously characterised as "leaky" in the absence of inducer (M. Paget, personal communication).

In the presence of a range of thiostrepton concentrations, growth was restored however the phenotype was only partially complemented. At the highest concentrations of thiostrepton, colonies still exhibited a *whi* phenotype, typical of mutants blocked in the later stages of sporulation (Flärdh and Buttner, 2009). It is likely that this was due to low expression from the *tipAp* inducible promoter in aerial mycelium, potentially caused by the inability of the characterisitically insoluble thiostrepton to diffuse into the aerial mycelium.

carD under the control of its own promoter and integrated at an alternative locus was able to fully complement the inducible mutant in the presence and absence of thiostrepton. This likely confirms that the partially complemented S200 phenotype was due to lack of CarD and not unwanted polar effects or mutations introduced when creating the strain.

In addition to creation of a strain dependent on a thiostrepton-inducible copy of *carD*, a method of inducible, directed proteolysis was implemented for the first time in *Streptomyces* and used to investigate the role of CarD. Active depletion of CarD yielded

similar results, with colonies appearing small and white. One limitation of this method is that it is not possible to ascertain the extent to which CarD is depleted. If CarD is an essential gene then it is possible that small colonies are still able to form due to low concentrations of CarD present in the cell.

4.8.2 Purification of CarD and RNAP core

Experiments in this chapter outline the successful purification of CarD protein. This was achieved through heterologous expression of CarD with an N-terminal His-tag with the pET system in *E. coli*. Following cell lysis, His-CarD was purified from the cleared cell lysate by Ni-affinity chromatography followed by size-exclusion chromatography. Purification via this method achieved a yield of approximately 1.04 mg of protein from 250 mL culture, identified as functional by *in vitro* transcription assays.

This chapter also outlined the purification of RNAP from *S. coelicolor*. The method combined published methods for purification of RNAP utilisating initial PEG precipitation of nucleic acids and associated proteins (Gross et al., 1976; Fong et al., 2010), Ni-affinity against the His-tagged β' subunit (Babcock et al., 1997) and finally, separation of RNAP holoenzyme from RNAP core by ion exchange chromatography (Hahn et al., 2003). This method successfully produced functional RNAP core with no activity in the absence of additional purified σ. This procedure is not without limitations; in addition to the length of time the purification takes and relatively low yield, it is also not always possible to purify RNAP ensuring absence of additional products. Whilst this study outlined purification of RNAP from a Δ*rbpA* strain, it was not possible to purify RNAP in a Δ*carD* background. Additionally Hahn et al. (2003) reported that the final ion exchange chromatography necessary for removal of σ subunits may result in loss of ω subunit. Recently Czyz et al. (2014) reported a method for purification of recombinant, active *M. bovis* RNAP core from *E. coli*, with four subunit genes co-expressed from a single plasmid (used in chapter 6). This method, also used for purification of *T. aquaticus* (Minakhin et al., 2001) and *E. coli* RNAP core (Artsimovitch et al., 2003), would be an attractive solution for purification of *S. coelicolor* RNAP in future studies.

4.8.3 CarD activates transcription from σ^{HrdB}-dependent promoters *in vitro*

Following purification of CarD and RNAP, experiments performed in this study have shown that *S. coelicolor* CarD activates transcription from a selection of σ^{HrdB}-dependent promoters *in vitro*, including rRNA promoters *rrnDp3* and *rrnDp4*. These data replicated findings showing that *M. tuberculosis* CarD activates transcription from σ^{A}-dependent rRNA promoters *in vitro* (Srivastava et al., 2013). Experiments also showed that CarD activates transcription from a diverse selection of σ^{HrdB}-dependent including genes involved in central metabolism, *atpI* and *rplJ*, and the actinorhodin CSR, *actII-ORF4*. Results obtained in this experiment are reminiscent of studies performed on RbpA which identify the protein as an activator of transcription from σ^{HrdB}-dependent promoters, including *rrnDp3, atpIp* and *rplJp* (Newell et al., 2006; Tabib-Salazar et al., 2013).

4.8.4 CarD and RbpA selectively activate transcription from *rrnD* promoters *in vitro*

Further *in vitro* transcription experiments performed on the entire *rrnD* promoter region, containing four promoters *rrnDp1–4*, identified differences in regulation by CarD and RbpA. The *rrnDp2* promoter was active in the absence of both CarD and RbpA and addition of either or both proteins had no effect on transcription from this promoter *in vitro*. The *rrnDp3* and *rrnDp4* promoters were very poorly expressed in the absence of CarD and RbpA however both promoters were activated by addition of the two proteins. When *in vitro* transcription experiments were performed on the *rrnDp2* and *rrnDp3* promoters in isolation, similar results were found.

These results draw remarkable comparisons with studies performed by Hahn and Roe, (2007) into regulation of rRNA in *S. coelicolor*. The study showed that partially purified RNAP preparations were able to support transcription from *rrnDp3* and *rrnDp4* howevever reconstituted RNAP core and σ^{HrdB} did were not. Fractionation of the partially purified RNAP preparation by size and reconstitution into an *in vitro* transcription reaction identified a protein roughly 30–35 kDa able to restore transcription from the *rrnDp3* and *rrnDp4* promoters. Whilst the molecular of weights of CarD and RbpA are 18 and 14 kDa, respectively, the results of this experiment would otherwise replicate the findings from this study. Further proteomic analysis of this fraction would either identify

the protein in question as CarD or RbpA, or alternatively identify a novel protein with a potentially similar function.

4.8.5 CarD and RbpA stimulate promoter activity in the absence of -35 σ_4 interactions

Following the discovery by Hahn and Roe (2007) that expression of *rrnDp3* and *rrnDp4* required the presence of an additional factor, this raised the questions about why the promoter was unable to support transcription without this factor in a minimal system. Alignments of the four *rrnD* promoters revealed that absence of a recognisable σ^{HrdB} -35 promoter elements in the p3 and p4 promoters. To investigate if this was responsible for the requirement of CarD and RbpA for transcription *in vitro*, synthetic promoter fragments were constructed exchanging the -35 hexamer from the p2 and p3 promoters. Indeed, altering the -35 element changed the requirement for CarD and RbpA.

The p2 promoter which was previously unaffected by presence or absence of both proteins now required CarD or RbpA and presence of both proteins had an additive effect.

Modifications to include a conserved -35 region abolished the requirement of the p3 promoter for CarD or RbpA. Remarkably, addition of either CarD, RbpA or both proteins inhibited transcription. Despite a number of studies into the properties of CarD and RbpA as transcriptional activators, this is the first reported incidence of either protein inhibiting transcription *in vitro.*

The phage Φ29 protein p4 has similar properties, able to activate or repress transcription from *B. subtilis* σ^A-dependent promoters. The late A3 promoter is activated by p4, whilst the early A2c promoter is repressed by p4 (Monsalve et al., 1996a, 1996b). Closer inspection of these promoters revealed that the difference in p4 activity, which binds DNA upstream and makes contact with the α-CTD, is dependent on promoter sequence (Monsalve et al., 1997). The activated A3 promoter lacks a -35 element, whilst the A2c has a conserved -35 element. Further *in vitro* transcription assays identified that the mechanism of inhibition was through preventing promoter escape into elongation.

It is possible that CarD and RbpA have the ability to repress transcription through a similar mechanism to Φ29 protein p4, by providing an extra point of contact on DNA and overstabilising RNAP at the promoter. Both proteins are thought to interact with DNA

upstream of the -10 element (Srivastava et al., 2013; Hubin et al., 2015). Indeed, increased contact with DNA upstream of the -10 element has been shown to stabilise RNAP through interaction between an extended -10 region and σ_3 (Barne et al., 1997). Additionally *in vitro* transcription data suggests that presence of an extended -10 element can absolve requirement the -35 element or σ_4 altogether (Yuzenkova et al., 2011). Most interestingly this result brings into question whether CarD and RbpA are able to inhibit transcription *in vivo*, particularly at naturally occurring promoters with similar properties to the synthetic promoter created in this study.

4.8.6 Depletion of CarD or RbpA has no effect on *rrnDp1–4* transcript levels

Having shown that the *rrnDp3* and *rrnDp4* promoters required CarD or RbpA for transcription *in vitro*, depletion studies were performed to investigate the requirement for transcription *in vivo*. Using the the SspB-dependent ClpX system demonstrated elsewhere in this study, CarD and RbpA were depleted and S1 nuclease protection analysis was performed. Follow depletion no changes to the *rrnDp1–4* transcript level was observed. Unfortunately it was not possible to determine the extent to which either protein was depleted. In future studies it would be possible to determine the presence or absence of either protein through generation of an antibody and western blot analysis.

Additionally, *in vitro* transcription data identified that presence of either CarD or RbpA was required for transcription of *rrnDp3* and *rrnDp4.* It therefore remains a possibility that CarD and RbpA were depleted in this experiment however no effect was observed due to the presence of RbpA and CarD, respectively. If so this would raise further questions about an overlapping role of CarD and RbpA in *S. coelicolor*.

Chapter 5:

Results III: ChIP-seq analysis of transcription initiation factors in *S. coelicolor*

5 ChIP-seq analysis of transcription initiation factors in *S. coelicolor*

5.1 Overview

This chapter first outlines the creation of a *S. coelicolor* strain possessing a single copy of *carD* with a C-terminal 3xFLAG tag, designated S202. ChIP-seq was performed on mid–late exponential phase cultures of this strain using antibodies specific to RNAP, σ^{HrdB}, and CarD-3xFLAG. Additional sequencing data was analysed from a similar experiment performed on RbpA in the presence and absence of rifampicin. ChIP-seq data revealed that CarD and RbpA co-localise with σ^{HrdB} at all σ^{HrdB}-dependent promoters in *S. coelicolor.*

5.2 ChIP-seq analysis of CarD

5.2.1 Construction of a 3xFLAG tagged *carD* allele

As an alternative to ordering a custom antibody, it is a common approach to fuse an epitope tag to the protein of interest allowing recognition of the recombinant protein with a readily available monoclonal antibody. One such example of this is the FLAG® tag system which encodes a small, hydrophilic 8-amino acid peptide (DYKDDDDK) (Hopp et al., 1988). Due to the hydrophilic nature of the FLAG tag, it is typically surface exposed allowing the anti-FLAG antibody to readily recognise and bind. As an improvement upon this system, a 3xFLAG tag comprises 3 consecutive FLAG epitopes (DYKDHDGDYKDHDIDYKDDDDK) and detection is up to 200 times more sensitive than any other epitope tag system ("FLAG® and 3xFLAG® Overview", Sigma Aldrich).

It was decided to express the *carD*-3xFLAG allele in a *S. coelicolor* strain in which the native *carD* gene had been deleted, ensuring that the epitope-tagged protein was the dominant cellular species. To allow this, pRT802, which integrates into the ϕBT1 attachment site, was adopted as cloning vector. Prior to the construction of a 3xFLAG-tagged *carD* allele in the appropriate vector, an initial plasmid was constructed in which *rbpA* was fused to a 3xFLAG tag with HindIII restriction sites upstream of the *rbpA* promoter and at the intersection between the *rbpA* ORF and the in-frame C-terminal FLAG tag. To achieve this, *rbpA*-3xFLAG was first amplified by PCR using the primers RbpA_3xFLAG_F/RbpA_3xFLAG_R and the plasmid pMT3000::*rbpA*-3xFLAG as a

template. The PCR product was cloned into EcoRV-cut pBlueScript SKII+ and sub-cloned into pRT802 as a SpeI/BamHI fragment, producing the plasmid pRT802::*rbpA*-3xFLAG. This created a construct that allowed the *rbpA* portion of *rbpA*-3xFLAG to be swapped as a HindIII/HindIII fragment with *carD*. Thus, *carD* was PCR amplified from M145 genomic DNA using the primers carD_3xFLAG_F/carD_3xFLAG_R, which incorporated SpeI-HindIII restriction sites upstream of the *carD* promoter region and a HindIII site immediately upstream of the *carD* stop codon. The PCR product was cloned into EcoRV-cut pBlueScript SKII+ and sub-cloned into pRT802::rbpA-3xFLAG as a HindIII/HindIII fragment, replacing the *rbpA* insert and producing the plasmid pRT802::*carD*-3xFLAG.

5.2.2 CarD-3xFLAG fully complements a Δ*carD* mutant

Before conducting any further experiments on the biological role of CarD using the CarD-3xFLAG expressing construct, it was necessary to confirm that the epitope-tagged recombinant protein was fully functional *in vivo* and that the addition of an N-terminal 3xFLAG tag does not affect protein function. The plasmid pRT802::*carD*-3xFLAG was used to transform ET12567/pR9406 and transferred via conjugation into the conditional *carD* mutant S200. The 3xFLAG-tagged construct fully complemented the mutant phenotype and removed any dependence of the strain for thiostrepton, compared with the positive and negative controls pRT802::*carD* and pRT802, respectively (Figure 5.1). This confirmed that a 3xFLAG tagged copy of CarD was fully functional in *S. coelicolor*.

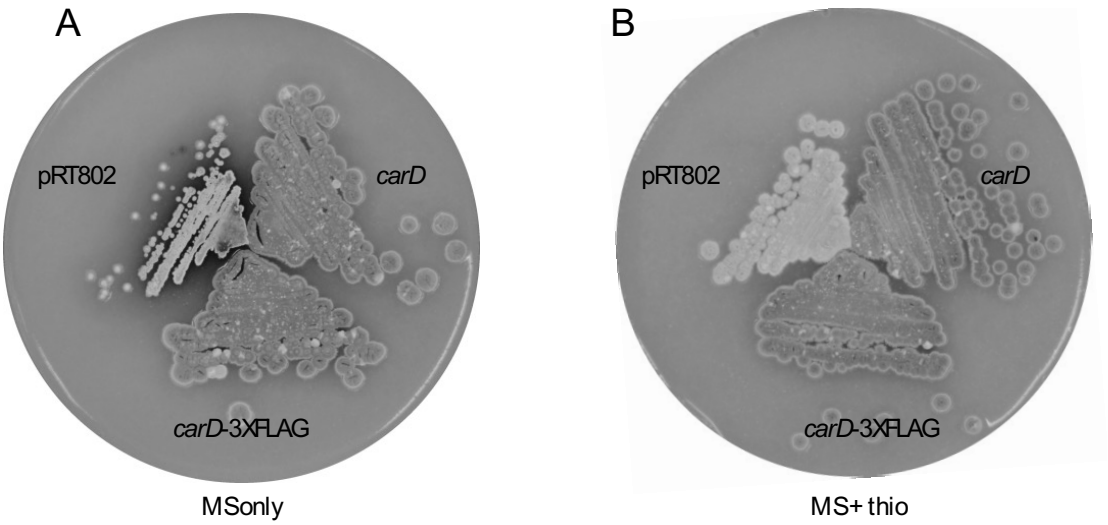


Figure 5.1 - ***carD*-3xFLAG complements a Δ*carD* deletion mutant.** S200/pRT802, S200/pRT802::*carD* and S200/pRT802::*carD*-3xFLAG streaked onto (A) MS agar (B) MS agar + 15 µg/mL thiostrepton and incubated for 5 days at 30 °C.

5.2.3 Creation of a Δ*carD*-3xFLAG mutant

It was necessary to create a strain possessing only one chromosomal copy of *carD* with a 3xFLAG tag. Due to the essential nature of CarD, it was not possible to first create a Δ*carD* mutant and complement this with *carD*-3xFLAG. Instead, the plasmid pRT802::*carD*-3xFLAG was transferred via conjugation into a single crossover recombinant strain J1915/pIJ6650Δ*carD::hyg* (obtained in chapter 4). As the strain now possessed an additional functional copy of *carD* (*carD*-3xFLAG integrated at the phage ϕBT1 site), following a round of non-selective growth it was now possible to obtain Apr^S Hyg^R Kan^R double crossover deletion mutants representing the strain J1915Δ*carD::hyg/pRT802::carD-3xFLAG*. This strain was named S202.

5.2.4 Chromatin immunoprecipitation (ChIP) on *S. coelicolor* grown in liquid cultures

Chromatin immunoprecipitation (ChIP) is a technique used to study protein-DNA interactions. It involves crosslinking all protein-DNA complexes, precipitating the protein of interest with an appropriate antibody and analysing the DNA fragments purified, typically through qPCR or high-throughput DNA sequencing.

To perform the crosslinking, cultures of S202 were grown in YEME-10 (with glycerol as a carbon source) to mid-late exponential phase (OD_{450} = 1.5–2). 37% formaldehyde solution was added to a final concentration of 1% before incubation for a further 20 min at 30 °C. To quench the remaining formaldehyde, glycine was added to a final concentration of 0.5 M before incubation for 5 min at room temperature. For each immunoprecipitation, 70 mL aliquots of cells were harvested by centrifugation at 6,000 x g for 2 min at 4 °C, washed twice with 25 mL ice-cold PBS and cell pellets were stored at -80 °C until required.

Keeping frozen in liquid nitrogen, pellets were cryogenically ground for 3 x 90 s before resuspension in 2.2 mL IP buffer (with 0.1 mg/mL RNase A + protease inhibitor). Cells were further disrupted and DNA fragmented by sonication using a Diagenode Bioruptor. To do this, each pellet suspension was divided into 6 x 400 µL aliquots and sonicated for 35 cycles of sonication (30 s on, 30 s off) at 4 °C. Samples were centrifuged at 16,000 x g for 30 min to remove cell debris and the supernatant transferred to a new tube.

To visualise the degree of DNA fragmentation, a 20 µL sample of the supernatant was treated with 2 µL Riboshredder RNase Blend at room temperature for 30 min followed by 2 µL proteinase K at 55 °C for 2 hours and de-crosslinked at 65 °C for 6 hours. This sample was run on a 1.2% agarose gel to check DNA fragment size; a smear ranging from ~150 bp to 600 bp was seen and judged suitable for sequencing.

To perform the immunoprecipitation, the chromatin samples were first pre-cleared with the appropriate Protein A/G magnetic beads; this is a process that reduces non-specific binding by removing proteins that may bind immunoglobulins. Protein G magnetic beads (New England Biolabs) were used for RNAP and 3xFLAG immunoprecipitations and Protein A magnetic beads (New England Biolabs) were used for σ^{HrdB} immunoprecipitation. Pre-clearing was performed with 50 µL magnetic beads incubated at 4 °C for 1.5 hours rotating at 20 rpm. Following incubation, the magnetic beads were immobilised with a magnetic rack and the cleared supernatant was transferred to a new tube. 100 µL of this pre-cleared chromatin was removed and stored at -20 °C as an "input" control. 1–5 µg of the following antibodies was added to the remaining cleared chromatin and incubated at 4 °C for 2.5 hours rotating at 20 rpm:

- 7 µL anti-FLAG M2 monoclonal antibody (Sigma F18041MG)
- 7 µL anti-σ^{HrdB} polyclonal antibody (a gift from P. Doughty)
- 2 µL anti-RNAP β monoclonal antibody (Abcam ab12087)

30 µL of appropriate magnetic beads were added and incubated for a further 1.5 hours. The beads were immobilised with a magnetic rack and the spent supernatant was discarded. The beads were gently resuspended in 750 µL ice-cold IP buffer, transferred to a new tube and incubated at 4°C for 10 min rotating at 20 rpm. The same wash procedure was performed with 1 mL ice-cold IP + salt buffer, IP wash buffer and finally pH 8 TE buffer. After washing, the magnetic beads were resuspended in 100 µL 1 x IP elution buffer + 5 µL RiboShredder RNase blend and incubated at room temperature for 30 min. At this point, the “input” control sample was thawed and incubated with 5 µL RiboShredder RNase blend at room temperature for 30 min before de-crosslinking both the IP and input samples in a 65°C water bath overnight. The following day the beads were immobilised and washed with an additional 50 µL pH 7.5 TE buffer. 50 µL pH 7.5 TE

buffer was added to the input sample. 150 µL IP and input samples were finally purified through a Qiagen MinElute PCR purification kit and eluted with 22 µL MQ water.

5.2.5 Aligning and visualising ChIP-seq data

The samples were sequenced using an Illumina HiSeq 2500 platform (Rapid-Run mode) with 50 bp single-end reads. 16 samples were simultaneously sequenced via multiplex sequencing. This process was performed externally (The Genome Analysis Centre (TGAC), Norwich, UK) and de-multiplexed sequencing results were received as individual FASTQ files for each ChIP sample.

For downstream processing, files were uploaded to Galaxy (https://usegalaxy.org), an "open, web-based platform for data intensive biomedical research". To allow downstream processing, FASTQ files were first converted from Illumina to Sanger format with FASTQ Groomer (Galaxy Tool version 1.0.4).

Before alignment and analysis, it was necessary to perform quality control on the FASTQ sequence files as evidence of successful sequencing of uncontaminated samples using FASTQC (Galaxy Tool Version 0.63). For all samples the mean GC content passed QC with the around 72% (Figure 5.2A). Results indicated a decrease in average quality and deviation from the expected GC content at positions 1–10 of the 50 bp reads for every result (Figure 5.2B). It has been reported that this phenomena caused by random hexamer mis-priming is often seen with Illumina NGS data and may not affect downstream applications. However, to ensure this did not affect the ability of Bowtie to map the reads onto the *S. coelicolor* genome, the first 10 bases of every read were removed with Trim sequences as a precaution (Galaxy Tool version 1.0.0). For all immunoprecipitation results, FASTQC failed on the sequence duplication level (Figure 5.2C). Errors at this step in ChIP-seq are typically introduced via the PCR amplification step. As library preparation did not include any PCR steps and the sequence duplication level was normal for the input samples (Figure 5.2D) it was considered that this result was due by high coverage and would not affect results.

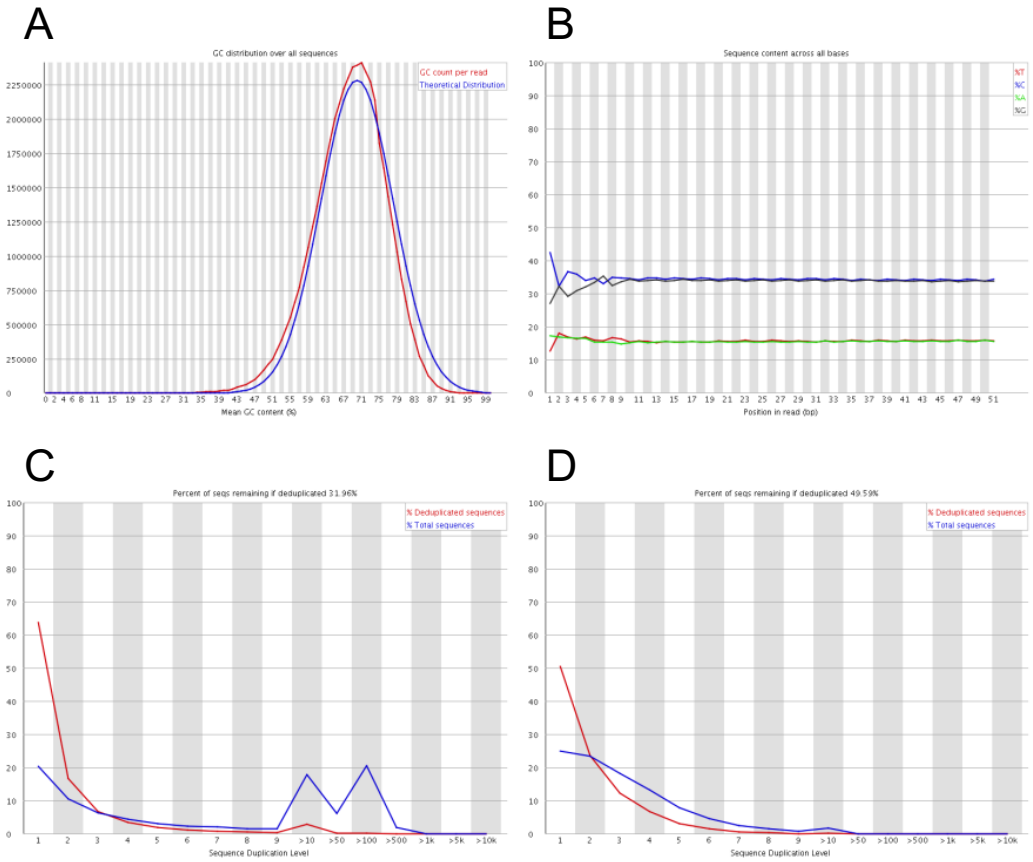


Figure 5.2 - **FASTQC analysis of ChIP-seq results**. (A) Mean GC content result for RNAP sample (B) Per base sequence content across all bases for RNAP sample (C) Sequence duplication level result for RNAP sample (D) Sequence duplication level result for input sample.

Once the FASTQ files were suitably prepared, the processed reads were aligned to the *S. coelicolor* A3(2) genome sequence (NCBI reference sequence: NC_003888.3) with Bowtie for Illumina (Galaxy Tool version 1.1.2, using the settings "-n 2 -e 70 -S --un") (Langmead et al., 2009). By using the --un mode, unmapped reads were written to a separate file and compared to the number of mapped reads as a measure of read mapping efficiency. Percentage mapped reads of 97.6% to 98.9% revealed that the alignment was successful (Table 5.1).

Sample	Mapped Reads	Unmapped Reads	Total Reads	% Mapped
Input	22,998,668	242,041	23,240,709	99.0%
RNAP	16,256,217	253,948	16,510,165	98.5%
σ^{HrdB}	24,146,596	370,590	24,517,186	98.5%
CarD	28,004,757	675,407	28,680,164	97.6%

Table 5.1 - **Percentage of reads mapped with Bowtie**. Mapped and unmapped read counts were obtained with IdxStats (Galaxy Tool Version 2.0).

The Bowtie output SAM file was compressed to a binary BAM file using SAM to BAM (Galaxy Tool version 2.0) before sorting by chromosomal co-ordinates with the Sort BAM dataset tool (Galaxy Tool version 2.0).

To create histogram plots of the mapped reads, the sorted BAM files were uploaded to the Galaxy deepTools server (http://deeptools.ie-freiburg.mpg.de) and the bamCoverage tool (Galaxy Tool Version 1.5.9.1.0) was used for creation of a bigWig file (Ramírez et al., 2014). The average fragment size captured from the ChIP samples in the final stages of library preparation by TGAC was 200 bp; this fragment size was entered into the bamCoverage parameters and used to extend the length of every mapped read. As ChIP-seq involves high throughput sequencing of very short reads, in this case just 50 bp, likely to map to the 5' ends of each strand, the reads are likely to map to the left and right of the true protein location (Myers et al., 2015). The process of extending the read length compensates for this and gives a truer representation of the protein-DNA interaction *in vivo*. To normalise for sequencing depth, all samples were normalised to reads per kilobase per million mapped reads (RPKM) (Mortazavi et al., 2008; Bailey et al., 2013). The process of aligning and visualising ChIP-seq reads is outlined in Figure 5.3. Integrated Genome Browser (IGB) (Version 8.3.4) was used to visualise the bigWig histogram against annotated *S. coelicolor* A3(2) genome sequence (Nicol et al., 2009) (Figure 5.4).

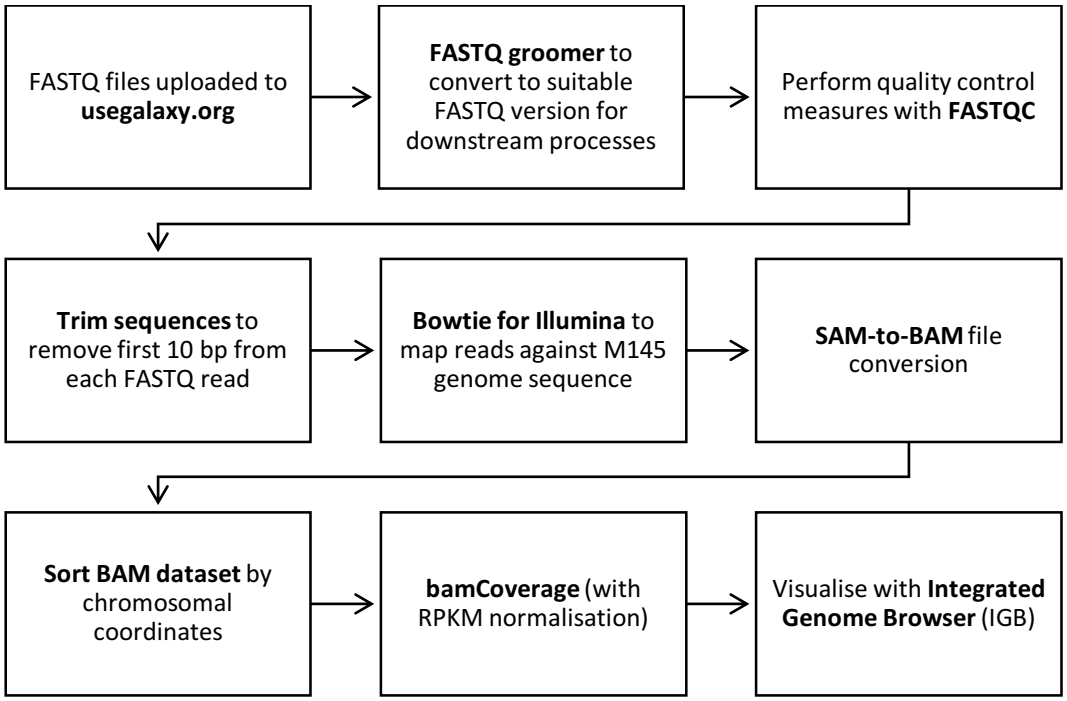


Figure 5.3 - **Flowchart outlining the procedures for ChIP-seq alignment and analysis.**

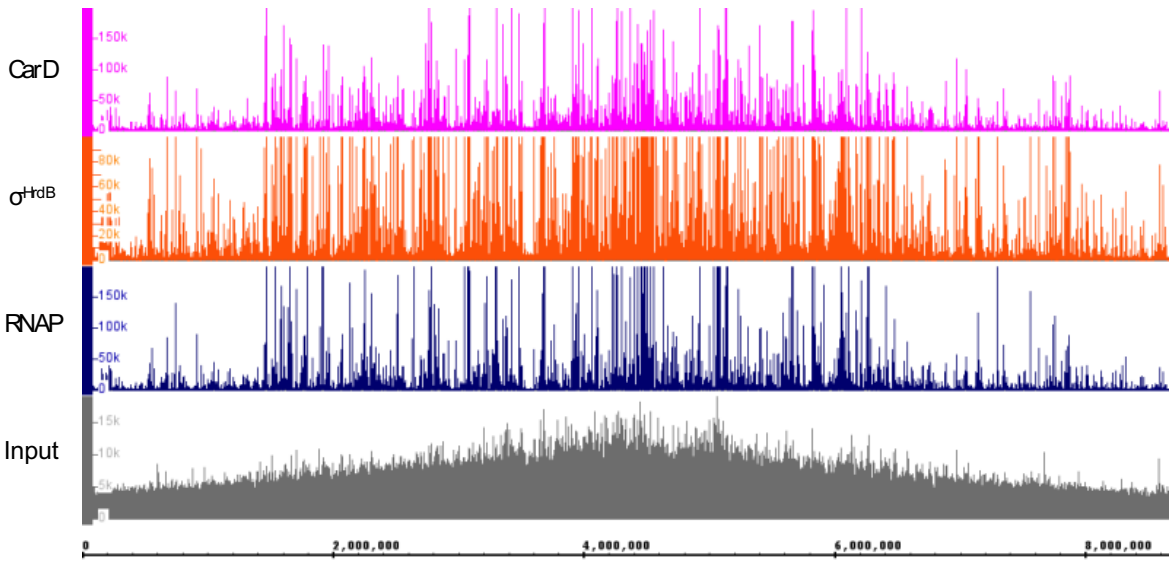


Figure 5.4 - **Visualisation of bigWig histogram files with Integrated Genome Browser** ChIP-seq analysis performed on S202. bigWig histograms represent total number of aligned reads, RPKM normalised for sequencing depth. The entire *S. coelicolor* genome is represented. Input tracks are grey, RNAP are blue, σ^{HrdB} are orange and CarD are pink.

5.3 CarD and σ^{HrdB} co-localise at σ^{HrdB}-dependent promoters *in vivo*

Initial analysis of the distribution of RNAP, CarD and σ^{HrdB} was performed by calculating the correlation of the entire datasets using deepTools bamCorrelate (Galaxy Tool version 1.5.9.1). Pearson correlation analysis revealed that the distribution of CarD and σ^{HrdB} was highly correlated with a value of 0.89 (Figure 5.5). Data also revealed that distribution of RNAP correlated with CarD (0.55) and σ^{HrdB} (0.60). Positive correlation, albeit with low scores, between the input and RNAP (0.22), σ^{HrdB} (0.29), and CarD (0.37) is likely to be due, at least in part, to the relationship between gene position and expression. In *S. coelicolor*, the highly expressed core genes responsible for essential functions including replication, transcription and translation are positioned in the central core region of the linear chromosome (Bentley et al., 2002). Accessory genes, including cryptic/silent secondary metabolic gene clusters, tend to be located in the chromosome arms. In this ChIP-seq experiment, alignment of the sequenced input DNA revealed a higher number of reads for DNA derived from the central region of the genome, as expected given the centrally located origin of replication (Figure 5.4). Therefore, a positive correlation would be expected due to the presence of highly expressed genes being in the core of the genome where the gene copy number is higher due active growth and replication.

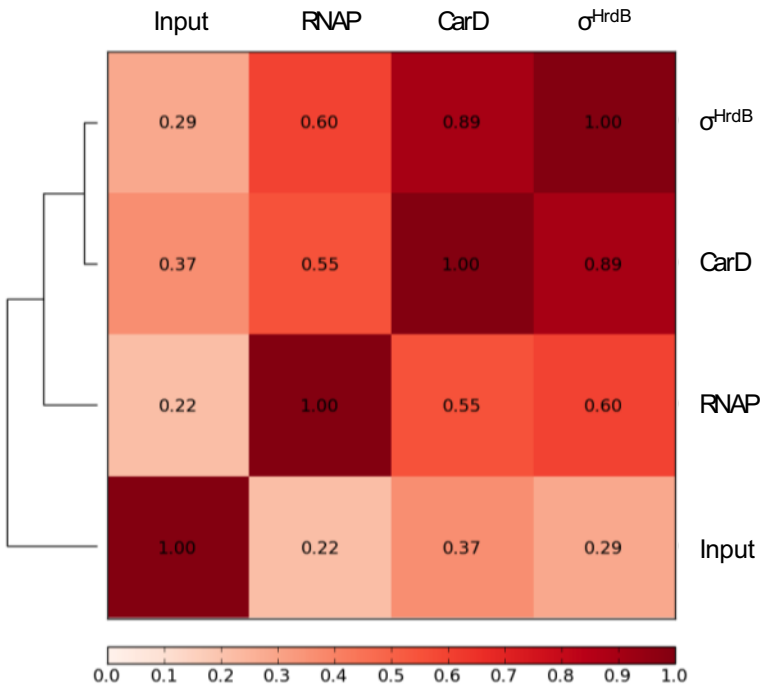


Figure 5.5 - **Distribution of CarD and σ^{HrdB} is highly correlated throughout *S. coelicolor* genome** Pearson correlation of aligned ChIP-seq files performed with deepTools bamCorrelate, using a bin size of 500 bp.

The regions surrounding σ^{HrdB}-dependent genes SCO4808 (succinyl-CoA synthetase beta chain, *sucC*), SCO1947 (glyceraldehyde-3-phosphate dehydrogenase, *gap1*), SCO2136 (putative secreted protein) and SCOr07–9 (*rrnD*) are four examples that represent observations made throughout the genome (Figure 5.6A–D).

σ^{HrdB} enrichment was distributed throughout the genome primarily in intergenic regions, upstream of genes, typical of promoter location (Figure 5.6A–D). These enriched regions are consistent with predicted and known transcription start sites, thus confirming that this ChIP-seq method is reliable for identification of σ^{HrdB}-dependent promoters. Little or no σ^{HrdB} enrichment was detected within the main body of genes, consistent with the efficient dissociation of σ from RNAP core during elongation (Raffaelle et al., 2005). However, it should be noted that cross-linking efficiency between σ and DNA is likely to decrease during the transition from initial transcribing complex to elongation and so the absence of enrichment does not necessarily correlate with an absence of σ in the elongating complex.

RNAP enrichment was distributed more widely throughout the genome, at expected promoter regions as well as downstream into ORFs. However, enrichment was significantly greater at promoter regions than downstream into the gene (Figure 5.6A–D) suggesting that the transition from initiation to elongation is limiting factor for most promoters, as has been suggested for *E. coli* (Reppas et al., 2006). Furthermore, as also seen in *E. coli* (Reppas et al., 2006), the enrichment of RNAP at promoters compared to the main gene body was highly variable; for example the travelling ratio for exemplar genes in Figure 5.6, defined by the enrichment 800 bp into the gene divided by enrichment at the promoter, was 0.03 for *sucC*, 0.09 for SCO1947 and 0.15 for SCO2136. The reason for this is not currently understood although one possibility is that the transition from initiation to elongation is highly variable, possibly reflecting regulatory differences between promoters. Analysis of the region surrounding rRNA operons revealed that while RNAP accumulates at the promoter, it is proportionally distributed more evenly throughout the operon (Figure 5.6D). To maintain high levels of rRNA expression it is likely that RNAP proceeds more rapidly from these promoters into elongation. Under the growth conditions used in this experiment, the vast majority of RNAP peaks upstream from ORFs correlated with the presence of σ^{HrdB}. Two notable

exceptions were *hrdB* (SCO5820) and *cwgA* (SCO6179) which are controlled by the alternative sigma factors σ^{ShbA} and σ^{E}, respectively (Otani et al., 2013; Hong et al., 2002) (see chapter 6).

CarD enrichment revealed that the protein co-localises with σ^{HrdB} and RNAP at σ^{HrdB}-dependent promoters throughout the genome (Figure 5.6A–D). It was found exclusively at promoter regions, suggesting that CarD is not a component of elongating RNAP. Interestingly, despite the absence of σ^{HrdB}, CarD was enriched in the *hrdB* (SCO5820) and *cwgA* (SCO6179) promoter regions, which suggested that CarD might be present in RNAP initiation complexes that contain alternative sigma factors; this discovery is further investigated in chapter 6.

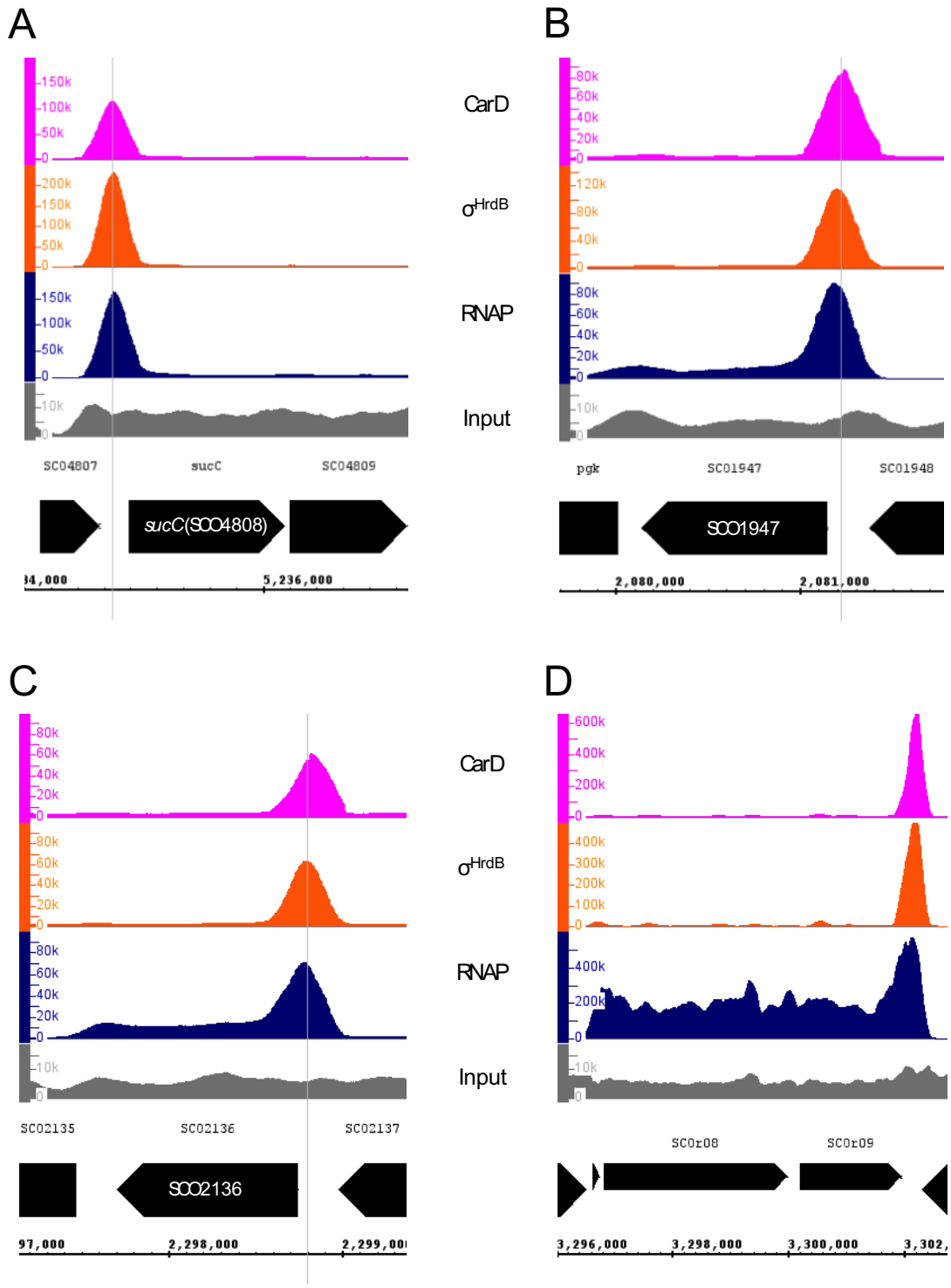


Figure 5.6 - **CarD and σ^{HrdB} co-localise at σ^{HrdB}-dependent promoters *in vivo*.** ChIP-seq analysis performed on S202. bigWig histograms represent total number of aligned reads, RPKM normalised for sequencing depth. (A) *sucC* (SCO4808) (B) *gap1* (SCO1947) (C) SCO2136 (putative secreted protein) (D) *rrnD* (SCOr07-9). Input tracks are grey, RNAP are blue, σ^{HrdB} are orange and CarD are pink. Predicted transcription start site (if available) is indicated with a grey line (B.K. Cho, personal communication).

5.4 RbpA and σ^{HrdB} co-localise at σ^{HrdB}-dependent promoters *in vivo*

ChIP experiments were also performed as described above on the strain S115/pRT802::*rbpA*-3xFLAG (M145Δ*rbpA::apr/pRT802::rbpA-3xFLAG*) by A. Tabib-Salazar. Sequencing results were received and analysed as part of this study. Due to the presence of *glkA* in the parent strain, cultures were grown in YEME-10 with glucose as a carbon source. As before, immunoprecipitations were performed using antibodies specific to RNAP β, σ^{HrdB} and 3xFLAG tagged RbpA. Otherwise, ChIP and sequencing procedures were performed as with S202. Samples from this experiment were run across two HiSeq lanes so produced two FASTQ files per sample. Following conversion to Sanger format with FASTQ groomer, the two FASTQ files from each sample were combined with the Concatenate Datasets" function (Galaxy Tool version 1.0.0). Once combined, data were prepared and aligned as described above (Figure 5.3). As before, comparison of the number of mapped to unmapped reads was performed and percentage mapped reads of 97.7% to 99% revealed that DNA fragment alignment was successful (Table 5.2).

Sample	Mapped Reads	Unmapped Reads	Total Reads	% Mapped
Input	18,431,999	194,693	18,626,692	99.0%
RNAP	39,283,744	582,635	39,866,379	98.5%
σ^{HrdB}	34,865,132	571,610	35,436,742	98.4%
RbpA	34,508,616	823,139	35,331,755	97.7%

Table 5.2 - **Percentage of reads mapped with Bowtie**. Mapped and unmapped read counts were obtained with IdxStats (Galaxy Tool Version 2.0).

The correlation of RNAP, σ^{HrdB} and RbpA distribution was analysed as described in Section 5.3. RbpA and σ^{HrdB} displayed a very high positive correlation (0.98; Figure 5.7), suggesting that they co-localise throughout the *S. coelicolor* genome. RNAP distribution showed a slightly lower positive correlation score with σ^{HrdB} (0.72) and RbpA (0.69), consistent with data presented for σ^{HrdB} and CarD. Furthermore, input samples were poorly correlated with and RNAP (0.24), RbpA (0.25) and σ^{HrdB} samples (0.30).

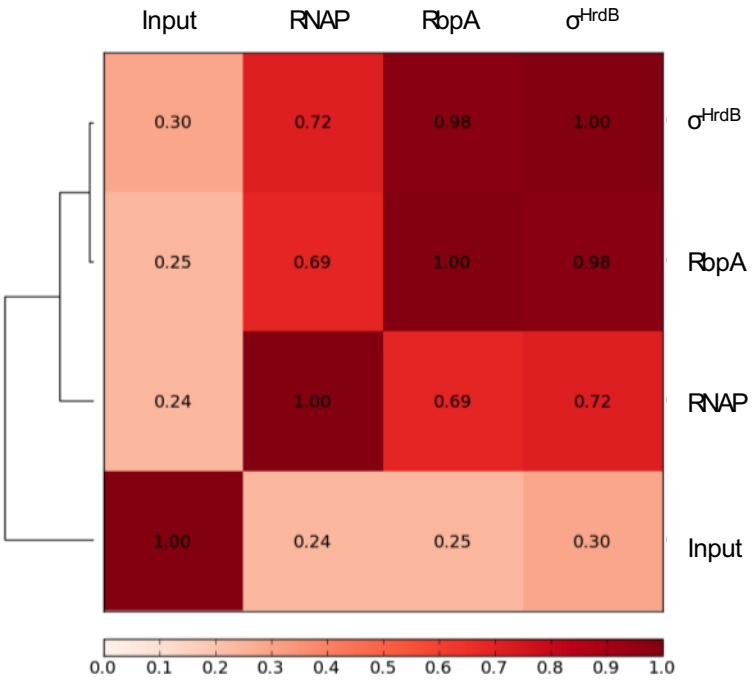


Figure 5.7 - **Distribution of RbpA and σ^{HrdB} is highly correlated throughout *S. coelicolor* genome** Pearson correlation of aligned ChIP-seq files performed with deepTools bamCorrelate, using a bin size of 500 bp.

The four representative σ^{HrdB}-dependent genes described in Section 5.3 were again chosen to illustrate key points: SCO4808 (succinyl-CoA synthetase beta chain, *sucC*), SCO1947 (glyceraldehyde-3-phosphate dehydrogenase, *gap1*), SCO2136 (putative secreted protein) and SCOr07–9 (*rrnD*) (Figure 5.8A–D).

The distribution of σ^{HrdB} and RNAP was highly similar to that observed in previous S202 ChIP-seq experiments (Figure 5.8) demonstrating the reproducibility of this technique for observing DNA binding *in vivo*. Consistent with the remarkably high correlation scores, RbpA and σ^{HrdB} co-localised throughout the genome at σ^{HrdB}-dependent promoters (Figure 5.8). There were no incidences observed throughout the entire genome where RbpA was present in the absence of σ^{HrdB} enrichment and *vice versa*.

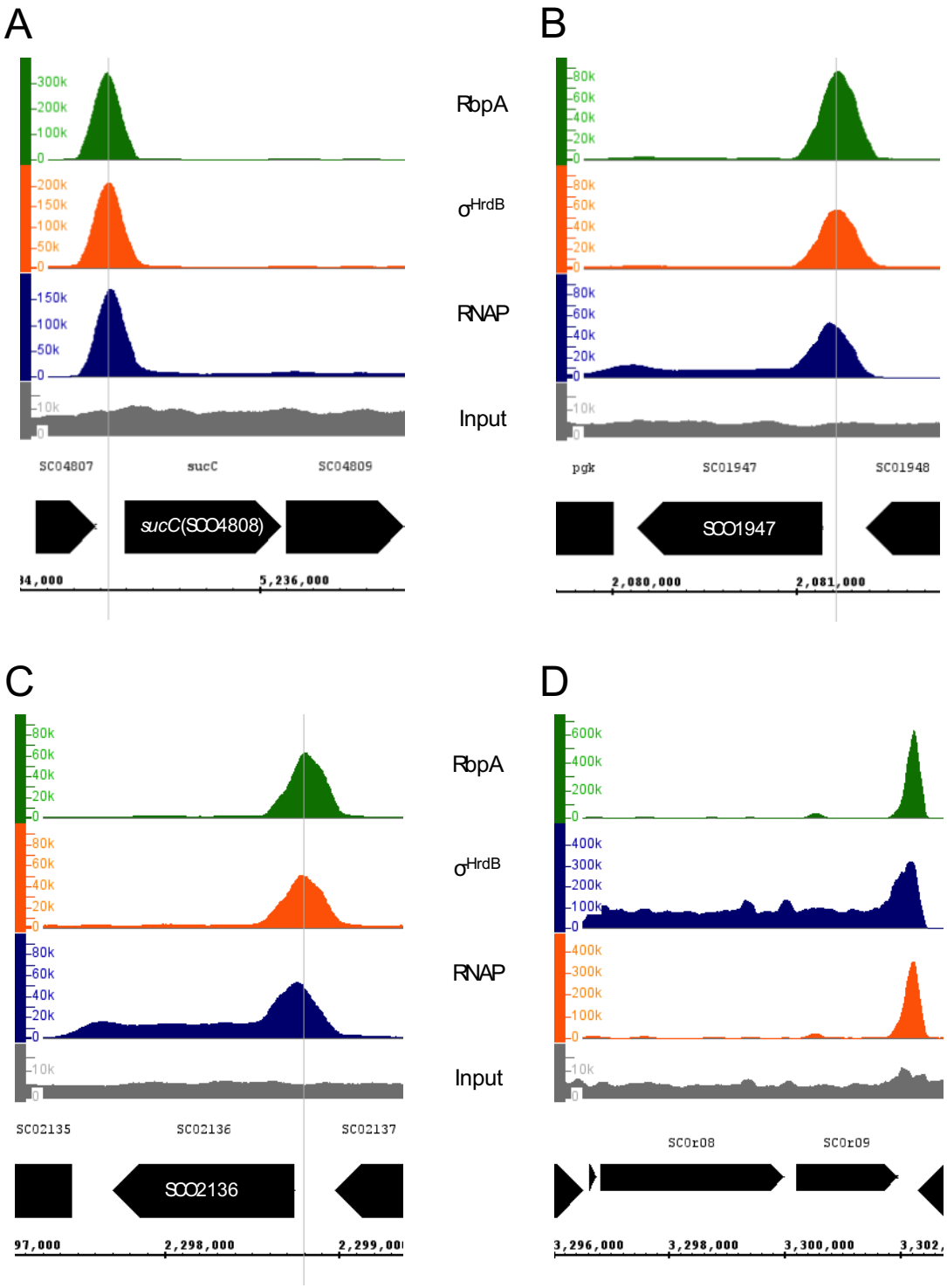


Figure 5.8 - **RbpA and σ^{HrdB} co-localise at σ^{HrdB}-dependent promoters *in vivo*** ChIP-seq analysis performed on S115/pRT802::*rbpA*-3xFLAG. bigWig histograms represent total number of aligned reads, RPKM normalised for sequencing depth. (A) *sucC* (SCO4808) (B) *gap1* (SCO1947) (C) SCO2136 (putative secreted protein) (D) *rrnD* (SCOr07–9). Input tracks are grey, RNAP are blue, σ^{HrdB} are orange and RbpA are green. Predicted transcription start site (if available) is indicated with a grey line (B.K. Cho, personal communication).

5.5 ChIP-seq analysis of effects of rifampicin

In previous ChIP-chip studies performed on RNAP, rifampicin has been used to treat cells prior to harvesting to aid identification of promoter regions (Herring et al., 2005; Grainger et al., 2005). Rifampicin is an antibiotic that inhibits bacterial transcription by binding to the β subunit in the DNA/RNA channel, preventing the enzyme from entering the elongation phase of transcription (Campbell et al., 2001). Consequently, rifampicin-bound RNA polymerase molecules are prevented from escaping initiation and are confined to promoter regions (Herring et al., 2005).

In addition to samples prepared from exponentially growing cultures, ChIP experiments were performed on S115/pRT802::*rbpA*-3xFLAG following addition of rifampicin by A. Tabib-Salazar. Mycelia was harvested from cultures exposed to the 140 μg/mL rifampicin for 20 min (conditions shown previously to be suitably for studying effects of rifampicin on transcription *in vivo*) (Herring et al., 2005). Immunoprecipitations were performed as above with antibodies specific to RNAP β, σ^{HrdB} and 3xFLAG tag. Samples were sequenced externally and sequencing was analysed as described above.

Comparison of the number of mapped to unmapped reads was performed and percentage mapped reads of 98.2% to 98.9% revealed that the alignment was successful (Table 5.3).

Sample	Mapped Reads	Unmapped Reads	Total Reads	% Mapped
Input + RIF	30,000,367	320,836	30,321,203	98.9%
RNAP + RIF	38,969,681	701,081	39,670,762	98.2%
σ^{HrdB} + RIF	34,013,740	440,525	34,454,265	98.7%
RbpA + RIF	38,020,694	626,876	38,647,570	98.4%

Table 5.3 - **Percentage of reads mapped with Bowtie**. Mapped and unmapped read counts were obtained with IdxStats (Galaxy Tool Version 2.0).

The correlation of RNAP, σ^{HrdB} and RbpA distribution was analysed as described above. Following addition of rifampicin, RNAP, RbpA and σ^{HrdB} were all distributed throughout the *S. coelicolor* with a very high correlation (Figure 5.7). With comparison to untreated samples, the correlation of RbpA and σ^{HrdB} was unchanged. However, RNAP correlation with σ^{HrdB} (and RbpA) increased considerably following addition of rifampicin (0.72 to 0.97), suggesting that RNAP was globally trapped in initiating transcription complexes, unable to escape promoters into elongation.

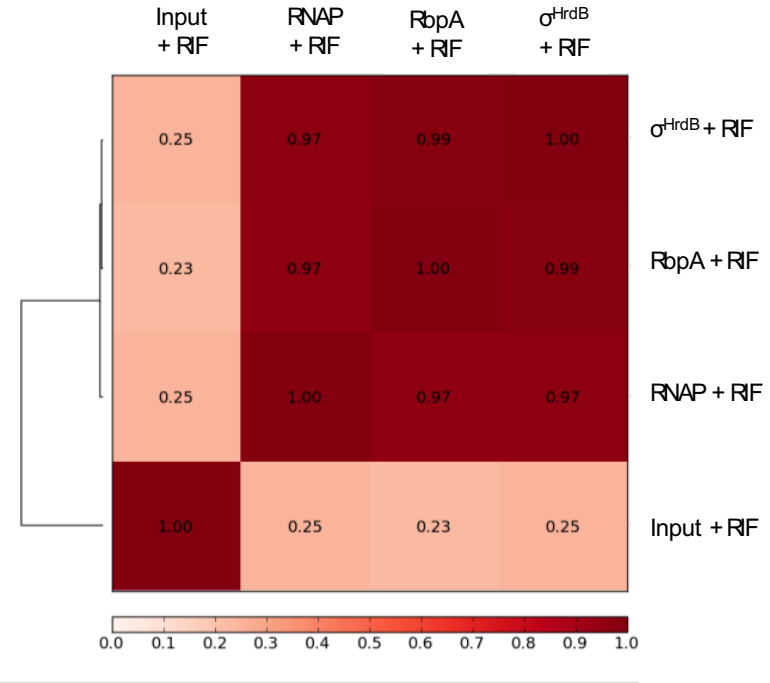


Figure 5.9 - **Addition of rifampicin increases the correlation of RNAP, RbpA and σ^{HrdB} across the *S. coelicolor* genome**
Pearson correlation of aligned ChIP-seq files performed with deepTools bamCorrelate using a bin size of 500 bp.

This was confirmed by visualisation of bigWig files using IGB, which indicated a general increase in the enrichment of RNAP and σ^{HrdB} at the promoter regions, and a general decrease in coding sequences. As illustration, there is a large difference in RNAP enrichment enriched throughout the ATP synthase operon (*atpIBEFHAGDC*) prior to and after the addition of rifampicin (Figure 5.10A). The 20 min period following the addition of rifampicin, prior to cross-linking would be expected to be long enough to allow RNAP to complete the current round of transcription at most genes. However, *Streptomyces* spp. are characterised by some extremely large antibiotic biosynthetic operons where this might not be the case. The CDA peptide synthetase operon at 40 kb is the largest operon in the *S. coelicolor* genome. Prior to the addition of rifampicin, RNAP was

enriched in the promoter-proximal region of the *cda* operon, suggesting either that expression of the *cda* operon had only recently started at the time of sampling, or that the initial stages of transcription elongation are limiting. Following rifampicin treatment, there was a striking redistribution of RNAP to the promoter-distal region of the operon (Figure 5.10B). Notwithstanding the unknown time taken for rifampicin to inhibit further rounds of transcription initiation, the location of the trailing edge of RNAP enrichment should indicate distance travelled by RNAP in 20 min, and allow an estimation of transcription speed. Thus, the position of the trailing edge of RNAP at ~23 kb into the operon suggests a transcription speed of 19 nt per second. This can be compared to that in *E. coli*, where the speed of RNAP elongation has been estimated to be approximately 30–100 nt per second (Vogel and Jensen, 1994).

In addition to a global decrease in elongating RNAP, treatment with rifampicin caused the appearance of a number of new RNAP and σ^{HrdB} peaks throughout the genome where transcription would not be expected to initiate (Figure 5.10). This is consistent with previous observations from ChIP experiments performed on rifampicin-treated cells (Herring et al., 2005; Grainger et al., 2005). One possibility is that these sites represent unannotated ORFs or small regulatory RNAs. It is also possible that an increase in free RNAP following rifampicin treatment allows RNAP and σ^{HrdB} to bind at genomic loci not observed under normal growth conditions.

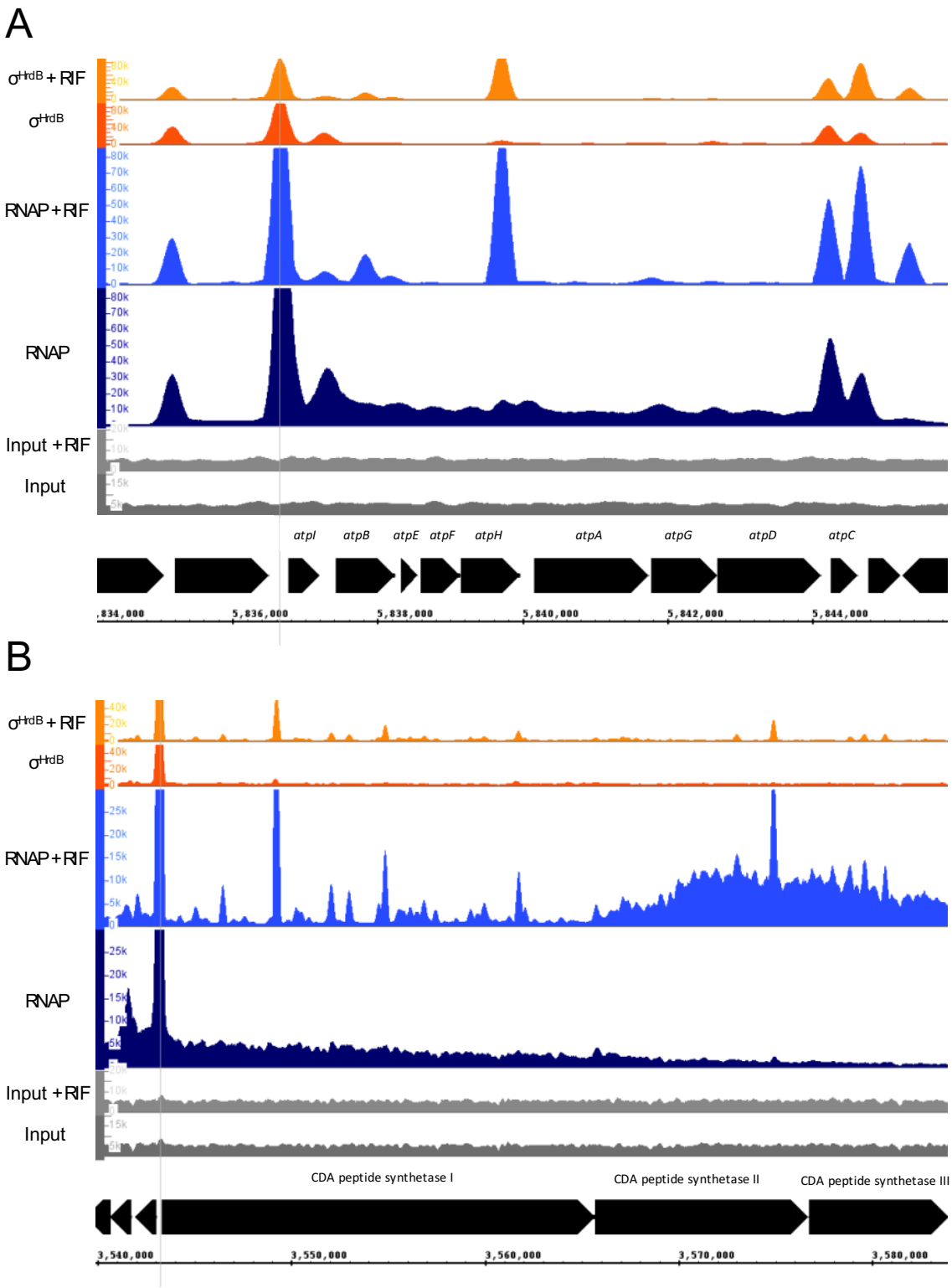


Figure 5.10 - **Rifampicin inhibits transcription initiation.** ChIP-seq analysis performed on S115/pRT802::*rbpA*-3xFLAG before and 20 mins after addition of 140 µg/mL (A) ATP synthase operon (B) CDA synthetase operon. Input tracks are grey, RNAP are blue and σ^{HrdB} are orange. Predicted transcription start site (if available) is indicated with a grey line (B.K. Cho, personal communication).

Although rifampicin inhibited transcription from the majority of promoters, it was possible to identify a number of promoters where transcription was activated. For example *hflX*, which encodes a universally conserved GTP-binding protein, displayed increased RNAP enrichment throughout the ORF following rifampicin treatment, which correlated with increased RNAP enrichment at the promoter. Interestingly, however, σ^{HrdB} enrichment was unaffected suggesting that the increased transcription activity did not involve σ^{HrdB} (Figure 5.11A). A similar effect was also observed at the *rpmE*/*prfA* operon (Figure 5.11B), although in this case level of σ^{HrdB} enrichment at the promoter actually decreased. Whilst both promoters bind σ^{HrdB}, these genes have also been identified as σ^{R} targets (Paget et al., 2001a). This suggests that rifampicin treatment increases transcription from promoters dependent on σ^{R} or a related ECF σ factor.

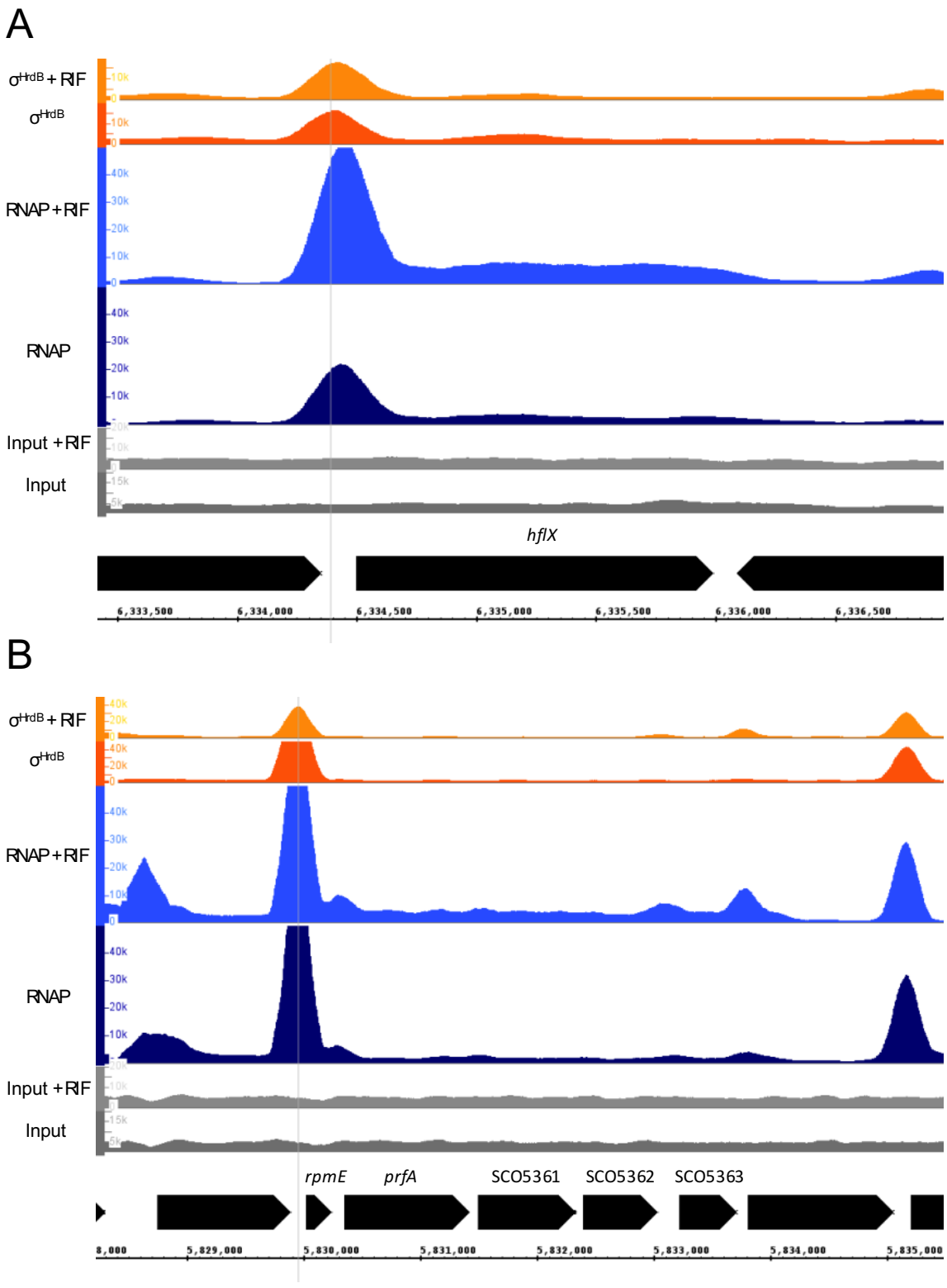


Figure 5.11 - **Rifampicin increases transcription of selected genes.** ChIP-seq analysis performed on S115/pRT802::*rbpA*-3xFLAG before and 20 mins after addition of 140 µg/mL (A) *hflX* (B) *rpmE*/*prfA* operon. Input tracks are grey, RNAP are blue and σ^{HrdB} are orange. Predicted transcription start site (if available) is indicated with a grey line (B.K. Cho, personal communication).

5.6 Discussion

5.6.1 CarD and σ^{HrdB} co-localise at σ^{HrdB}-dependent promoters *in vivo*

ChIP-seq has emerged as a key technique for studying the localisation of transcription factors in all organisms and has recently been applied to *Streptomyces* for the location of a range of DNA binding proteins including the developmental regulator WhiA (Bush et al., 2013) and the oxidative stress regulator NdgR (Kim et al., 2015). The current work demonstrates that the technique is also suitable for studying the distribution of the transcription machinery in *S. coelicolor* despite the much wider distribution of RNA polymerase. Sequencing depth is an important consideration when studying such broadly distributed transcription factors, since an insufficient numbers of reads might lead to saturation at highly expressed genes or the inability to detect occupancy at poorly expressed genes (Sims et al., 2014). While this point had not been extensively investigated here, with as low as ~16 million reads per sample, there was clearly no saturation within gene bodies, and RNAP was readily detected in coding sequences throughout the genome. The application of RNAP ChIP-seq to quantify gene expression has been termed "RNA polymerase-omics" and it can be argued that it provides a better snapshot of active transcription compared to the analysis of the transcriptome (Grainger and Busby, 2008).

The data presented here demonstrate that in mid-late exponential cultures, CarD co-localises with σ^{HrdB} at transcription intiation complexes. Similarly, during the course of these experiments, a related study revealed that CarD co-localises with the principal σ factor σ^{A} at promoters in *M. smegmatis* (Srivastava et al., 2013; Landick et al., 2014). Together, these data suggest that CarD is a genome-wide transcription initiation factor in all actinobacteria, and possibly in other phyla. Data presented in chapter 4 demonstrate that CarD is essential for growth and can activate a wide range of promoters, and the global distribution of this transcription activator is consistent with this.

Sigma factors dissociate from elongating RNA polymerase soon after the enzyme enters the elongation phase (Raffaelle et al., 2005). Since CarD was not detected in main gene bodies, it too appears to dissociate as soon as RNAP escapes the promoter. However, the

mechanism for this dissociation remains unknown. In the case of σ, dissociation is a multi-step process that begins when the nascent RNA displaces the region 3.2 linker from the RNA exit channel. It is thought that this leads to disruption of the σ_4/β flap interactions, followed by stochastic dissociation of σ_2 from the β' coiled coil as RNAP enters elongation. Furthermore, it has been proposed that NusG competes with σ for binding to the β' coiled coil, which drives the process forward, swapping initiation factor for elongation factor and trapping RNAP in elongation mode (Sevostyanova et al., 2008). It is possible that the structure of elongating RNAP is somehow incompatible with CarD binding, leading to its dissociation. Furthermore, similar to the case of σ and NusG, the CarD binding site is shared with other cellular factors. CarD binds to the β1 domain via its RNAP interacting domain (RID). The same region of β is also contacted by other RID-containing proteins including TRCF, which is involved in transcription coupled repair (Westblade et al., 2010). However, TRCF is unlikely to be involved in the displacement of CarD since it is thought to specifically associate with stalled RNAP elongation complexes.

5.6.2 RbpA and σ^{HrdB} co-localise at σ^{HrdB}-dependent promoters *in vivo*

Experiments performed in this chapter detail the first time RbpA has been localised at a genome level. Similar to experiments performed with CarD, the data presented here demonstrate that in mid-late exponential cultures, RbpA co-localises with σ^{HrdB} at transcription intiation complexes. These findings are consistent with recently published interaction and structural data showing that RbpA specifically binds σ^{HrdB} and σ^{A} in *S. coelicolor* and *M. tuberculosis*, respectively. (Tabib-Salazar et al., 2013; Hubin et al., 2015). It has also been proposed that RbpA binds elsewhere on RNAP through the uncharacterised RCD or N-terminal domains, with one possible location being the β' clamp domain (Hubin et al., 2015). It is therefore conceivable that RbpA is present in the absence of σ^{HrdB} on elongation complexes. However, ChIP-seq data suggests this is not the case, with enrichment data showing no RbpA present in the absence of σ^{HrdB} throughout the genome. Additionally, observations show that enrichment peaks for RbpA and σ^{HrdB} are identical in shape and relative size. It is therefore proposed that RbpA dissociates with σ^{HrdB} following transcription initiation.

RbpA was previously not thought to be a DNA binding protein; however, recent structural studies show that RbpA contacts DNA near the upstream edge of the -10 element through

conserved arginine and lysine residues. Indeed, formaldehyde has been shown to crosslink RbpA with DNA (Hubin et al., 2015). This therefore suggests that the crosslinking step used for ChIP-seq analysis is likely to capture interactions between RbpA and DNA, as well as DNA crosslinks via RNAP.

The presence of RbpA at transcription initiation complexes does not appear to indicate whether RbpA plays a role in initiation events. As described in chapter 4, not all promoters are likely to require RbpA for activity. With this in mind, the key question for the future is understanding what factors determine whether RbpA plays a role in transcription initiation. Considering RbpA appears to be present at all σ^{HrdB}-dependent promoters, this suggests that such determinants may lie in the intrinsic properties of the promoter or potential regulation of RbpA transcription, translation or post-translational modifications.

5.6.3 CarD and RbpA co-localise at σ^{HrdB}-dependent promoters *in vivo*

Data from independent ChIP-seq experiments shows that CarD and RbpA both co-localise with σ^{HrdB} at initiation complexes (Figure 5.12). This suggests that both proteins may be present on the same initiation complex. ChIP-seq experiments performed in this study were unable to identify whether this is the case although structural modelling has suggested that binding of both proteins to one holoenzyme is possible (E. Hubin, E. Campbell and S. Darst, personal communication) (Figure 5.13). Interestingly, both proteins are suggested to interact with the minor groove upstream of the -10 element, on the opposite side of DNA to σ_3, which interacts with an extended -10 element (Barne et al., 1997). Therefore, in principle, if CarD and RbpA binding is compatible *in vivo*, all three interactions could occur to stabilise transcription complexes. This raises interesting questions regarding how the stoichiometry of CarD and RbpA may affect transcription. For example, following oxidative stress, RbpA transcription is upregulated and CarD is downregulated (Kallifidas et al., 2010). It is possible that such changes in the stoichiometry of CarD and RbpA may have interesting regulatory consequences.

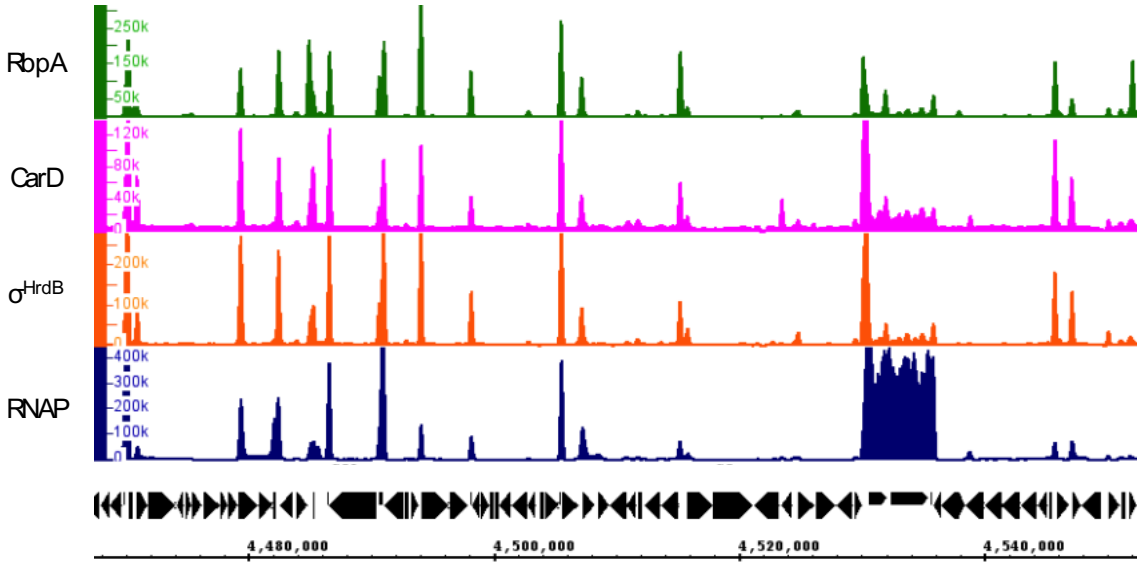


Figure 5.12 - **RbpA and CarD co-localise at σ^{HrdB}-dependent promoters *in vivo*.** ChIP-seq analysis performed on S202 (RNAP, σ^{HrdB}, CarD) and S115/pRT802::*rbpA*-3xFLAG (RbpA). bigWig histograms represent total number of aligned reads, RPKM normalised for sequencing depth. RNAP track is blue, σ^{HrdB} is orange, CarD is pink and RbpA is green.

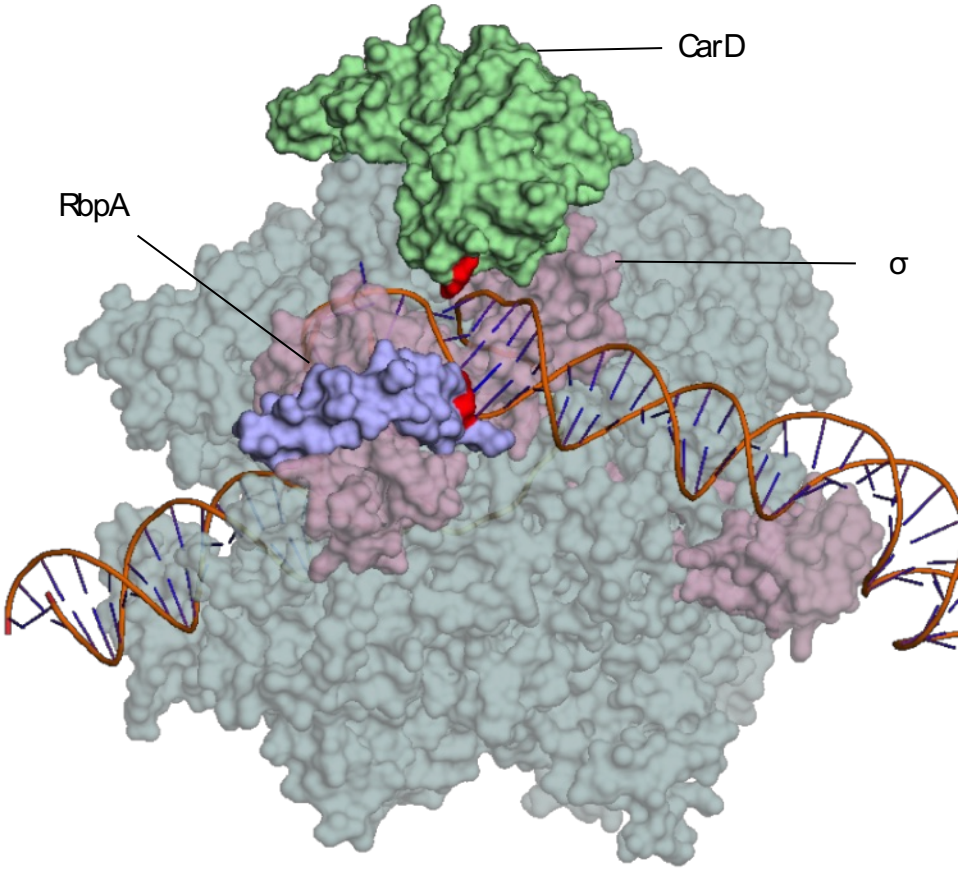


Figure 5.13 - **Structural model of CarD and RbpA.** Structural model of *T. aquaticus* RNAP holoenzyme, *T. thermophilus* CarD and *M. tuberculosis* RbpA in complex with promoter DNA. RNAP core is light blue, σ is pink, CarD is green and RbpA is purple, DNA is shown as a phosphate backbone worm. Reproduced with permission from E. Hubin, E. Campbell and S. Darst.

5.6.4 Rifampicin has diverse effects on transcription

Rifampicin is an antibiotic that inhibits bacterial transcription by binding to the β subunit in the DNA/RNA channel, preventing the enzyme from entering the elongation phase of transcription (Campbell et al., 2001). Consequently, rifampicin-bound RNA polymerase molecules are prevented from escaping initiation and are confined to promoter regions (Herring et al., 2005). ChIP-seq experiments performed in this chapter identify the localisation of RNAP, σ^{HrdB} and RbpA before and after treatment of cells with rifampicin. Consistent with published data (Herring et al., 2005; Grainger et al., 2005), treatment with rifampicin reduced distribution of RNAP throughout coding regions but increased enrichment at promoter regions. Additionally, RNAP, σ^{HrdB} and RbpA peaks appeared at new, unexpected locations where transcription would not be expected to initiate. In actively growing cells, where the majority of RNAP is performing transcription and free RNAP is limited, inhibition of transcription initiation is likely to lead to a large increase in the level of free core RNAP. The increase of free core RNAP might therefore allow RNAP to bind at weaker sites and could explain the appearance of new peaks. Although this is unlikely to be a physiologically relevant response, comparisons can be made to the stringent response in which inhibition of transcription by ppGpp leads to redistribution of RNAP. In the passive model for positive control of promoters unregulated as part of the stringent response, it is proposed that an increase in the concentration of free RNAP leads to stimulation of amino acid biosynthetic promoters (Barker et al., 2001a).

The concentration of rifampicin used in this experiment is lethal and was expected to inhibit all transcription. However, preliminary evidence suggests that some genes were still expressed even after 20 minutes of treatment. Two of the promoters that have been identified as still active, *hflX* and *rpmE*/*prfA,* are members of the σ^{R}-regulon (Kallifidas et al., 2010). It has previously been observed that σ^{R} activity is induced by sub-lethal concentrations of rifampicin and it has been proposed that this is due to increased core concentration able to compete with anti-sigma RsrA (Newell et al., 2006). Indeed, modelling experiments have also suggested that an increase in core RNAP would lead to a passive rise of alternative holoenzyme forms (Mauri and Klumpp, 2014). Despite this, discovery of active transcription at high concentrations of rifampicin is surprising and suggests that holoenzymes may differ in their sensitivity to rifampicin although the

reason for this is unknown. This finding may also have medical implications, with rifampicin used as a first-line drug in the treatment of tuberculosis. *M. tuberculosis* has a σ^R orthologue, σ^H that is also thought to play a key role in protein quality control (Manganelli et al., 2002). Consequently, results from this study suggest that the σ^H regulon could potentially be induced following rifampicin treatment.

Chapter 6:

Results IV: The sigma specificity of CarD and RbpA

6 The sigma specificity of CarD and RbpA

6.1 Overview

ChIP-seq studies revealed that CarD and RbpA are present at all σ^{HrdB}-dependent promoters in *S. coelicolor* and *in vitro* transcription experiments demonstrated that each activate transcription from σ^{HrdB}-dependent promoters. To investigate the association of CarD with alternative holoenzymes, ChIP-seq and ChIP-qPCR were performed following induction of the σ^{R} regulon with diamide. The enrichment of CarD, σ^{R} and RNAP at σ^{R}-dependent promoters, together with the absence of σ^{HrdB}, revealed that CarD is a component of alternative forms of holoenzyme during transcription initiation. Despite this, *in vitro* transcription experiments performed on σ^{R}-dependent promoters showed that addition of CarD had no effect on transcription.

RbpA has been shown to bind specifically to primary and primary-related but not alternative sigma factors (Tabib-Salazar et al., 2013). To investigate this specificity, the structure of RbpA-SID in complex with domain 2 of σ^{A} was solved and, following sequence alignments, site-directed mutagenesis and *in vitro* transcription experiments, a number of key residues essential for RbpA-σ^{A} interaction were identified.

6.2 CarD is present at alternative sigma factor promoters in association with alternative holoenzymes

6.2.1 CarD is present at σ^{ShbA} and σ^{E}-dependent promoters under normal growth conditions

ChIP-seq experiments revealed that CarD is present at all σ^{HrdB}-dependent promoters in *S. coelicolor in vivo*. This finding is concurrent with results showing that CarD is present at σ^{A}-dependent promoters throughout the *M. smegmatis* chromosome (Srivastava et al., 2013). Whilst ChIP-seq studies show that CarD and σ^{A} are highly co-localised and both appear to dissociate from RNAP following initiation, structural studies have shown no interaction between the two proteins when modelled on a *T. thermophilus* open complex (Srivastava et al., 2013). This raises the question of whether CarD is able to bind to RNAP holoenzymes containing alternative sigma factors.

Under normal growth conditions used for ChIP-seq experiments in chapter 5, the vast majority of transcription is directed from σ^{HrdB}-dependent promoters. Despite this, it was possible to identify two active promoters regulated by alternative sigma factors. The *hrdB* (SCO5820) promoter is regulated by the alternative sigma factor σ^{ShbA} (Otani et al., 2013). Analysis of the ChIP-seq data for this promoter region revealed RNAP and CarD present at the promoter (Figure 6.1A). Small amounts of σ^{HrdB} enrichment was also observed meaning it was not possible to attribute presence of CarD at the *hrdB* promoter as bound to a σ^{ShbA} holoenzyme. The *cwgA* (SCO6179) promoter is regulated by the alternative sigma factor σ^{E} (Hong et al., 2002). Analysis of this promoter region revealed presence of RNAP and CarD but no σ^{HrdB} (Figure 6.1B). This suggests that CarD is present at the *cwgA* promoter bound to a σ^{E} holoenzyme.

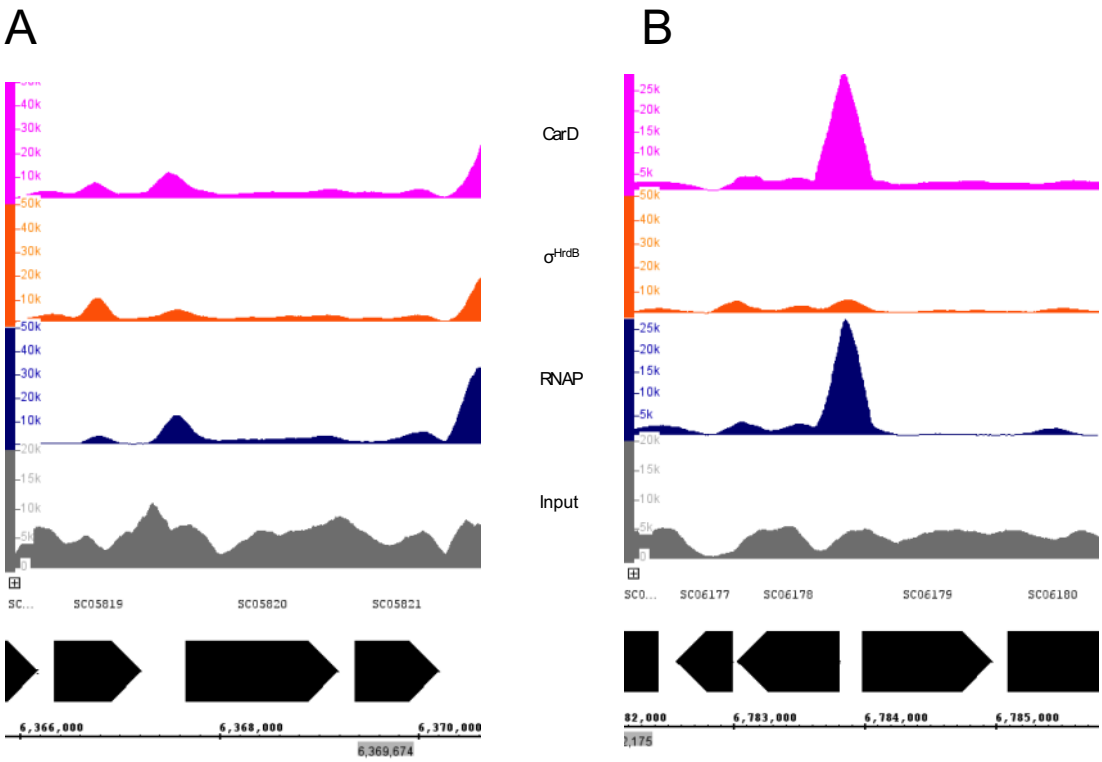


Figure 6.1 - **CarD is present at σ^{ShbA} and σ^{E}-dependent promoters.** ChIP-seq analysis performed on S202. bigWig histograms represent total number of aligned reads, RPKM normalised for sequencing depth. **(A)** *hrdB* (SCO5820) **(B)** *cwgA* (SCO6179). Input tracks are grey, RNAP are blue, σ^{HrdB} are orange and CarD are pink.

6.2.2 Chromatin Immunoprecipitation following diamide treatment

To further investigate whether CarD is able to associate with non-σ^{HrdB} holoenzymes, it was necessary to induce an alternative sigma factor regulon. Cultures of S202 were grown in YEME-10 (with glycerol as a carbon source) to mid-late exponential phase (OD_{450} = 1.5-2). To induce the σ^R regulon, 0.5 mM diamide was added to cultures followed by incubation for 15 minutes prior to crosslinking with formaldehyde. This concentration and time-point was chosen for maximum expression of the σ^R regulon (Paget et al., 1998, 2001). Following crosslinking, ChIP protocols were followed as described in chapter 5. Immunoprecipitations were performed as before with the following antibodies:

- 7 µL anti-FLAG M2 monoclonal antibody (Sigma F18041MG)
- 7 µL anti-σ^{HrdB} polyclonal antibody (a gift from P. Doughty)
- 2 µL anti-RNAP β monoclonal antibody (Abcam ab12087)
- 7 µl anti-σ^R polyclonal antibody (a gift from M. Feeney and M. Buttner)

6.2.3 ChIP-seq: CarD is present at all σ^R-dependent promoters

Library preparation and sequencing was performed by TGAC as described in chapter 5. De-multiplexed sequencing results were received as individual FASTQ files for each ChIP sample. FASTQ files were analysed, processed and aligned as described in chapter 5. All samples were efficiently aligned with proportion of mapped reads ranging from 96-99% (Table 6.1). Samples obtained from control cultures that had not been treated with diamide were described in chapter 5. In the absence of diamide, the σ^R antibody failed to immunoprecipitate sufficient DNA for ChIP-seq analysis and so this sample was not sequenced.

Sample	Mapped Reads	Unmapped Reads	Total Reads	% mapped
Input*	22,998,668	242,041	23,240,709	99.0%
RNAP*	16,256,217	253,948	16,510,165	98.5%
σ^{HrdB}*	24,146,596	370,590	24,517,186	98.5%
CarD*	28,004,757	675,407	28,680,164	97.6%
Input + DIA	17,169,352	624,660	17,794,012	96.5%
RNAP + DIA	16,731,644	284,060	17,015,704	98.3%
σ^{HrdB} + DIA	14,073,796	221,455	14,295,251	98.5%
σ^{R} + DIA	28,041,346	1,078,579	29,119,925	96.3%
CarD + DIA	23,264,961	489,931	23,754,892	97.9%

Table 6.1 - **Percentage of reads mapped with Bowtie**. Mapped and unmapped read counts were obtained with IdxStats (Galaxy Tool Version 2.0). Samples marked with an asterix are from experiments performed in chapter 5.

bigWig histogram plots of mapped reads were created for each sample as described in chapter 5 and visualised with IGB. The region surrounding σ^{R}-dependent genes SCO1936 (probable transaldolase tal2), SCO2763 (ABC transporter protein), SCO4109 (possible oxidoreductase) and SCO6551 (probable oxidoreductase) are four examples that represent observations throughout the genome (Figure 6.2). Each gene was previously reported as a σ^{R} target (Kim et al., 2012)., consistent with presence of enrichment peaks upstream from each ORF in the "σ^{R} + DIA" sample. This confirmed that these genes are σ^{R} targets and that σ^{R} was present at the promoter under these experimental conditions. When comparing untreated and treated samples, addition of diamide increased CarD and RNAP enrichment at the promoter region. The co-enrichment of RNAP, σ^{R}, and CarD suggests that upon addition of diamide, σ^{R} recruits RNAP to the promoters with CarD present on the holoenzyme. The lack of σ^{HrdB} enrichment in both the presence and absence of diamide demonstrates that CarD does not require σ^{HrdB} to associate with transcription initiation complexes.

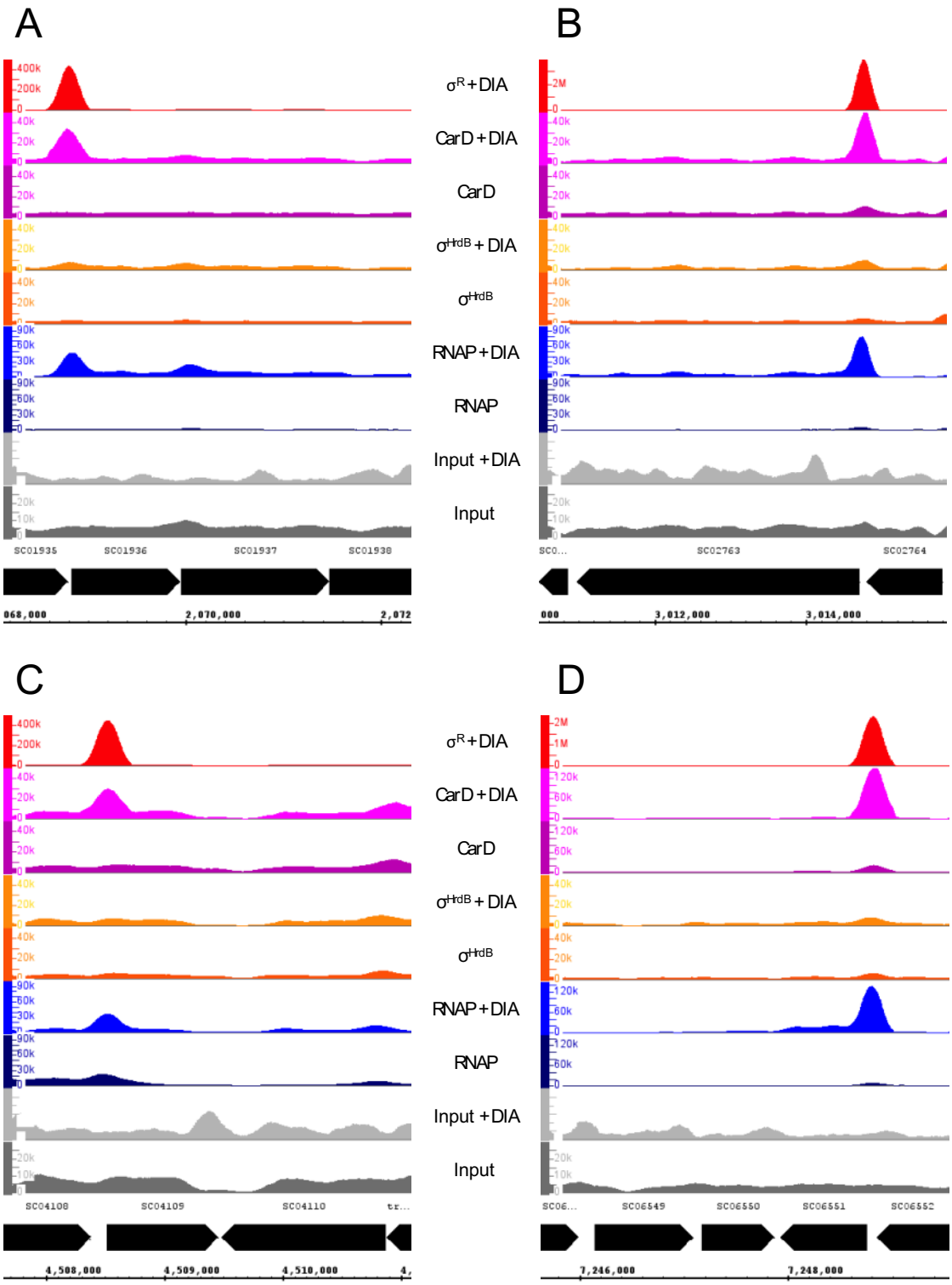


Figure 6.2 - **CarD is present at σ^R-dependent promoters.** ChIP-seq analysis performed on S202 in the presence and absence of 0.5 mM diamide ("+DIA"). bigWig histograms represent total number of aligned reads, RPKM normalised for sequencing depth. Regions with σ^R targets **(A)** SCO1936 (probable transaldolase tal2) **(B)** SCO2763 (ABC transporter protein) **(C)** SCO4109 (possible oxidoreductase) **(D)** SCO6551 (probable oxidoreductase). Input tracks are grey, RNAP are blue, σ^{HrdB} are orange, CarD are pink and σ^R is red.

6.2.4 ChIP-qPCR: CarD is present at σ^R-dependent promoters

qPCR was performed with the primer sets PSCO6551_F/PSCO6551_R and PSCO2763_F/PSCO2763_R, which amplify the promoter region of SCO6551, and SCO2763, respectively. SCO6551 is a probably oxidoreductase, previously identified as a σ^R target by S1 nuclease, microarray and ChIP-on-chip experiments (Paget et al., 2001a; Kallifidas et al., 2010; Kim et al., 2012). SCO2763 is a putative ATP-binding cassette (ABC) transporter and has also been identified as a σ^R target by S1 nuclease mapping and ChIP-on-chip (Kim et al., 2012). In addition to being σ^R-regulated promoters, these two promoters were identified as suitable for qPCR analysis due to their non-divergent nature and lack of additional known promoters within 1 kb upstream or downstream. A standard curve was prepared with serial dilutions of *S. coelicolor* gDNA containing 3,000,000 to 300 copies of the genome. ChIP input and elutions samples were diluted 1:1000 and 1:100, respectively.

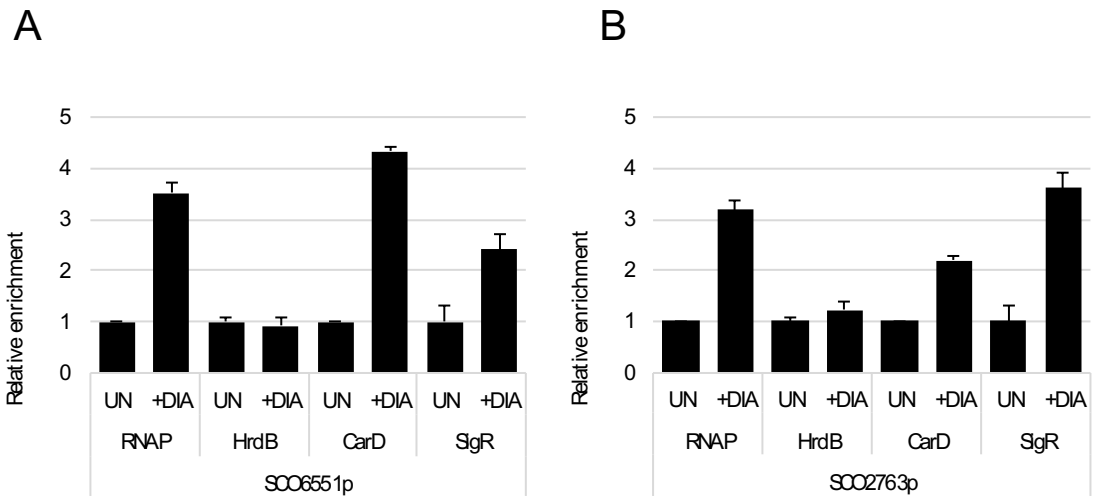


Figure 6.3 - **CarD is present at σ^R-dependent promoters.** Occupancy of RNAP, HrdB, CarD and SigR at **(A)** SCO6551p **(B)** SCO2763p following induction of oxidative stress with diamide. ChIP-qPCR data is quantified against a standard curve prepared with *S. coelicolor* gDNA, normalised to input and shown relative to untreated sample data. Standard deviations from two technical replicates are indicated as error bars.

ChIP-qPCR results show that following induction of oxidative stress with diamide, RNAP, CarD and SigR accumulate at the promoter region of SCO6551 (Figure 6.3A) and SCO2763 (Figure 6.3B). Enrichment of σ^{HrdB} was unaffected by treatment with diamide.

6.2.5 Identification of novel σ^R targets

Transcriptome (Kallifidas et al., 2010) and ChIP-on-chip studies (Kim et al., 2012) have provided a detailed picture of the σ^R regulon with 108 promoters identified, expressing a total of 163 genes. However, it was possible that these microarray-based studies might not have identified all σ^R binding sites, for example, if a region was missing from the microarray tiles. Therefore, to identify possible additional members of the σ^R regulon, peak calling was performed on "σ^R + DIA" ChIP alignments. MACS2 (version 2.1.0.20140616.0, using the settings "q=1e-7 -nolambda -nomodel -keepdup=0 -slocal=100 -llocal=500") (Zhang et al., 2008) was used to call peaks on the sorted BAM file with the "input + DIA" sample as a negative control. The list was further filtered with stringent conditions ($-\log_{10}$ p-value > 500, $-\log_{10}$ q-value > 2000 and fold enrichment > 10). Loci previously reported by Kallifidas et al. (2010) and Kim et al. (2012) and were excluded from the dataset. Finally, as the immunoprecipitation was performed with a polyclonal antibody and to ensure hits were true σ^R targets, only genes that had an increase in RNAP enrichment but no change in σ^{HrdB} enrichment following exposure to diamide (Figure 6.4) were reported. The process identified 14 previously unreported σ^R targets (Table 6.2).

The majority of the targets represented previously uncharacterised genes. Interestingly, the RNA polymerase β-subunit gene *rpoB* was identified as a σ^R target. Sequence analysis revealed it possesses a σ^R-dependent promoter upstream of the σ^{HrdB}-dependent promoter (data not shown) and RNAP enrichment suggests it is upregulated upon addition of diamide (Figure 6.4C). Two cold-shock proteins, scoF1 and F40, were identified as σ^R targets which may remove potentially harmful aggregates forms during oxidative stress.

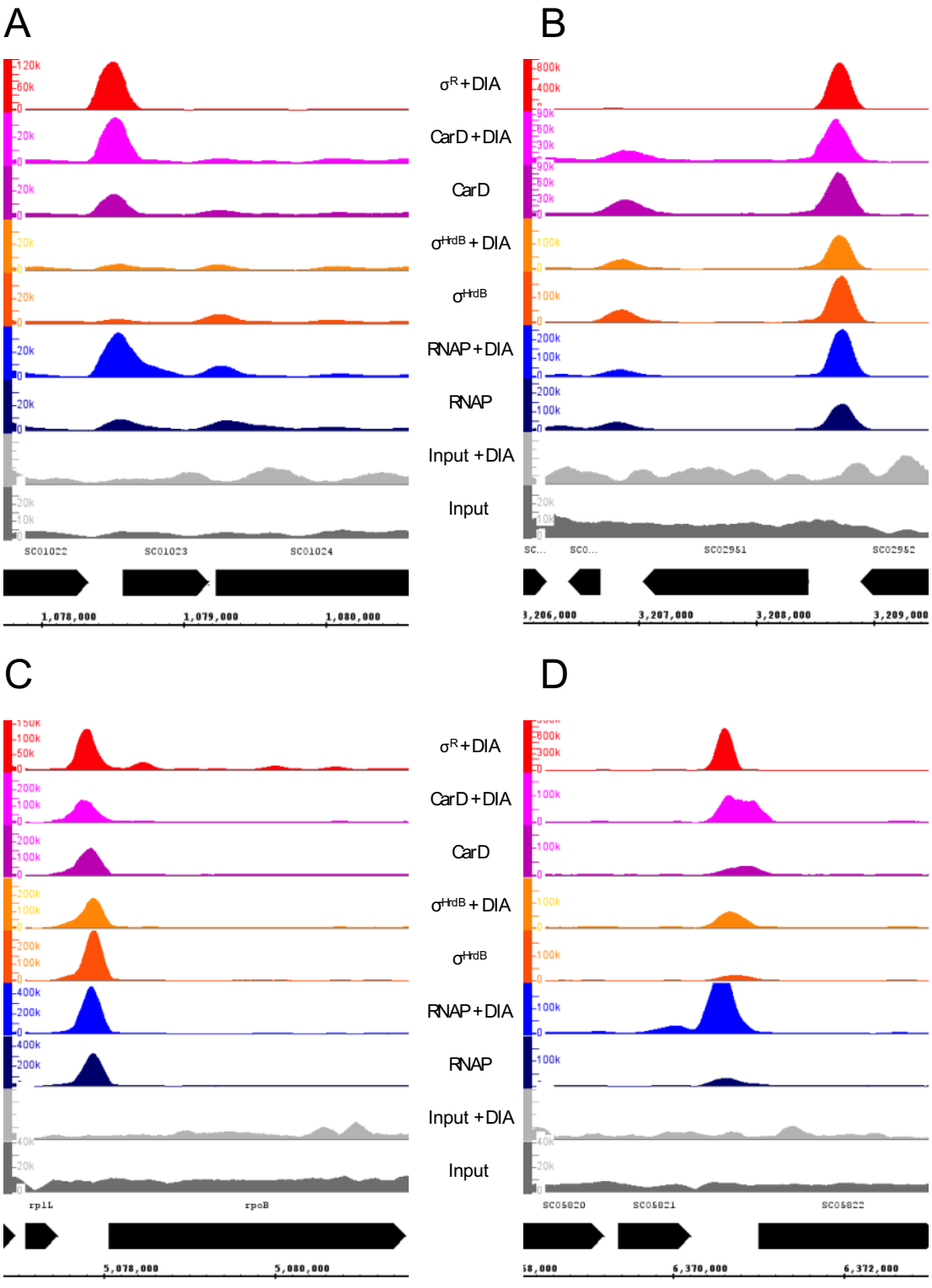


Figure 6.4 - **Validation of novel σ^R targets.** ChIP-seq analysis performed on S202 in the presence and absence of 0.5 mM diamide ("+DIA"). bigWig histograms represent total number of aligned reads, RPKM normalised for sequencing depth. Regions with σ^R targets **(A)** SCO1023 **(B)** SCO2951 **(C)** *rpoB* **(D)** SCO5821a. Input tracks are grey, RNAP are blue, σ^{HrdB} are orange, CarD are pink and σ^R is red.

Gene	Peak summit position	Annotation
SCO0792	838732	Sugar epimerase SCF43.03
SCO1023	1078499	Possible membrane protein SCG20A.03
SCO1556	1668919	Putative acetyltransferase SCL11.12c
SCO2191	2357697	Probable oxidoreductase SC5F7.10
SCO2570	2775451	Putative secreted protein SCC123.08c
SCO2951	3208719	Putative malate oxidoreductase SCE59.10c
SCO3568	3948807	Conserved hypothetical protein SCH17.02c
SCO3731	4105176	Cold shock protein scoF1
SCO3748	4119290	Cold shock protein F40
SCO4654	5077793	DNA-directed RNA polymerase beta subunit, *rpoB*
SCO4934	5368616	Putative lipoprotein SCK13.26
SCO5393	5862698	Putative ABC transporter ATP-binding subunit 2SC6G5.37c
SCO5439	5912247	Conserved hypothetical protein SC6A11.15
SCO5821a	6370622	Hypothetical protein (see below)

Table 6.2 - **Novel σ^R targets identified by ChIP-seq.** A list of genes identified as novel σ^R targets with the summit position identified by MACS2 and gene annotation as listed on StrepDB (http://strepdb.streptomyces.org.uk).

6.2.6 Discovery of SCO5821a, a previously unidentified small ORF

Further analysis of novel σ^R targets revealed a peak with summit position 6370622 bp, located between SCO5821 (serine protease) and SCO5822 (DNA gyrase B) that did not correspond to an annotated open reading frame (Figure 6.4D). Sequence analysis of the region revealed a small previously unidentified 228 bp ORF that encodes a 76 amino acid protein, which was designated SCO5821a.

Analysis of ChIP-seq data for the region surrounding SCO5821a show that the gene is upregulated in response to oxidative stress (Figure 6.4D). RNA-seq data from experiments performed by Romero et al. (2014) further confirm presence of the ORF and indicate that it is expressed under normal growth conditions (Figure 6.5A) and a promoter matching the σ^R consensus sequence (GGAT/C-N16-GTT) was located 41 bp upstream from the start codon (Figure 6.5B). Despite σ^{HrdB} enrichment upstream of the gene, it could not be established whether this represented σ^{HrdB}-dependent expression of SCO5821a or the divergent DNA gyrase B (SCO5822) gene.

Analysis of the protein sequence revealed presence of a CXXC motif that could potentially co-ordinate a metal ion such as zinc (Figure 6.5D). A structural prediction performed with Phyre2 revealed a predicted structure based upon a number of different proteins (Figure 6.5C). The region containing the CXXC motif was predicted as homologous to a known metal-binding rubredoxin protein (PDB: 1DXG) while the region containing the large alpha helix was predicted with high homology to an RNA binding protein (PDB: 3VRH). Sequence homology searches revealed that the protein was conserved in *M. tuberculosis* as well as actinophages including *Streptomyces* and *Mycobacterium* phages Peebs and Gaia (Actinobacteriophage Database) (Figure 6.5D).

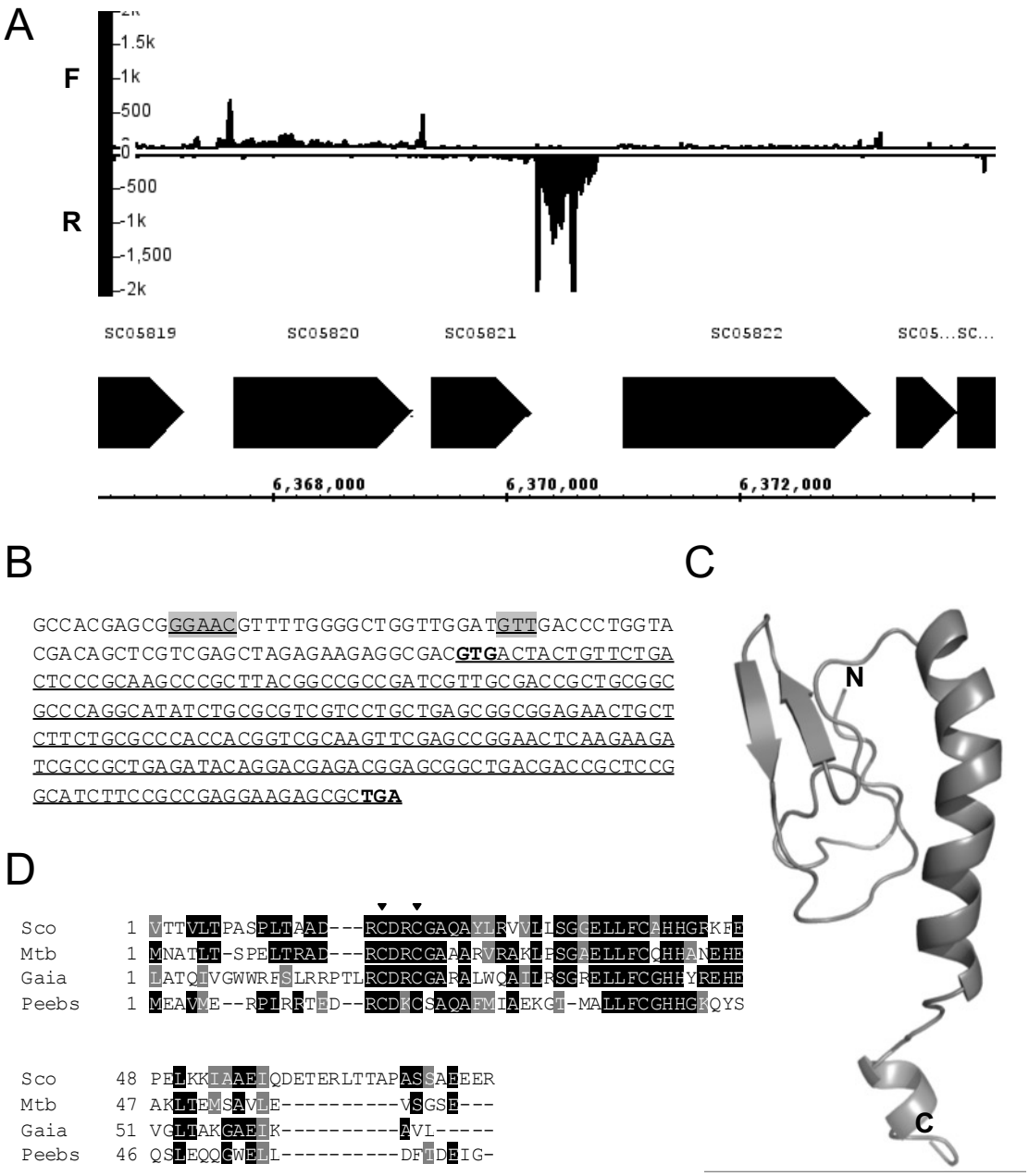


Figure 6.5 - **SCO5821a is a previously unreported ORF.** **(A)** RNA-seq analysis of the region around SCO5821a. RNA-seq data was downloaded from the GEO archive under the accession number GSM1126846 in the form of aligned BEDGRAPH files representing the forward and reverse strands. Data was visualised with Integrated Genome Browser (IGB) (Version 8.3.4) **(B)** Sequence of SCO5821a σ^R-dependent promoter highlighted in grey, SCO5821a ORF underlined with start and stop codons in bold. **(C)** Structural prediction of SCO5821a, prediction performed with Phyre2 and visualised with PyMOL (version 1.3). **(D)** Multiple sequence alignment of *S. coelicolor* SCO5821a (Sco), and hypothetical proteins from *M. tuberculosis* (Mtb), Streptomyces phage Peebs (Peebs) and Mycobacteria phage Gaia (Gaia) performed using CLUSTALW multiple sequence alignment tool (version 2.1). Box shading analysis was performed using BOXSHADE 3.21. Identical amino acid residues are shaded black, similar residues are shaded grey and conserved CXXC motif cysteine residues are indicated with a black triangle.

6.3 CarD does not activate transcription from the σ^R-dependent *trxCp* promoter

ChIP-qPCR and ChIP-seq results suggested that, unlike RbpA, CarD is a component of alternative sigma factor holoenzymes and is present at transcription initiation complexes (section 6.2). CarD was shown to be a transcriptional activator of σ^{HrdB}-dependent promoters. To investigate if CarD activates transcription from alternative sigma factor promoters, *in vitro* transcription experiments were performed using the σ^R-dependent *trxC* promoter region, PCR amplified from M145 gDNA using the primers PTRXC_EXT/PTRXC_INT. The *trxCp* promoter had previously been identified as a σ^R-target by S1 nuclease mapping and microarray analysis of a Δ*sigR* mutant (Paget et al., 2001a; Kallifidas et al., 2010). Using these data and a well conserved σ^R promoter consensus we expected a single transcription product from the *in vitro* transcription template 129 bp in size. To confirm this, *in vitro* transcription reactions were performed with core RNAP in the presence and absence of purified σ^R (a gift from P. Doughty). A transcript from the *trxCp* was only observed in the presence of additional purified σ^R.

To test the effect of CarD, reactions were performed with increasing concentrations of CarD, from CarD:RNAP molar ratios of 0.5:1 to 10:1. As a positive control for function of CarD as a transcriptional activator, reactions were performed simultaneously under the same conditions with core RNAP, purified σ^{HrdB} and the *atpIp in vitro* transcription template (see chapter 4). The inclusion of CarD had no effect on transcription from the *trxCp* promoter at all concentrations (Figure 6.6), whereas CarD increased transcriptional output from the *atpIp* promoter by up to 3-fold. This result suggests that while CarD can bind to alternative holoenzymes it does not activate transcription from alternative promoters.

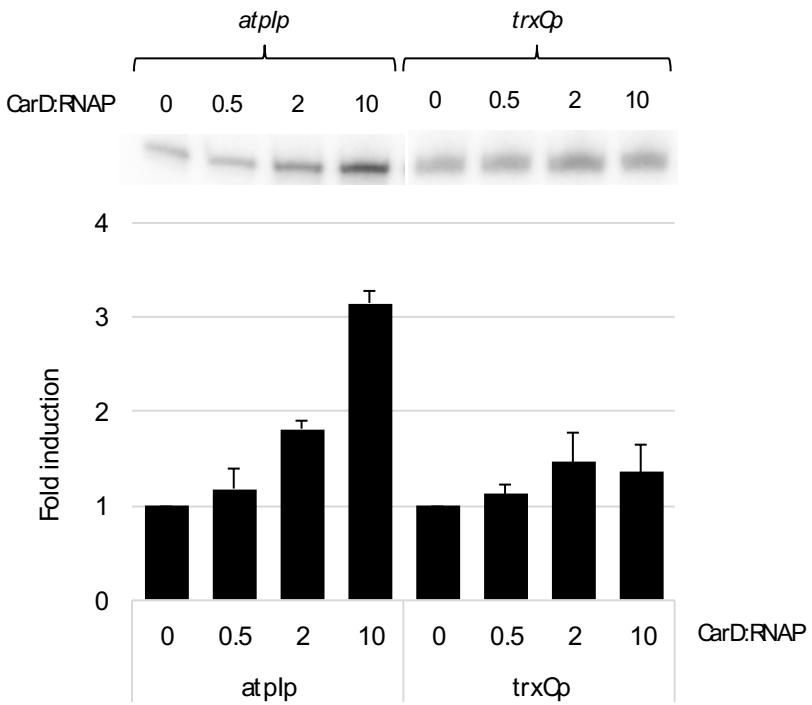


Figure 6.6 - **CarD does not affect transcription from the *trxCp* promoter *in vitro*.** Multi-round i*n vitro* transcription reactions performed on the *atplp* and *trxCp* promoter templates (5 nM) with core RNAP (50 nM), σ^{HrdB} (250 nM) and CarD at CarD:RNAP ratios of 0, 0.5 (25 nM), 2 (100 nM), and 10 (500 nM).

6.4 RbpA is not present at alternative sigma factor promoters

It has been shown that RbpA makes key contacts with σ^{HrdB} which are necessary for its function as a transcriptional activator (Hubin et al., 2015). Additionally, RbpA distribution is highly correlated with σ^{HrdB} and found both proteins are found co-localised at σ^{HrdB}-dependent promoters (see chapter 5). Despite this, it is possible that RbpA makes additional contacts with RNAP. It has previously been proposed that RbpA could bind to the β/β' active-site channel near the rifampicin binding site (Dey et al., 2011) or the β subunit (Hu et al., 2012). However, these binding sites are incompatible with the most recently model of RbpA in complex with with σ_2 of σ^A (see below). A more likely binding site is the β' clamp (M. Paget, personal communication). Therefore, it is conceivable that RbpA could associate with RNAP in the absence of σ^{HrdB}. To test this, ChIP-seq data was used to study the occupancy of RbpA at alternative sigma promoters. For ChIP-seq experiments performed under normal growth conditions, neither σ^{HrdB} nor RbpA was observed at the *hrdB* (SCO5820) or *cwgA* (SCO6179) promoters, known to be controlled by σ^{ShbA} and σ^{E}, respectively (Figure 6.7). Addition of rifampicin, previously shown to

increase RNAP, σ^{HrdB} and RbpA enrichment, increased enrichment of RNAP but had no effect on σ^{HrdB} or RbpA at these promoters. This suggests that RbpA does not associate with alternative holoenzymes.

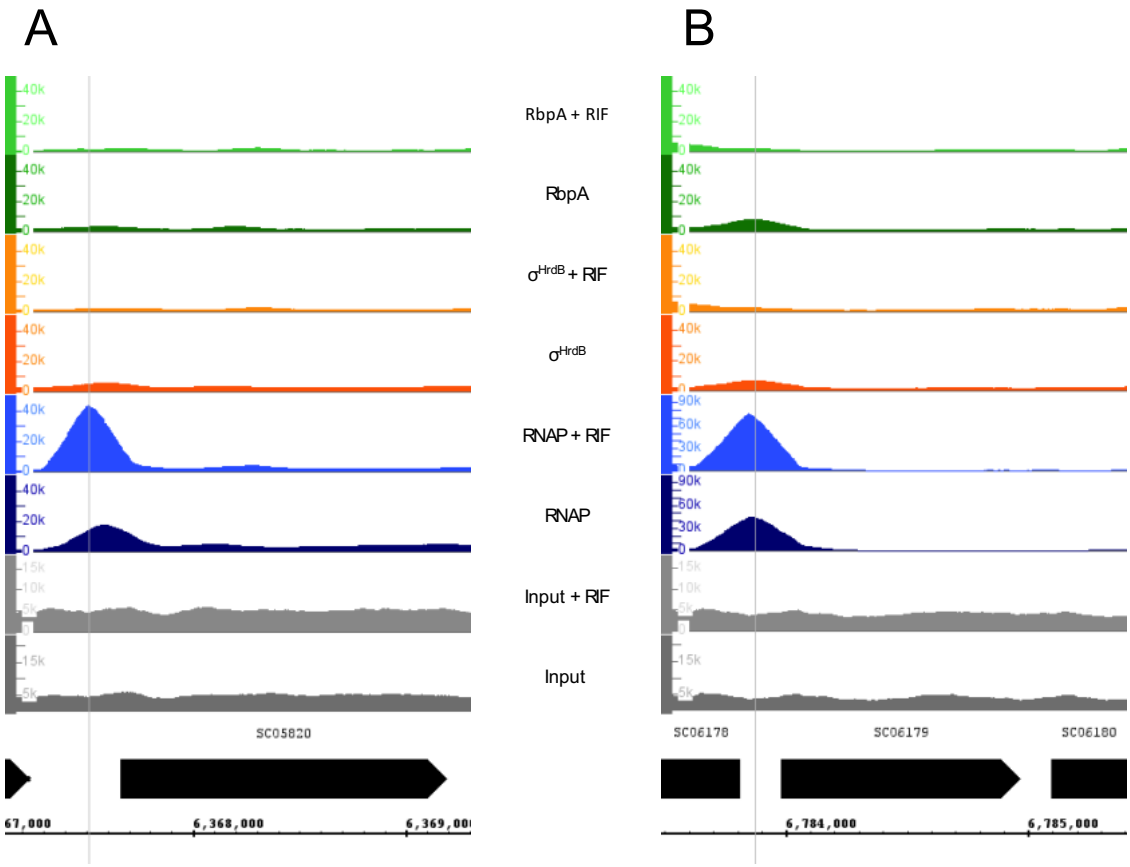


Figure 6.7 - **RbpA is not present at σ^{ShbA} and σ^{E}-dependent promoters.** ChIP-seq analysis performed on S115/pRT802::*rbpA*-3xFLAG. bigWig histograms represent total number of aligned reads, RPKM normalised for sequencing depth. **(A)** *hrdB* (SCO5820) **(B)** *cwgA* (SCO6179). Input tracks are grey, RNAP are blue, σ^{HrdB} are orange and RbpA are green. Predicted transcription start site is indicated with a grey line (B.K. Cho, personal communication).

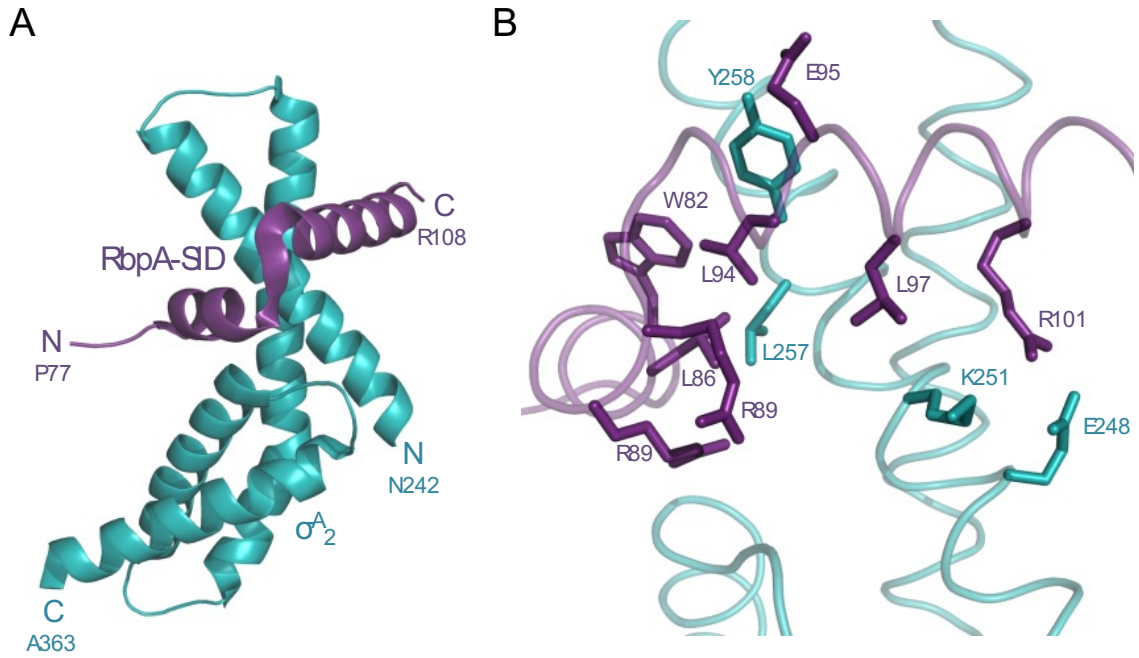
6.5 Identification of residues involved in sigma selectivity of RbpA.

It was previously demonstrated that *S. coelicolor* and *M. tuberculosis* RbpA binds specifically to principal and related sigma factors (Tabib-Salazar et al., 2013). BACTH analysis showed that RbpA binds to *S. coelicolor* σ^{HrdB} (group I) and σ^{HrdA} (group II) but does not interact with sigma factors σ^{HrdC}, σ^{HrdD} (group II) σ^{B}, σ^{WhiG} (group III), σ^{E} or σ^{R} (group IV). Similarly it was shown that *M. tuberculosis* RbpA binds σ^{A} (group I) and σ^{B} (group II) but does not affect σ^{F} (group III) holoenzymes *in vitro* (Hu et al., 2012). Further *in vitro* pull down assays and BACTH analysis on truncated fragments of *S. coelicolor* RbpA identified the C-terminal region as necessary and sufficient for interaction with the σ_2

domain of σ^{HrdB}. The same was also found with *M. tuberculosis* RbpA and σ^{A}, suggesting that the σ_2-RbpA interaction was conserved across the actinobacteria. These data raise the question of how RbpA distinguishes between different sigma factors, including the closely-related group II sigma factors in *S. coelicolor.*

6.5.1 Alignment of conserved regions 1.2-2.3 in σ_2-domains of *M. tuberculosis* σ^{A}, σ^{B}, and σ^{F} and *S. coelicolor* σ^{HrdB}, σ^{HrdA} and σ^{HrdC}

Through work completed in collaboration with Elizabeth Campbell and Seth Darst and colleagues at The Rockefeller University, we have recently obtained a crystal structure of the RbpA sigma interacting domain (RbpA-SID) in complex with σ_2 of σ^{A} (Figure 6.8A and B). From this crystal structure, the σ^{A} residues that potentially contact RbpA could be identified (Figure 6.8B). To probe the sigma specificity of RbpA, the amino acid regions of σ_2 included in the structure were aligned for the RbpA-interacting *M. tuberculosis* σ^{A} and σ^{B}, *S. coelicolor* σ^{HrdB}, σ^{HrdA} and non-interacting *M. tuberculosis* σ^{F} and *S. coelicolor* σ^{HrdC}. Although there is extensive conservation between *M. tuberculosis* σ^{A} and *S. coelicolor* σ^{HrdC} among most of these RbpA interacting residues, there were several positions in σ^{HrdC} that were substituted with physicochemically dissimilar amino acids that might impede RbpA binding.



C

```
SigA Mtb   226 DSVRAYLKQIGKVALLNAEEEVELAKRIEAGLYATQLMTELSERGEKLPAAQRRDMMWIC
SigB Mtb    25 DLVRVYLNGIGKTALLNAAGEVELAKRIEAGLYAEHLLETR----KRLGENRKRDIAAVV
SigF Mtb     1 --MTARAAG-GSASRANEYADVPEMFRELVGLPAG-------------SPEFQR------
HrdB Sco   212 DPVKDYLKQIGKVPLLNAEQEVELAKRIEAGLFAEDKLANS----DKLAPKIKRELEIIA
HrdA Sco    98 DLFRQYLREIGRIPLLSAAEEVDLARRVEAGLFAEEKLRCS----PGLDDRLALDLDRLV
HrdC Sco    37 DLLGQYLTQIGATPLLTAEDEVRLATRIEAGVRAREELETADTGEPAPTPRRRRTLEETV

SigA Mtb   287 RDGDRAKNHLLEANLRLVVSLAKRYTGRGMAFLDLIQEGNLGLIRAVEKFDYTKGYKFST
SigB Mtb    82 RDGEAARRHLLEANLRLVVSLAKRYTGRGMPLLDLIQEGNLGLIRAMEKFDYTKGFKFST
SigF Mtb    39 -----HRDKIVQRCLPLADHLARRFEGRGEPRDDLIQVARVGLVNAAVRFDVKTGSDFVS
HrdB Sco   269 EDGRRAKNHLLEANLRLVVSLAKRYTGRGMLFLDLIQEGNLGLIRAVEKFDYTKGYKFST
HrdA Sco   154 VLGRLAKRRLIEANLRLVVSVAKRYVGRGLTMLDLVQEGNLGLIRAVEKFDYARGYKFST
HrdC Sco    98 HDGQEAKDHMVRANLRLVVSMAKRHAHRGLPLLDVIQEGNLGLIRAVEKFDHTKGFKFST
```

Figure 6.8 - **The X-ray crystal structure of the *M. tuberculosis* σ^A-RbpA complex and prediction of interacting residues**. **(A)** Crystal structure of the *M. tuberculosis* σ^A_2-RbpA-SID complex. Accessed from Protein Data Bank (PDB ID 4X8K) and visualised with PyMOL (version 1.3). RbpA-SID (residues 77-108) is shown in purple, σ^A_2 (residues 242-363) is shown in teal. **(B)** Close up of *M. tuberculosis* σ^A_2-RbpA-SID complex showing σ^A_2 residues selected for mutagenesis and their interacting residues on RbpA. **(C)** Amino acid multiple sequence alignment of conserved σ regions 1.2-2.3 of *M. tuberculosis* σ^A, σ^B and σ^F and *S. coelicolor* σ^{HrdB}, σ^{HrdA} and σ^{HrdC}. Alignment performed using CLUSTALW multiple sequence alignment tool (version 2.1). Box shading analysis was performed using BOXSHADE 3.21. Identical residues are shaded black, similar residues are shaded grey. σ^A residues that contact RbpA indicated with red dots and those selected for mutagenesis outlined in black.

Four of these residues not conserved between *M. tuberculosis* σ^A and *S. coelicolor* σ^{HrdC} identified as possibly essential for σ-selectivity of RbpA were:

- σ^A E248/σ^{HrdC} R59
- σ^A K251/σ^{HrdC} T62
- σ^A L257/σ^{HrdC} V68
- σ^A Y258/σ^{HrdC} R69

To test this, site-specific mutagenesis was performed to introduce these substitutions into σ^A and the ability of RbpA to bind and activate transcription was assessed.

6.5.2 Purification of σ^A and σ^A mutants

To overexpress and purify *M. tuberculosis* σ^A from *E. coli,* the *sigA* ORF (Rv2703) was amplified from *M. tuberculosis* H37Rv genomic DNA by PCR using primers SigA_F/SigA_R, designed to introduce an NdeI site overlapping the start codon and a BglII site downstream from the stop codon respectively. The PCR product was cloned into EcoRV-cut pBlueScript SKII+, sequenced to confirm amplification with no mutations and subcloned into pET15b as an NdeI/BglII fragment, producing the plasmid pET15b-sigA.

To introduce the mutations into *sigA*, inverse PCR for site directed mutagenesis was performed. Two mutant σ^A proteins were designed with L257V and Y258R mutations, and E248R, K251T, L257V and Y258R mutations named $\sigma^{A(VR)}$ and $\sigma^{A(RTVR)}$, respectively. Using pBlueScript-sigA as a template, the primer sets SigA-VRA_F/SigA-VRA_R and SigA-double_F/SigA-double_R created $\sigma^{A(VR)}$ and $\sigma^{A(RTVR)}$ mutants, respectively, and following sequencing to confirm presence of the expected mutations, the σ^A mutant alleles were subcloned into pET15b as NdeI/BglII fragments.

The recombinant plasmids pET15b-sigA, pET15b-sigA$^{(VR)}$ and pET15b-sigA$^{(RTVR)}$ were used to transform *E. coli* BL21 (pLysS) and a single colony used to inoculate 500 mL LB. The culture was grown in an orbital shaker at 37 °C to OD_{600} 0.5 - 0.6, the flask was submerged in ice-water for 10 minutes before addition of 1 mM IPTG and returning the culture to the orbital shaker for 3 hours at 30 °C. Cells were harvested by centrifugation, resuspended in 15 mL binding buffer + 1.5 mL protease inhibitor cocktail + 25 ug/mL PMSF and disrupted by sonication (6 x 10s @ 35% ampl.). The cleared cell lysate (CCL)

was separated from the cell debris by centrifugation at 12,000 x g for 15 minutes at 4 °C. To purify 6His-tagged σ^A, the CCL was loaded onto a Ni-NTA sepharose column and washed with 10 C.V. of 10 mM imidazole wash buffer followed by 20 C.V of 25 mM imidazole wash buffer.

To cleave the 6His-tag and elute the untagged protein, the column was incubated with 10 NIH units thrombin in 4 mL thrombin cleavage buffer for 5 hours at 4 °C. Samples were analysed at this point by SDS-PAGE (Figure 6.9A). Following digestion, approximately 4 mL of eluate was further purified by size-exclusion chromatography (HiLoad 16/60Superdex 200; GE Healthcare) (Figure 6.9B). Elutions were collected as 0.5 mL fractions, analysed by SDS-PAGE (Figure 6.9C) and those containing σ^A were combined, concentrated to ~5 µM by centrifugal filtration (VivaSpin 6, 3000 MWCO, Sartorius), analysed by SDS-PAGE (Figure 6.9D) and flash frozen in liquid nitrogen and stored at -80 °C.

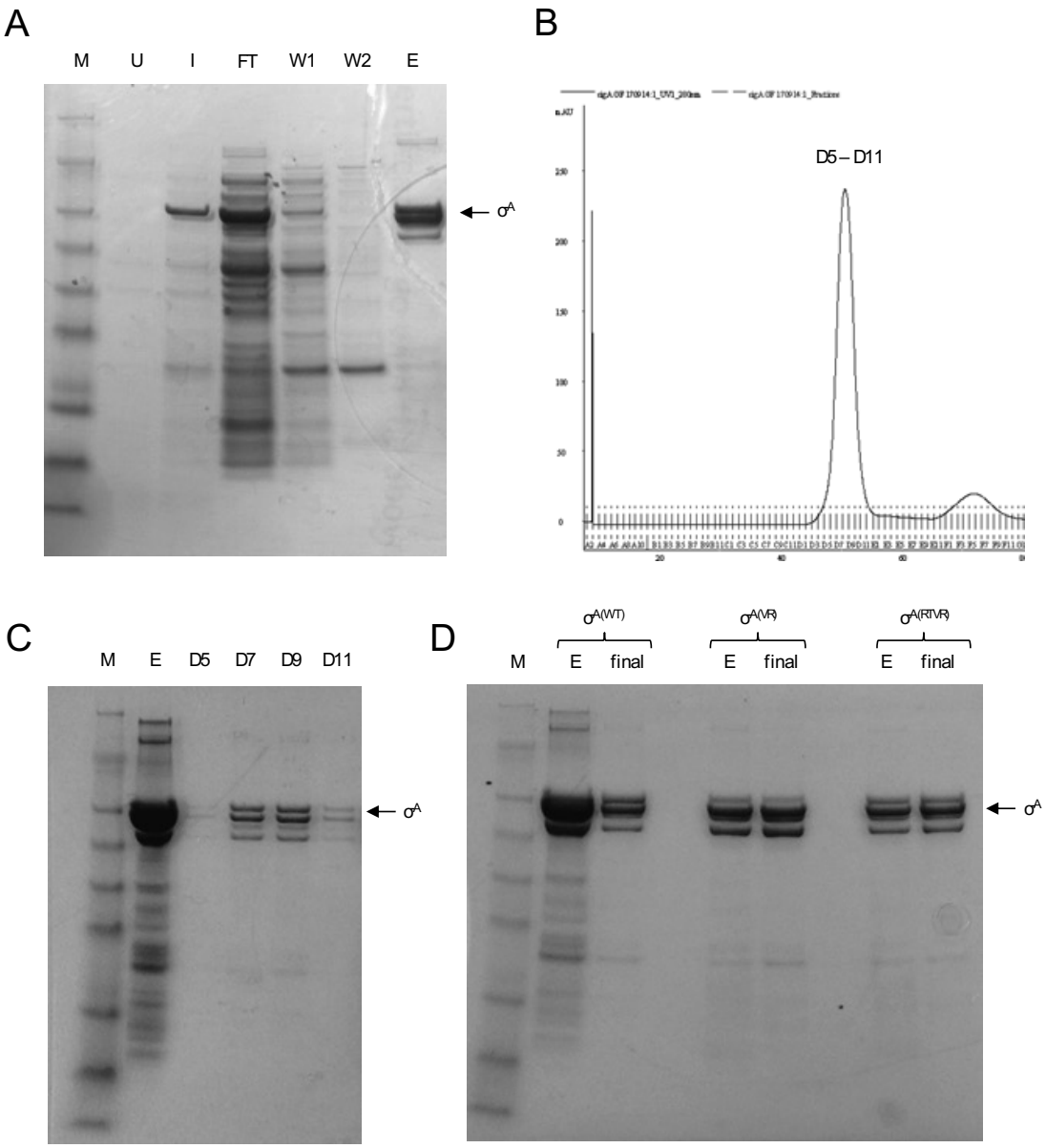


Figure 6.9 - **Purification of σ^A and σ^A mutants.** **(A)** SDS-PAGE analysis of σ^A Ni-NTA purification. Lane 1 (M): SeeBlue Plus2 marker; lane 2 (U): uninduced culture; lane 3 (I): IPTG induced culture; lane 4 (FT): Ni-NTA flowthrough; lane 5 (W1): Ni-NTA wash 1; lane 6 (W2): Ni-NTA wash 2; lane 7 (E): thrombin-cleavage elution. **(B)** Size exclusion purification of σ^A. HiLoad™ 16/60 Superdex™ 200 column elution trace for purification of σ^A **(C)** SDS-PAGE analysis of σ^A size exclusion purification. Lane 1 (M): SeeBlue Plus2 marker; lane 2 (E): thrombin-cleaved elution; lanes 3-6 (D5) size exclusion elution fractions D5-D11. **(D)** SDS-PAGE analysis of σ^A, $\sigma^{A(VR)}$ and $\sigma^{A(RTVR)}$ final purification products. Lane 1 (M): SeeBlue Plus2 marker; lane 2 (σ^A E): thrombin-cleavage σ^A elution; lane 3 (σ^A final): σ^A final concentrated purification product; lane 5 ($\sigma^{A(VR)}$ E): thrombin-cleavage $\sigma^{A(VR)}$ elution; lane 6 ($\sigma^{A(VR)}$ final): $\sigma^{A(VR)}$ final concentrated purification product; lane 8 ($\sigma^{A(RTVR)}$ E): thrombin-cleavage $\sigma^{A(RTVR)}$ elution; lane 9 ($\sigma^{A(RTVR)}$ final): $\sigma^{A(RTVR)}$ final concentrated purification product.

6.5.3 $\sigma^{A(VR)}$ and $\sigma^{A(RTVR)}$ mutants are functional but unresponsive to RbpA

Multi-round i*n vitro* transcription reactions performed on the *vapB10L* promoter template with *M. bovis* RNAP (50 nM) (a gift from E. Campbell); σ^{A} (250 nM), $\sigma^{A(VR)}$ (250 nM) or $\sigma^{A(RTVR)}$ (250 nM) in the presence and absence of excess *M. tuberculosis* RbpA (500 nM) (a gift from A. Tabib-Salazar). In the absence of RbpA, $\sigma^{A(VR)}$ and $\sigma^{A(RTVR)}$ performed similarly to σ^{A} in directing transcription from the *vapB10L* promoter (Figure 6.10). However, unlike wild-type σ^{A} which was activated up to 3-fold, neither $\sigma^{A(VR)}$ nor $\sigma^{A(RTVR)}$ were affected by the addition of RbpA to the reaction mixture. This confirms that the interaction between σ and RbpA is essential for the role of RbpA as a transcriptional activator.

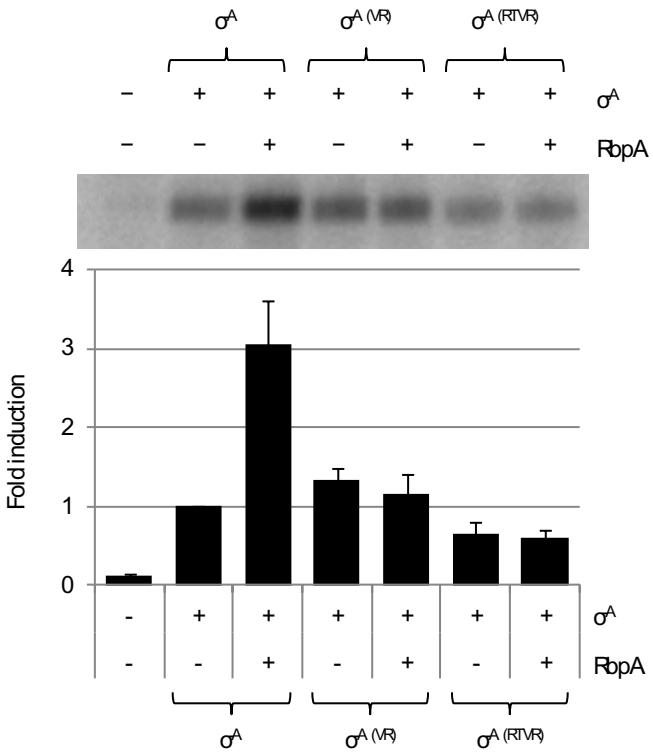


Figure 6.10 - **$\sigma^{A(VR)}$ and $\sigma^{A(RTVR)}$ mutants are functional but unresponsive to RbpA.** Multi-round i*n vitro* transcription reactions performed on the *vapB10L* promoter template (5 nM) with *M. bovis* RNAP (50 nM); σ^{A} (250 nM), $\sigma^{A(VR)}$ (250 nM) or $\sigma^{A(RTVR)}$ (250 nM) in the presence and absence of excess RbpA (500 nM). Data are presented as fold-difference relative to data obtained with σ^{A} in the absence of RbpA. Transcript levels were quantified by phosphorimaging from triplicate data, and standard deviation is indicated.

6.6 Discussion

6.6.1 CarD is present on alternative holoenzymes

Structural studies have shown that CarD binds to the β1 lobe of the β subunit of RNAP (Srivastava et al., 2013). Consistent with predictions based on the amino acid sequence, the CarD RNAP interaction domain (RID) strongly resembles the TRCF-RID in structure as well as sequence. Therefore using a crystal structure of TRCF in complex with the β1-lobe, Srivastava et al. (2013) were able to model CarD onto the RNAP open complex structure. Bound to the β1 lobe, the model positioned CarD on the opposite face of RNAP to the σ subunit. Based on this it seemed unlikely for CarD to interact with σ.

ChIP-seq experiments performed under normal growth conditions in *M. tuberculosis* showed that CarD distribution highly correlates with σ^{A} (Srivastava et al., 2013). Similar experiments performed in *S. coelicolor* showed that CarD localised with σ^{HrdB} throughout the genome under normal growth conditions. Under these conditions, it was only possible to conclusively identify the presence of CarD and absence of σ^{HrdB} at one alternative sigma promoter, σ^{E}-dependent *cwgA*. During oxidative stress caused by addition of diamide to cultures of *S. coelicolor*, expression of the σ^{R} regulon is induced. As part of this response, ChIP-seq experiments revealed that CarD co-localises with RNAP and σ^{R} at σ^{R}-dependent promoters (Figure 6.2). The absence of σ^{HrdB} at these loci reveals that CarD must be bound to σ^{R} holoenzymes. This is the first time it has been reported that CarD binding is not σ factor specific.

6.6.2 CarD does not activate transcription from σ^{R}-dependent *trxCp* promoter

CarD activates transcription from σ^{HrdB}-dependent promoters *in vitro*. The discovery that CarD binding is not σ specific and is present on alternative holoenzymes raised the possibility that CarD could also activate transcription from alternative promoters. However, the addition of CarD at CarD:RNAP ratios of up to 10:1 has no effect on the σ^{R}-dependent *trxC* promoter *in vitro*. While the activity of CarD at other alternative sigma factor dependent promoters has not been tested, it seems possible that CarD specifically activates promoters dependent on the principal sigma factor. The reason for this is not clear, but may be due to differences in promoter sequence and/or mechanism of transcription initiation.

Following the structural elucidation of CarD and modelling onto the *T. thermophilus* RNAP open complex structure, it was revealed that the CarD-CTD was positioned to interact directly with the promoter DNA just upstream of the -10 element (Srivastava et al., 2013). Structural studies performed on the *E. coli* ECF sigma factor σ^E revealed a number of differences in the mechanism of -10 element recognition and open complex formation compared to housekeeping sigma factors (Campagne et al., 2014). Biochemical studies showed that similar differences were observed in a number of other ECF sigma factors included *S. coelicolor* σ^R. Notably, whilst two nucleotides are flipped from the base pair stack at positions -11 and -7 in recognition of housekeeping sigma factors, a single nucleotide at position -10 is flipped into a binding pocket on σ_2 of σ^E and other ECF sigma factors (Feklistov and Darst, 2011; Campagne et al., 2014). It is possible that such differences in structural recognition and bubble formation in ECF sigma factors are incompatible with the mechanism through which CarD activates transcription by interacting with DNA upstream edge of the -10 element.

ECF sigma factors also differ in the absence of $\sigma_{1.1}$ and σ_3. During initiation, $\sigma_{1.1}$ sits in the the active-site cleft downstream of the transcription start site and must be displaced during open complex formation (Vuthoori et al., 2001). It is possible that CarD is important in suitably stabilising promoters to allow displacement of $\sigma_{1.1}$ and therefore has no effect on transcription from ECF sigma promoters.

6.6.3 Heterologous expression and purification of σ^A and σ^A mutants

This chapter demonstrated a method for heterologous expression and purification of *M. tuberculosis* σ^A from *E. coli* without the presence of additional affinity tags. Following Ni affinity, His-tag cleavage, and size-exclusion chromatography, wild-type σ^A and σ^A mutants were successfully purified. Purification profiles for σ^A, $\sigma^{A(VR)}$ and $\sigma^{A(RTVR)}$ were similar indicating that presence of point mutants didn't affect protein expression or solubility. SDS-PAGE analysis of all three proteins revealed the presence of a number of smaller bands, potentially σ^A degradation products (Figure 6.9). Whilst the presence of these products could not be prevented or explained, they did not appear to affect σ^A, $\sigma^{A(VR)}$, or $\sigma^{A(RTVR)}$ performance *in vitro*.

6.6.4 Identification of residues involved in sigma selectivity of RbpA

It had previously been identified in BACTH experiments that RbpA interacts with *S. coelicolor* σ^{HrdB} and σ^{HrdA}, and *M. tuberculosis* σ^{A} and σ^{B} (Tabib-Salazar et al., 2013; Hu et al., 2012). Additionally, these experiments identified a number σ factors in both organisms that would not bind to RbpA. Multiple sequence alignment of these interacting and non-interacting proteins, in combination with predicted interacting residues based on a structure of RbpA in complex with σ^{A}, revealed substitutions that might determine RbpA sigma specificity. An ELAK/LYA motif, conserved throughout RbpA-binding σ factors but absent in σ^{HrdC} and σ^{F} was identified as a region for site-directed mutagenesis. In σ^{HrdC}, an RLAT/VRA motif is present at the corresponding position. Substitution of VRA or RLAT/VRA motif into σ^{A} created the point mutants $\sigma^{A(VR)}$ and $\sigma^{A(RTVR)}$. *In vitro* transcription experiments showed that while the σ^{A} point mutants were still functional in directing transcription, addition of RbpA to the reaction had no effect. Published BACTH experiments on σ^{A}, $\sigma^{A(RT)}$, $\sigma^{A(VR)}$ and $\sigma^{A(RTVR)}$ showed similar results, with only wild-type σ^{A} showing interaction with RbpA (Hubin et al., 2015). Together these results served as biochemical evidence to complement the structure of RbpA in complex with σ^{A}.

Chapter 7:

General discussion

7 General discussion

7.1 Summary of findings

The aims of this study were to investigate global regulators of transcription initiation in *S. coelicolor*. This study added to the body of knowledge about the function of three RNAP-binding proteins in *S. coelicolor*: DksA, CarD and RbpA.

DksA is a protein implicated in the function of ppGpp and the stringent response in *E. coli*. Recently it has been proposed that the mechanism of action of ppGpp in a number of organisms including *B. subtilis* differs from that in *E. coli*. In these organisms, instead of directly binding RNAP, ppGpp regulates intracellular GTP concentration, the initiating NTP at stringently regulated promoters (Kriel et al., 2012; Liu et al., 2015). Furthermore it was proposed that this was consistent with the absence of a DksA homologue in such organisms (Liu et al., 2015). The situation in *S. coelicolor* was unclear because although *S. coelicolor* appears to lack a ppGpp binding site on RNAP, a DksA homologue was identified. However, it was demonstrated here that deletion of *dksA* had no visible effect on growth and antibiotic production, in contrast to the known effect of deleting the ppGpp synthase *relA*. Nonetheless, overexpression of *dksA* induced the overproduction of the blue-pigmented antibiotic actinorhodin in a manner that appears to require binding to the RNA polymerase secondary channel. These studies suggest a role for DksA that does not involve ppGpp, although the mechanism for actinorhodin overproduction and biological function of DksA in *S. coelicolor* remains unknown.

This study also initiated investigations into the function of CarD in *S. coelicolor.* CarD is essential in *M. smegmatis* and *M. tuberculosis* (Stallings et al., 2009) and was shown here to be essential for growth in *S. coelicolor*. *In vitro* transcription experiments indicated that CarD activates housekeeping promoters dependent on σ^{HrdB}. CarD ChIP-seq experiments demonstrated that it is found exclusively at promoter regions. Interestingly, despite CarD only activating transcription from σ^{HrdB}-dependent promoters *in vitro*, it associates with RNAP containing alternative sigma factors, indicating that it does not display holoenzyme specificity.

This study also detailed ChIP-seq experiments performed on RbpA, an RNAP binding transcriptional activator exclusive to the Actinobacteria. RbpA binds specifically to σ^{HrdB}

and co-localised with σ^{HrdB} throughout the genome. Thus, σ^{HrdB} co-localises with both CarD and RbpA suggesting that all three factors are present at transcription initiation complexes. Furthermore, structural modelling suggests that both proteins are able to bind RNAP simultaneously, each interacting with DNA upstream of the -10 element. *In vitro* transcription analysis of natural and synthetic promoters suggested that RbpA and CarD can overcome the absence of a consensus -35 element, a common occurrence in actinobacteria. Together, these data present a convincing model for an overlapping role of CarD and RbpA in *S. coelicolor.*

7.2 General discussion

Historically, *E. coli* has served as the model organism for studying prokaryotic transcription, owing to its amenability to genetic analysis, protein purification, and this has recently been bolstered by recent crystal structures of RNAP transcription initiation complexes. However, while transcription initiation is fundamentally conserved across all bacteria, recent studies in a range of organisms are revealing key mechanistic and regulatory differences.

Understanding transcription initiation in diverse bacterial species comes with far reaching implications. For instance, mycobacterial RNAP is a key target for the first line drug rifampicin. However, the emergence and spread of rifampicin resistance worldwide has potentially dire consequences for human health. Further research into transcription in actinobacteria will be necessary to fully exploit this proven anti-mycobacterial target through the discovery of novel compounds. Understanding of transcriptional regulation in diverse bacterial species also has implications within the growing field of synthetic biology. For example, to use an organism like *S. coelicolor* as a production chassis, or to express *S. coelicolor* genes or systems heterologously will require full understanding of gene transcription to ensure that a controlled, engineered approach can be taken.

A key finding in this study is that RbpA and CarD are present at all vegetative promoters dependent on σ^{HrdB}. RbpA is confined to the actinobacteria, while CarD is more widespread, present in diverse phyla such as Firmicutes and α-proteobacteria but notably absent from all γ-proteobacterial species including *E. coli.* CarD has been shown as an essential protein in both mycobacteria and *S. coelicolor*. It was previously thought that

RbpA is not an essential protein although recent findings have shown that *S. coelicolor* possesses a paralogue, RbpB, which shares a functional redundancy with RbpA (M. Paget, personal communication). The basis for the essentiality of both CarD and RbpA remains unknown. Whilst there appears to be striking similarities between the two proteins and potentially an overlap in function, there is no genetic redundancy suggesting they have specific essential roles.

Whilst it has been shown that both proteins activate transcription from σ^{HrdB}-dependent promoters *in vitro*, much is still unknown about both proteins *in vivo.* The cellular stoichiometry of CarD, RbpA and σ^{HrdB} could be important to the role of both proteins. Evidence suggests that RbpA is a major component of RNAP, following its discovery as a major band in RNAP preparations (Paget et al., 2001a). One way the stoichiometry could affect activity is whether both proteins are present on the same transcription complexes. Modelling suggests that both proteins may be able to simultaneously bind RNAP although this is something that would need to be confirmed by crosslinking/pulldown studies. It is possible that different outputs on transcription could occur if one or both proteins are present. In this study, an additive effect of both proteins was observed at saturating concentrations.

Present on all at all housekeeping promoters and able to activate or potentially inhibit transcription, CarD and RbpA are well placed to orchestrate major global transcriptional changes. For example, for streptomycetes to enter stationary phase requires major reprogramming of the transcriptome (Nieselt et al., 2010). The basis of this change is not yet understood but it is possible that global transcription factors such as CarD and RbpA could be involved in such widespread changes. This could either be for example through regulation of CarD/RbpA levels or post-translational modifications. Interestingly differences have already been observed in how transcription of the two genes is regulated. Diamide has been shown to have opposing effects on transcription of these two proteins, repressing transcription of *carD* but activating transcription of RbpA (Kallifidas et al., 2010). Post-translational modification might involve acetylation which is an emerging phenomenon in bacteria and occurs extensively in *M. tuberculosis* (Liu et al., 2014). A potential target for acetylation of RbpA is the basic linker that contains three lysine residues. Reversible acetylation of these lysine residues would be expected to

influence the interaction between RbpA and the -10 upstream region, which might in turn affect promoter activity at RbpA-controlled promoters. Another unexplored mechanism is the possibility that CarD and RbpA have unidentified binding partners that are able to control their activity.

7.3 Future directions

One of the most striking results obtained in this study was *in vitro* transcription data showing that CarD and RbpA are both able to inhibit transcription from specifically designed promoters possessing a conserved -35 region. Through screening of additional promoters by *in vitro* transcription, it may be possible to identify naturally occurring promoters that are similarly affected by CarD or RbpA. Characterisation of these promoters *in vivo* would give novel insights into the function of both proteins.

As an essential gene, performing *in vivo* studies on the function of CarD presents a number of difficulties, primarily the inability to obtain and characterise a stable deletion mutant. As an alternative approach, this study has succeeded in creating two strains suitable for depletion of CarD. To further understand the importance of CarD in regulation of transcription, it would be of interest to perform transcriptomic analysis on strains depleted of CarD. Whilst ChIP-seq experiments have identified CarD is present at all promoters *in vivo*, comparisons in gene expression between a wild-type strain and one depleted of CarD may identify differences in role of CarD at individual promoters. Additionally, this study has demonstrated the power of studying distribution of RNAP by ChIP-seq. It might therefore be advantageous to perform such “RNA polymerase-omics” experiments over transcriptomic techniques such as RNA-seq.

Chapter 8:

Bibliography

8 Bibliography

Aldridge, M., Facey, P., Francis, L., et al. (2013) A novel bifunctional histone protein in Streptomyces: A candidate for structural coupling between DNA conformation and transcription during development and stress? Nucleic Acids Research, 41 (9): 4813–4824

Alting-Mees, M.A. and Short, J.M. (1989) pBluescript II: gene mapping vectors. Nucleic acids research, 17 (22): 9494

Angelis, K.J. (1986) Glover, D. M., (ed.): DNA Cloning. A Practical Approach, Vol. I and II. Biologia Plantarum, 28 (5): 395–395

Artsimovitch, I., Patlan, V., Sekine, S.I., et al. (2004) Structural basis for transcription regulation by alarmone ppGpp. Cell, 117 (3): 299–310

Artsimovitch, I., Svetlov, V., Murakami, K.S., et al. (2003) Co-overexpression of Escherichia coli RNA polymerase subunits allows isolation and analysis of mutant enzymes lacking lineage-specific sequence insertions. The Journal of biological chemistry, 278 (14): 12344–55

Babcock, M.J., Buttner, M.J., Keler, C.H., 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): 31–42

Bailey, T., Krajewski, P., Ladunga, I., et al. (2013) Practical guidelines for the comprehensive analysis of ChIP-seq data. PLoS computational biology, 9 (11): e1003326

Banta, A.B., Cuff, M.E., Lin, H., et al. (2014) Structure of the RNA polymerase assembly factor Crl and identification of its interaction surface with sigma S. Journal of Bacteriology, 196 (18): 3279–88

Barker, M.M., Gaal, T. and Gourse, R.L. (2001a) Mechanism of regulation of transcription initiation by ppGpp. II. Models for positive control based on properties of RNAP mutants and competition for RNAP. Journal of molecular biology, 305 (4): 689–702

Barker, M.M., Gaal, T., Josaitis, C.A., et al. (2001b) 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): 673–688

Barne, K.A., Bown, J.A., Busby, S.J., et al. (1997) Region 2.5 of the Escherichia coli RNA polymerase sigma70 subunit is responsible for the recognition of the “extended-10” motif at promoters. The EMBO journal, 16 (13): 4034–40

Basu, R.S., Warner, B.A., Molodtsov, V., et al. (2014) Structural Basis of Transcription Initiation by Bacterial RNA Polymerase Holoenzyme. Journal of Biological Chemistry, 289 (35): 24549–24559

Battesti, A., Majdalani, N. and Gottesman, S. (2011) The RpoS-mediated general stress response in Escherichia coli. Annual review of microbiology, 65: 189–213

Baylis, H.A. and Bibb, M.J. (1988) Transcriptional analysis of the 16S rRNA gene of the rrnD gene set of Streptomyces coelicolor A3(2). Molecular Microbiology, 2 (5): 569–579

Becker, G. and Hengge-Aronis, R. (2001) What makes an Escherichia coli promoter sigma(S) dependent? Role of the -13/-14 nucleotide promoter positions and region 2.5 of sigma(S). Molecular Microbiology, 39 (5): 1153–1165

Bentley, S.D., Chater, K.F., Cerdeño-Tárraga, A.-M., et al. (2002) Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2). Nature, 417 (6885): 141–7

Bibb, M.J., Domonkos, A., Chandra, G., 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): 1033–49

Bibb, M.J., Janssen, G.R. and Ward, J.M. (1986) Cloning and analysis of the promoter region of the erythromycin-resistance gene (ermE) of Streptomyces erythraeus. Gene, 41 (2-3): E357–E368

Bibb, M.J., Molle, V. and Buttner, M.J. (2000) sigma(BldN), an extracytoplasmic function RNA polymerase sigma factor required for aerial mycelium formation in Streptomyces coelicolor A3(2). Journal of bacteriology, 182 (16): 4606–16

Bierman, M., Logan, R., O'Brien, K., et al. (1992) Plasmid cloning vectors for the conjugal transfer of DNA from Escherichia coli to Streptomyces spp. Gene, 116 (1): 43–49

Bloch, H. (1953) Acid-Fast Bacteria. Annual Review of Microbiology, 7 (1): 19–46

Bougdour, A., Lelong, C. and Geiselmann, J. (2004) Crl, a low temperature-induced protein in Escherichia coli that binds directly to the stationary phase sigma subunit of RNA polymerase. The Journal of biological chemistry, 279 (19): 19540–50

Brown, K.L., Wood, S. and Buttner, M.J. (1992) Isolation and characterization of the major vegetative RNA polymerase of Streptomyces coelicolor A3(2); renaturation of a sigma subunit using GroEL. Molecular microbiology, 6 (9): 1133–9

Buck, M., Gallegos, M.T., Studholme, D.J., et al. (2000) The bacterial enhancer-dependent σ54 (σ(N)) transcription factor. Journal of Bacteriology. 182 (15) pp. 4129–4136

Burgess, R.R., Travers, A.A., Dunn, J.J., et al. (1969) Factor stimulating transcription by RNA polymerase. Nature, 221 (5175): 43–6

Bush, M.J., Bibb, M.J., Chandra, G., et al. (2013) Genes required for aerial growth, cell division, and chromosome segregation are targets of WhiA before sporulation in Streptomyces venezuelae. mBio, 4 (5): e00684–13

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): 3367–78

Buttner, M.J. and Lewis, C.G. (1992) Construction and characterization of Streptomyces coelicolor A3(2) mutants that are multiply deficient in the nonessential hrd-encoded RNA polymerase sigma factors. Journal of Bacteriology, 174 (15): 5165–5167

Campagne, S., Damberger, F.F., Kaczmarczyk, A., et al. (2012) Structural basis for sigma factor mimicry in the general stress response of Alphaproteobacteria. Proceedings of the National Academy of Sciences of the United States of America, 109 (21): E1405–14

Campagne, S., Marsh, M.E., Capitani, G., et al. (2014) Structural basis for -10 promoter element melting by environmentally induced sigma factors. Nature structural & molecular biology, 21 (3): 269–76

Campbell, E.A., Greenwell, R., Anthony, J.R., et al. (2007) A conserved structural module regulates transcriptional responses to diverse stress signals in bacteria. Molecular cell, 27

(5): 793–805

Campbell, E.A., Korzheva, N., Mustaev, A., et al. (2001) Structural Mechanism for Rifampicin Inhibition of Bacterial RNA Polymerase. Cell, 104 (6): 901–912

Campbell, E.A., Muzzin, O., Chlenov, M., et al. (2002) Structure of the bacterial RNA polymerase promoter specificity sigma subunit. Molecular cell, 9 (3): 527–39

Campbell, E.A., Tupy, J.L., Gruber, T.M., et al. (2003) Crystal Structure of Escherichia coli σE with the Cytoplasmic Domain of Its Anti-σ RseA. Molecular Cell, 11 (4): 1067–1078

Capstick, D.S., Willey, J.M., Buttner, M.J., et al. (2007) SapB and the chaplins: connections between morphogenetic proteins in Streptomyces coelicolor. Molecular Microbiology, 64 (3): 602–13

Cashel, M. and Gallant, J. (1969) Two compounds implicated in the function of the RC gene of Escherichia coli. Nature, 221 (5183): 838–41

Cashel, M. and Kalbacher, B. (1970) The control of ribonucleic acid synthesis in Escherichia coli. V. Characterization of a nucleotide associated with the stringent response. Journal of Biological Chemistry, 245 (9): 2309–2318

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): 5854–5861

Challis, G.L. (2008) Genome mining for novel natural product discovery. Journal of medicinal chemistry, 51 (9): 2618–28

Chater, K.F., Biró, S., Lee, K.J., et al. (2010) The complex extracellular biology of Streptomyces. FEMS microbiology reviews, 34 (2): 171–98

Chater, K.F., Hopwood, D.A., Kieser, T., et al. (1982) Gene cloning in Streptomyces. Current topics in microbiology and immunology, 96: 69–95

Chen, C.Y. and Richardson, J.P. (1987) Sequence elements essential for rho-dependent transcription termination at lambda tR1. The Journal of biological chemistry, 262 (23): 11292–9

Choy, J., Aung, L. and Karzai, A. (2007) Lon protease degrades transfer-messenger RNA-tagged proteins. Journal of Bacteriology

Cole, S.T., Brosch, R., Parkhill, J., et al. (1998) Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. Nature, 393 (6685): 537–44

Cramer, P., Bushnell, D.A., Fu, J., et al. (2000) Architecture of RNA polymerase II and implications for the transcription mechanism. Science (New York, N.Y.), 288 (5466): 640–9

Czyz, A., Mooney, R.A., Iaconi, A., et al. (2014) Mycobacterial RNA polymerase requires a U-tract at intrinsic terminators and is aided by NusG at suboptimal terminators. mBio, 5 (2): e00931

Davis, C.A., Bingman, C.A., Landick, R., 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): 7833–8

Davis, C.A., Capp, M.W., Record, M.T., et al. (2005) The effects of upstream DNA on open complex formation by Escherichia coli RNA polymerase. Proceedings of the National Academy of Sciences of the United States of America, 102 (2): 285–90

Davis, E., Chen, J., Leon, K., et al. (2015) Mycobacterial RNA polymerase forms unstable open promoter complexes that are stabilized by CarD. Nucleic acids research, 43 (1): 433–45

Dey, A., Verma, A.K. and Chatterji, D. (2011) Molecular insights into the mechanism of phenotypic tolerance to rifampicin conferred on mycobacterial RNA polymerase by MsRbpA. Microbiology (Reading, England), 157 (Pt 7): 2056–71

Dombroski, A.J., Walter, W.A. and Gross, C.A. (1993) Amino-terminal amino acids modulate sigma-factor DNA-binding activity. Genes & development, 7 (12A): 2446–55

Dombroski, A.J., Walter, W.A., Record, M.T., et al. (1992) Polypeptides containing highly conserved regions of transcription initiation factor sigma 70 exhibit specificity of binding to promoter DNA. Cell, 70 (3): 501–12

Dougan, D.A., Weber-Ban, E. and Bukau, B. (2003) Targeted delivery of an ssrA-tagged substrate by the adaptor protein SspB to its cognate AAA+ protein ClpX. Molecular cell, 12 (2): 373–80

Dove, S.L., Darst, S.A. and Hochschild, A. (2003) Region 4 of σ as a target for transcription regulation. Molecular Microbiology, 48 (4): 863–874

England, P., Westblade, L.F., Karimova, G., et al. (2008) Binding of the unorthodox transcription activator, Crl, to the components of the transcription machinery. The Journal of biological chemistry, 283 (48): 33455–64

English, B.P., Hauryliuk, V., Sanamrad, A., et al. (2011) Single-molecule investigations of the stringent response machinery in living bacterial cells. Proceedings of the National Academy of Sciences of the United States of America, 108 (31): E365–E373

Epshtein, V., Dutta, D., Wade, J., et al. (2010) An allosteric mechanism of Rho-dependent transcription termination. Nature, 463 (7278): 245–9

Errington, J. (2003) Regulation of endospore formation in Bacillus subtilis. Nature reviews. Microbiology, 1 (2): 117–26

Farewell, A., Kvint, K. and Nyström, T. (1998) Negative regulation by RpoS: a case of sigma factor competition. Molecular Microbiology, 29 (4): 1039–1051

Feklistov, A., Barinova, N., Sevostyanova, A., et al. (2006) A Basal Promoter Element Recognized by Free RNA Polymerase σ Subunit Determines Promoter Recognition by RNA Polymerase Holoenzyme. Molecular Cell, 23 (1): 97–107

Feklistov, A. and Darst, S.A. (2011) Structural basis for promoter-10 element recognition by the bacterial RNA polymerase σ subunit. Cell, 147 (6): 1257–69

Feng, W.H., Mao, X.M., Liu, Z.H., et al. (2011) The ECF sigma factor SigT regulates actinorhodin production in response to nitrogen stress in Streptomyces coelicolor. Applied Microbiology and Biotechnology, 92 (5): 1009–1021

Flärdh, K. (2003) Essential role of DivIVA in polar growth and morphogenesis in Streptomyces coelicolor A3(2). Molecular Microbiology, 49 (6): 1523–36

Flärdh, K. and Buttner, M.J. (2009) Streptomyces morphogenetics: dissecting

differentiation in a filamentous bacterium. Nature reviews. Microbiology, 7 (1): 36–49

Flärdh, K., Richards, D.M., Hempel, A.M., et al. (2012) Regulation of apical growth and hyphal branching in Streptomyces. Current opinion in microbiology, 15 (6): 737–43

Fong, B.A., Gillies, A.R., Ghazi, I., et al. (2010) Purification of Escherichia coli RNA polymerase using a self-cleaving elastin-like polypeptide tag. Protein science : a publication of the Protein Society, 19 (6): 1243–52

Francke, C., Groot Kormelink, T., Hagemeijer, Y., et al. (2011) Comparative analyses imply that the enigmatic Sigma factor 54 is a central controller of the bacterial exterior. BMC genomics, 12 (1): 385

Furman, R., Tsodikov, O. V, Wolf, Y.I., et al. (2013) An insertion in the catalytic trigger loop gates the secondary channel of RNA polymerase. Journal of molecular biology, 425 (1): 82–93

Gaal, T., Bartlett, M.S., Ross, W., et al. (1997) Transcription regulation by initiating NTP concentration: rRNA synthesis in bacteria. Science (New York, N.Y.), 278 (5346): 2092–2097

Gaal, T., Mandel, M.J., Silhavy, T.J., et al. (2006) Crl facilitates RNA polymerase holoenzyme formation. Journal of Bacteriology, 188 (22): 7966–70

Gaal, T., Ross, W., Estrem, S.T., et al. (2001) Promoter recognition and discrimination by EsigmaS RNA polymerase. Molecular Microbiology, 42 (4): 939–954

Gao, B. and Gupta, R.S. (2012) Phylogenetic Framework and Molecular Signatures for the Main Clades of the Phylum Actinobacteria. Microbiology and Molecular Biology Reviews. 76 (1) pp. 66–112

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): 4275–4280

Gentry, D.R. and Burgess, R.R. (1989) rpoZ, encoding the omega subunit of Escherichia coli RNA polymerase, is in the same operon as spoT. Journal of Bacteriology, 171 (3): 1271–1277

Gerber, N.N. and Lechevalier, H.A. (1965) Geosmin, an earthly-smelling substance isolated from actinomycetes. Applied microbiology, 13 (6): 935–938

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 α 2 β subassembly. European Journal of Biochemistry, 268 (17): 4621–4627

Gottesman, S., Roche, E., Zhou, Y., et al. (1998) The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the SsrA-tagging system. Genes & Development, 12 (9): 1338–1347

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

Grainger, D.C. and Busby, S.J.W. (2008) Methods for studying global patterns of DNA binding by bacterial transcription factors and RNA polymerase. Biochemical Society transactions, 36 (Pt 4): 754–7

Grainger, D.C., Hurd, D., Harrison, M., et al. (2005) Studies of the distribution of Escherichia coli cAMP-receptor protein and RNA polymerase along the E. coli chromosome. Proceedings of the National Academy of Sciences, 102 (49): 17693–17698

Gregory, M. a., Till, R. and Smith, M.C.M. (2003) Integration site for Streptomyces phage ??BT1 and development of site-specific integrating vectors. Journal of Bacteriology, 185 (17): 5320–5323

Griffith, K.L. and Grossman, A.D. (2008) Inducible protein degradation in Bacillus subtilis using heterologous peptide tags and adaptor proteins to target substrates to the protease ClpXP. Molecular Microbiology, 70 (4): 1012–25

Gross, C., Engbaek, F., Flammang, T., et al. (1976) Rapid micromethod for the purification of Escherichia coli ribonucleic acid polymerase and the preparation of bacterial extracts active in ribonucleic acid synthesis. Journal of Bacteriology, 128 (1): 382–9

Gusarov, I. and Nudler, E. (1999) The mechanism of intrinsic transcription termination. Molecular cell, 3 (4): 495–504

Hahn, M.-Y., Bae, J.-B., Park, J.-H., et al. (2003) Isolation and characterization of Streptomyces coelicolor RNA polymerase, its sigma, and antisigma factors. Methods in enzymology, 370: 73–82

Hahn, M.-Y. and Roe, J.-H. (2007) Partial purification of factors for differential transcription of the rrnD promoters for ribosomal RNA synthesis in Streptomyces coelicolor. Journal of microbiology (Seoul, Korea), 45 (6): 534–540

Hamagishi, Y., Yoshimoto, A. and Oki, T. (1981) Determination of guanosine tetraphosphate (ppGpp) and adenosine pentaphosphate (pppApp) in various microorganisms by radioimmunoassay. Archives of Microbiology, 130 (2): 134–137

Harley, C.B. and Reynolds, R.P. (1987) Analysis of E. coli promoter sequences. Nucleic acids research, 15 (5): 2343–61

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): 1564–1568

Haug, I., Weissenborn, A., Brolle, D., et al. (2003) Streptomyces coelicolor A3(2) plasmid SCP2*: Deductions from the complete sequence. Microbiology. 149 (2) pp. 505–513

Haugen, S.P., Berkmen, M.B., Ross, W., et al. (2006) rRNA promoter regulation by nonoptimal binding of sigma region 1.2: an additional recognition element for RNA polymerase. Cell, 125 (6): 1069–82

Haugen, S.P., Ross, W. and Gourse, R.L. (2008a) Advances in bacterial promoter recognition and its control by factors that do not bind DNA. Nature reviews. Microbiology, 6 (7): 507–19

Haugen, S.P., Ross, W., Manrique, M., et al. (2008b) Fine structure of the promoter-sigma region 1.2 interaction. Proceedings of the National Academy of Sciences of the United States of America, 105 (9): 3292–3297

Hawley, D.K. and McClure, W.R. (1983) Compilation and analysis of Escherichia coli promoter DNA sequences. Nucleic acids research, 11 (8): 2237–55

Helmann, J.D. (1995) Compilation and analysis of Bacillus subtilis sigma A-dependent promoter sequences: evidence for extended contact between RNA polymerase and upstream promoter DNA. Nucleic acids research, 23 (13): 2351–60

Helmann, J.D. (2002) The extracytoplasmic function (ECF) sigma factors. Advances in Microbial Physiology. 46 pp. 47–110

Hempel, A.M., Wang, S., Letek, 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): 7579–83

Henard, C.A., Tapscott, T., Crawford, M.A., et al. (2014) The 4-cysteine zinc-finger motif of the RNA polymerase regulator DksA serves as a thiol switch for sensing oxidative and nitrosative stress. Molecular Microbiology, 91 (4): 790–804

Herring, C.D., Raffaelle, M., Allen, T.E., et al. (2005) Immobilization of Escherichia coli RNA polymerase and location of binding sites by use of chromatin immunoprecipitation and microarrays. Journal of Bacteriology, 187 (17): 6166–74

Hesketh, A., Chen, W.J., Ryding, J., et al. (2007) The global role of ppGpp synthesis in morphological differentiation and antibiotic production in Streptomyces coelicolor A3(2). Genome biology, 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): 136–144

Hett, E.C. and Rubin, E.J. (2008) Bacterial growth and cell division: a mycobacterial perspective. Microbiology and molecular biology reviews : MMBR, 72 (1): 126–156, table of contents

Hochschild, A. and Dove, S.L. (1998) Protein–Protein Contacts that Activate and Repress Prokaryotic Transcription. Cell, 92 (5): 597–600

Hodgson, D.A. (2000) Primary metabolism and its control in streptomycetes: a most unusual group of bacteria. Advances in microbial physiology, 42: 47–238

Hong, H.-J., Paget, M.S.B. and Buttner, M.J. (2002) A signal transduction system in Streptomyces coelicolor that activates the expression of a putative cell wall glycan operon in response to vancomycin and other cell wall-specific antibiotics. Molecular Microbiology, 44 (5): 1199–1211

Hook-Barnard, I.G. and Hinton, D.M. (2009) The promoter spacer influences transcription initiation via sigma70 region 1.1 of Escherichia coli RNA polymerase. Proceedings of the National Academy of Sciences of the United States of America, 106 (3): 737–42

Hopp, T.P., Prickett, K.S., Price, V.L., et al. (1988) A Short Polypeptide Marker Sequence Useful for Recombinant Protein Identification and Purification. Bio/Technology, 6 (10): 1204–1210

Hopwood, D.A. (1999) Forty years of genetics with Streptomyces: from in vivo through in vitro to in silico. Microbiology (Reading, England), 145 (Pt 9 (9): 2183–202

Hopwood, D.A. and Glauert, A.M. (1960) The fine structure of Streptomyces coelicolor. II. The nuclear material. The Journal of biophysical and biochemical cytology, 8: 267–78

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): 3575–9

Hu, Y., Morichaud, Z., Chen, S., et al. (2012) Mycobacterium tuberculosis RbpA protein is a new type of transcriptional activator that stabilizes the ?? a-containing RNA polymerase holoenzyme. Nucleic Acids Research, 40 (14): 6547–6557

Huang, J., Shi, J., Molle, V., et al. (2005) Cross-regulation among disparate antibiotic biosynthetic pathways of Streptomyces coelicolor. Molecular microbiology, 58 (5): 1276–87

Hubin, E.A., Tabib-Salazar, A., Humphrey, L.J., 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): 7171–7176

Jishage, M. and Ishihama, A. (1998) A stationary phase protein in Escherichia coli with binding activity to the major sigma subunit of RNA polymerase. Proceedings of the National Academy of Sciences of the United States of America, 95 (9): 4953–8

Jishage, M. and Ishihama, A. (1999) Transcriptional organization and in vivo role of the Escherichia coli rsd gene, encoding the regulator of RNA polymerase sigma D. Journal of bacteriology, 181 (12): 3768–76

Jones, A.C., Gust, B., Kulik, A., et al. (2013) Phage P1-Derived Artificial Chromosomes Facilitate Heterologous Expression of the FK506 Gene Cluster. PLoS ONE, 8 (7): e69319

Kallifidas, D., Thomas, D., Doughty, P., et al. (2010) The sigmaR regulon of Streptomyces coelicolor A32 reveals a key role in protein quality control during disulphide stress. Microbiology, 156 (Pt 6): 1661–1672

Kang, J.G., Hahn, M.Y., Ishihama, A., et al. (1997) Identification of sigma factors for growth phase-related promoter selectivity of RNA polymerases from Streptomyces coelicolor A3(2). Nucleic Acids Research, 25 (13): 2566–2573

Kang, J.G., Paget, M.S.B., Seok, Y.J., et al. (1999) RsrA, an anti-sigma factor regulated by redox change. The EMBO journal, 18 (15): 4292–4298

Kang, P.J. and Craig, E. a. (1990) Identification and characterization of a new Escherichia coli gene that is a dosage-dependent suppressor of a dnaK deletion mutation. Journal of Bacteriology, 172 (4): 2055–2064

Kang, S.G., Jin, W., Bibb, M., 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): 221–226

Kapanidis, A.N., Margeat, E., Ho, S.O., et al. (2006) Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. Science (New York, N.Y.), 314 (5802): 1144–7

Keiler, K.C. (2008) Biology of trans-translation. Annual review of microbiology, 62: 133–51

Kelemen, G.H., Plaskitt, K.A., Lewis, C.G., et al. (1995) Deletion of DNA lying close to the glkA locus induces ectopic sporulation in Streptomyces coelicolor A3(2). Molecular Microbiology, 17 (2): 221–230

Kelley, L.A., Mezulis, S., Yates, C.M., et al. (2015) The Phyre2 web portal for protein

modeling, prediction and analysis. **Nature protocols**, 10 (6): 845–858

Kieser, T., Bibb, M.J., Buttner, M.J., et al. (2000) **Practical Streptomyces Genetics**. John Innes Foundation

Kim, J.-N., Jeong, Y., Yoo, J.S., et al. (2015) Genome-scale analysis reveals a role for NdgR in the thiol oxidative stress response in Streptomyces coelicolor. **BMC genomics**, 16 (1): 116

Kim, J.H., Wei, J.R., Wallach, J.B., et al. (2011) Protein inactivation in mycobacteria by controlled proteolysis and its application to deplete the beta subunit of RNA polymerase. **Nucleic Acids Research**, 39 (6): 2210–2220

Kim, M.S., Dufour, Y.S., Yoo, J.S., et al. (2012) Conservation of thiol-oxidative stress responses regulated by SigR orthologues in actinomycetes. **Molecular Microbiology**, 85 (2): 326–344

Kinashi, H. and Shimaji-Murayama, M. (1991) Physical characterization of SCP1, a giant linear plasmid from Streptomyces coelicolor. **Journal of Bacteriology**, 173 (4): 1523–9

Koch, R. (1882) **I. Die Aetiologie der Tuberculose: Nach einem in der physiologischen Gesellschaft zu Berlin am 24. März cr. gehaltenen Vortrage**

Kodani, S., Hudson, M.E., Durrant, M.C., 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): 11448–53

Komissarova, N. and Kashlev, M. (1997) Transcriptional arrest: Escherichia coli RNA polymerase translocates backward, leaving the 3' end of the RNA intact and extruded. **Proceedings of the National Academy of Sciences of the United States of America**, 94 (5): 1755–60

Krásný, L. and Gourse, R.L. (2004) An alternative strategy for bacterial ribosome synthesis: Bacillus subtilis rRNA transcription regulation. **The EMBO journal**, 23 (22): 4473–4483

Kriel, A., Bittner, A.N., Kim, S.H., et al. (2012) Direct regulation of GTP homeostasis by (p)ppGpp: A critical component of viability and stress resistance. **Molecular Cell**, 48 (2): 231–241

Landick, R., Krek, A., Glickman, M.S., et al. (2014) Genome-Wide Mapping of the Distribution of CarD, RNAP σ(A), and RNAP β on the Mycobacterium smegmatis Chromosome using Chromatin Immunoprecipitation Sequencing. **Genomics data**, 2: 110–113

Lange, R. and Hengge-Aronis, R. (1994) The cellular concentration of the sigma S subunit of RNA polymerase in Escherichia coli is controlled at the levels of transcription, translation, and protein stability. **Genes & development**, 8 (13): 1600–12

Langmead, B., Trapnell, C., Pop, M., et al. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. **Genome biology**, 10 (3): R25

Lazzarini, R.A. and Dahlberg, A.E. (1971) The control of ribonucleic acid synthesis during amino acid deprivation in Escherichia coli. **The Journal of Biological Chemistry**, 246 (2): 420–429

Leibman, M. and Hochschild, A. (2007) A sigma-core interaction of the RNA polymerase holoenzyme that enhances promoter escape. The EMBO journal, 26 (6): 1579–90

Lesnik, E.A., Sampath, R., Levene, H.B., et al. (2001) Prediction of rho-independent transcriptional terminators in Escherichia coli. Nucleic acids research, 29 (17): 3583–94

Li, W., Stevenson, C.E.M., Burton, N., et al. (2002) Identification and structure of the anti-sigma factor-binding domain of the disulphide-stress regulated sigma factor sigma(R) from Streptomyces coelicolor. Journal of molecular biology, 323 (2): 225–36

Little, R., Ryals, J. and Bremer, H. (1983) rpoB mutation in Escherichia coli alters control of ribosome synthesis by guanosine tetraphosphate. Journal of Bacteriology, 154 (2): 787–92

Liu, F., Yang, M., Wang, X., et al. (2014) Acetylome analysis reveals diverse functions of lysine acetylation in Mycobacterium tuberculosis. Molecular & cellular proteomics : MCP, 13 (12): 3352–66

Liu, G., Chater, K.F., Chandra, G., et al. (2013) Molecular regulation of antibiotic biosynthesis in streptomyces. Microbiology and molecular biology reviews : MMBR, 77 (1): 112–43

Liu, K., Myers, A.R., Pisithkul, T., et al. (2015) Molecular Mechanism and Evolution of Guanylate Kinase Regulation by (p)ppGpp. Molecular cell, 57 (4): 735–49

Lonetto, M., Gribskov, M. and Gross, C.A. (1992) The sigma 70 family: sequence conservation and evolutionary relationships. Journal of Bacteriology. 174 (12) pp. 3843–3849

MacNeil, D.J., Gewain, K.M., Ruby, C.L., et al. (1992) Analysis of Streptomyces avermitilis genes required for avermectin biosynthesis utilizing a novel integration vector. Gene, 111 (1): 61–68

Magnusson, L.U., Gummesson, B., Joksimović, P., et al. (2007) Identical, independent, and opposing roles of ppGpp and DksA in Escherichia coli. Journal of Bacteriology, 189 (14): 5193–5202

Maillard, A.P., Girard, E., Ziani, W., et al. (2014) The crystal structure of the anti-σ factor CnrY in complex with the σ factor CnrH shows a new structural class of anti-σ factors targeting extracytoplasmic function σ factors. Journal of molecular biology, 426 (12): 2313–27

Manganelli, R., Voskuil, M.I., Schoolnik, G.K., et al. (2002) Role of the extracytoplasmic-function sigma factor sigma(H) in Mycobacterium tuberculosis global gene expression. Molecular microbiology, 45 (2): 365–74

Mao, X.-M., Zhou, Z., Cheng, L.-Y., et al. (2009) Involvement of SigT and RstA in the differentiation of Streptomyces coelicolor. FEBS letters, 583 (19): 3145–50

Marinelli, F. (2009) Antibiotics and Streptomyces: the future of antibiotic discovery. Microbiology Today, 36: 20–23

Mauri, M. and Klumpp, S. (2014) A model for sigma factor competition in bacterial cells. PLoS computational biology, 10 (10): e1003845

McGinness, K.E., Baker, T.A. and Sauer, R.T. (2006) Engineering controllable protein degradation. Molecular cell, 22 (5): 701–7

Mekler, V., Kortkhonjia, E., Mukhopadhyay, J., et al. (2002) Structural Organization of Bacterial RNA Polymerase Holoenzyme and the RNA Polymerase-Promoter Open Complex. Cell, 108 (5): 599–614

Minakhin, L., Nechaev, S., Campbell, E.A., et al. (2001) Recombinant Thermus aquaticus RNA polymerase, a new tool for structure-based analysis of transcription. Journal of bacteriology, 183 (1): 71–6

Miroux, B. and Walker, J.E. (1996) Over-production of proteins in Escherichia coli: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. Journal of molecular biology, 260 (3): 289–98

Mitchell, J.E., Zheng, D., Busby, S.J.W., et al. (2003) Identification and analysis of "extended -10" promoters in Escherichia coli. Nucleic acids research, 31 (16): 4689–95

Mittenhuber, G. (2002) An inventory of genes encoding RNA polymerase sigma factors in 31 completely sequenced eubacterial genomes. Journal of molecular microbiology and biotechnology, 4 (1): 77–91

Monsalve, M., Calles, B., Mencía, M., et al. (1997) Transcription activation or repression by phage psi 29 protein p4 depends on the strength of the RNA polymerase-promoter interactions. Molecular cell, 1 (1): 99–107

Monsalve, M., Mencia, M., Rojo, F., et al. (1996a) Activation and repression of transcription at two different phage phi29 promoters are mediated by interaction of the same residues of regulatory protein p4 with RNA polymerase. The EMBO journal, 15 (2): 383–91

Monsalve, M., Mencía, M., Salas, M., et al. (1996b) Protein p4 represses phage phi 29 A2c promoter by interacting with the alpha subunit of Bacillus subtilis RNA polymerase. Proceedings of the National Academy of Sciences of the United States of America, 93 (17): 8913–8

Monteil, V., Kolb, A., D'Alayer, J., et al. (2010) Identification of conserved amino acid residues of the Salmonella sigmaS chaperone Crl involved in Crl-sigmaS interactions. Journal of bacteriology, 192 (4): 1075–87

Moody, M.J., Young, R.A., Jones, S.E., et al. (2013) Comparative analysis of non-coding RNAs in the antibiotic-producing Streptomyces bacteria. BMC genomics, 14 (1): 558

Mortazavi, A., Williams, B.A., McCue, K., et al. (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nature methods, 5 (7): 621–8

Murakami, K.S. (2013) X-ray crystal structure of Escherichia coli RNA polymerase σ70 holoenzyme. The Journal of biological chemistry, 288 (13): 9126–34

Murakami, K.S. and Darst, S.A. (2003) Bacterial RNA polymerases: the wholo story. Current Opinion in Structural Biology, 13 (1): 31–39

Murakami, K.S., Masuda, S. and Darst, S.A. (2002) Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 A resolution. Science (New York, N.Y.), 296 (5571): 1280–1284

Myers, K.S., Park, D.M., Beauchene, N.A., et al. (2015) Defining bacterial regulons using ChIP-seq. Methods

Nene, V. and Glass, R.E. (1983) Relaxed mutants of Escherichia coli RNA polymerase. FEBS

letters, 153 (2): 307–10

Newell, K. V, Thomas, D.P., Brekasis, D., et al. (2006) The RNA polymerase-binding protein RbpA confers basal levels of rifampicin resistance on Streptomyces coelicolor. Molecular Microbiology, 60 (3): 687–696

Nguyen, K.T., Willey, J.M., Nguyen, L.D., et al. (2002) A central regulator of morphological differentiation in the multicellular bacterium Streptomyces coelicolor. Molecular Microbiology, 46 (5): 1223–1238

Nicol, J.W., Helt, G.A., Blanchard, S.G., et al. (2009) The Integrated Genome Browser: Free software for distribution and exploration of genome-scale datasets. Bioinformatics, 25 (20): 2730–2731

Nieselt, K., Battke, F., Herbig, A., et al. (2010) The dynamic architecture of the metabolic switch in Streptomyces coelicolor. BMC genomics, 11 (1): 10

Nyström, T. (2004) Growth versus maintenance: a trade-off dictated by RNA polymerase availability and sigma factor competition? Molecular Microbiology, 54 (4): 855–62

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): 2621–2631

Ochi, K. (1987a) Changes in Nucleotide Pools during Sporulation of Streptomyces griseus in Submerged Culture. Microbiology, 133 (10): 2787–2795

Ochi, K. (1987b) Metabolic initiation of differentiation and secondary metabolism by Streptomyces griseus: Significance of the stringent response (ppGpp) and GTP content in relation to a factor. Journal of Bacteriology, 169 (8): 3608–3616

Ochi, K. (1990) A relaxed (rel) mutant of Streptomyces coelicolor A3(2) with a missing ribosomal protein lacks the ability to accumulate ppGpp, A-factor and prodigiosin. Journal of general microbiology, 136 (12): 2405–2412

Oki, T., Yoshimoto, A., Sato, S., et al. (1975) Purine nucleotide pyrophosphotransferase from Streptomyces morookaensis, capable of synthesizing pppApp and pppGpp. Biochimica et biophysica acta, 410 (2): 262–72

Organization, W.H. (2010) Multidrug and extensively drug-resistant TB (M/XDR-TB) : 2010 global report on surveillance and response. Geneva : World Health Organization

Otani, H., Higo, A., Nanamiya, H., et al. (2013) An alternative sigma factor governs the principal sigma factor in Streptomyces griseus. Molecular Microbiology, 87 (6): 1223–1236

Paget, M.S. (2015) Bacterial Sigma Factors and Anti-Sigma Factors: Structure, Function and Distribution. Biomolecules, 5 (3): 1245–65

Paget, M.S., Kang, J.G., Roe, J.H., et al. (1998) sigmaR, 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): 5776–5782

Paget, M.S., Leibovitz, E. and Buttner, M.J. (1999a) A putative two-component signal transduction system regulates sigmaE, a sigma factor required for normal cell wall integrity in Streptomyces coelicolor A3(2). Molecular Microbiology, 33 (1): 97–107

Paget, M.S., Molle, V., Cohen, G., et al. (2001a) Defining the disulphide stress response in

Streptomyces coelicolor A3(2): identification of the sigmaR regulon. Molecular Microbiology, 42 (4): 1007–1020

Paget, M.S.B., Bae, J.-B., Hahn, M.-Y., et al. (2001b) Mutational analysis of RsrA, a zinc-binding anti-sigma factor with a thiol-disulphide redox switch. Molecular Microbiology, 39 (4): 1036–1047

Paget, M.S.B., Chamberlin, L., Atrih, A., et al. (1999b) Evidence that the extracytoplasmic function sigma factor ??(E) is required for normal cell wall structure in Streptomyces coelicolor A3(2). Journal of Bacteriology, 181 (1): 204–211

Paget, M.S.B. and Helmann, J.D. (2003) The sigma70 family of sigma factors. Genome biology, 4 (1): 203

Park, Y.-H., Lee, C.-R., Choe, M., et al. (2013) HPr antagonizes the anti-σ70 activity of Rsd in Escherichia coli. Proceedings of the National Academy of Sciences of the United States of America, 110 (52): 21142–7

Patikoglou, G.A., Westblade, L.F., Campbell, E.A., et al. (2007) Crystal structure of the Escherichia coli regulator of sigma70, Rsd, in complex with sigma70 domain 4. Journal of molecular biology, 372 (3): 649–59

Paul, B.J., Barker, M.M., Ross, W., 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): 311–322

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): 7823–7828

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): 147–51

Pedersen, F.S. and Kjeldgaard, N.O. (1977) Analysis of the relA gene product of Escherichia coli. European journal of biochemistry / FEBS, 76 (1): 91–97

Pedersen, L.B., Birkelund, S. and Christiansen, G. (1996) Purification of recombinant Chlamydia trachomatis histone H1-like protein Hc2, and comparative functional analysis of Hc2 and Hc1. Molecular microbiology, 20 (2): 295–311

Perederina, A., Svetlov, V., Vassylyeva, M.N., et al. (2004) Regulation through the secondary channel - Structural framework for ppGpp-DksA synergism during transcription. Cell, 118 (3): 297–309

Peters, J.M., Mooney, R.A., Kuan, P.F., et al. (2009) Rho directs widespread termination of intragenic and stable RNA transcription. Proceedings of the National Academy of Sciences of the United States of America, 106 (36): 15406–11

Pratt, L.A. and Silhavy, T.J. (1998) Crl stimulates RpoS activity during stationary phase. Molecular microbiology, 29 (5): 1225–36

Raffaelle, M., Kanin, E.I., Vogt, J., et al. (2005) Holoenzyme switching and stochastic release of sigma factors from RNA polymerase in vivo. Molecular Cell, 20 (3): 357–366

Ramírez, F., Dündar, F., Diehl, S., et al. (2014) deepTools: a flexible platform for exploring deep-sequencing data. Nucleic acids research, 42 (Web Server issue): W187–91

Rammohan, J., Ruiz Manzano, A., Garner, A.L., et al. (2015) CarD stabilizes mycobacterial open complexes via a two-tiered kinetic mechanism. Nucleic Acids Research, p. gkv078–

Reppas, N.B., Wade, J.T., Church, G.M., et al. (2006) The transition between transcriptional initiation and elongation in E. coli is highly variable and often rate limiting. Molecular cell, 24 (5): 747–57

Reyrat, J.-M. and Kahn, D. (2001) Mycobacterium smegmatis: an absurd model for tuberculosis? Trends in Microbiology, 9 (10): 472–473

Rigali, S., Titgemeyer, F., Barends, S., et al. (2008) Feast or famine: the global regulator DasR links nutrient stress to antibiotic production by Streptomyces. EMBO reports, 9 (7): 670–5

Roberts, J.W., Shankar, S. and Filter, J.J. (2008) RNA polymerase elongation factors. Annual review of microbiology, 62: 211–33

Rodrigue, S., Provvedi, R., Jacques, P.-E., et al. (2006) The sigma factors of Mycobacterium tuberculosis. FEMS microbiology reviews, 30 (6): 926–41

Roghanian, M., Yuzenkova, Y. and Zenkin, N. (2011) Controlled interplay between trigger loop and Gre factor in the RNA polymerase active centre. Nucleic acids research, 39 (10): 4352–9

Roghanian, M., Zenkin, N. and Yuzenkova, Y. (2015) Bacterial global regulators DksA/ppGpp increase fidelity of transcription. Nucleic acids research, 43 (3): 1529–36

Romero, D.A., Hasan, A.H., Lin, Y., et al. (2014) A comparison of key aspects of gene regulation in S treptomyces coelicolor and E scherichia coli using nucleotide-resolution transcription maps produced in parallel by global and differential RNA sequencing. Molecular Microbiology, 94 (5): 963–987

Ross, W., Vrentas, C.E., Sanchez-Vazquez, P., et al. (2013) The magic spot: A ppGpp binding site on E. coli RNA polymerase responsible for regulation of transcription initiation. Molecular Cell, 50 (3): 420–429

Rudd, B.A. and Hopwood, D.A. (1980) A pigmented mycelial antibiotic in Streptomyces coelicolor: control by a chromosomal gene cluster. Journal of general microbiology, 119 (2): 333–40

Ruff, E.F., Record, M.T. and Artsimovitch, I. (2015) Initial events in bacterial transcription initiation. Biomolecules, 5 (2): 1035–62

Russell, D.G. (2007) Who puts the tubercle in tuberculosis? Nature reviews. Microbiology, 5 (1): 39–47

Santos-Beneit, F., Barriuso-Iglesias, M., Fernández-Martínez, L.T., 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): 7586–94

Schatz, A., Bugle, E. and Waksman, S.A. (1944) Streptomycin, a Substance Exhibiting Antibiotic Activity Against Gram-Positive and Gram-Negative Bacteria.* . Experimental Biology and Medicine, 55 (1): 66–69

Schöbel, S., Zellmeier, S., Schumann, W., et al. (2004) The Bacillus subtilis sigmaW anti-sigma factor RsiW is degraded by intramembrane proteolysis through YluC. Molecular

Microbiology, 52 (4): 1091–105

Schwartz, E.C., Shekhtman, A., Dutta, K., et al. (2008) A Full-Length Group 1 Bacterial Sigma Factor Adopts a Compact Structure Incompatible with DNA Binding. Chemistry and Biology, 15 (10): 1091–1103

Sermonti, G. and Spada-Sermonti, I. (1955) Genetic Recombination in Streptomyces. Nature, 176 (4472): 121–121

Sevostyanova, A., Svetlov, V., Vassylyev, D.G., et al. (2008) The elongation factor RfaH and the initiation factor bind to the same site on the transcription elongation complex. Proceedings of the National Academy of Sciences, 105 (3): 865–870

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): 11004–11008

Sharma, U.K. and Chatterji, D. (2010) Transcriptional switching in Escherichia coli during stress and starvation by modulation of sigma activity. FEMS microbiology reviews, 34 (5): 646–57

Sims, D., Sudbery, I., Ilott, N.E., et al. (2014) Sequencing depth and coverage: key considerations in genomic analyses. Nature reviews. Genetics, 15 (2): 121–32

Skordalakes, E. and Berger, J.M. (2003) Structure of the Rho Transcription Terminator. Cell, 114 (1): 135–146

Sorenson, M.K., Ray, S.S. and Darst, S.A. (2004) Crystal structure of the flagellar sigma/anti-sigma complex sigma(28)/FlgM reveals an intact sigma factor in an inactive conformation. Molecular cell, 14 (1): 127–38

Srivastava, D.B., Leon, K., Osmundson, J., 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): 12619–24

Stallings, C.L., Stephanou, N.C., Chu, L., et al. (2009) CarD Is an Essential Regulator of rRNA Transcription Required for Mycobacterium tuberculosis Persistence. Cell, 138 (1): 146–159

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): 4389–4393

Strauch, E., Takano, E., Baylis, H.A., et al. (1991) The stringent response in Streptomyces coelicolor A3(2). Molecular Microbiology, 5 (2): 289–298

Tabib-Salazar, A., Liu, B., Doughty, P., et al. (2013) The actinobacterial transcription factor RbpA binds to the principal sigma subunit of RNA polymerase. Nucleic Acids Research, 41 (11): 5679–5691

Tanaka, K., Shiina, T. and Takahashi, H. (1988) Multiple principal sigma factor homologs in eubacteria: identification of the “rpoD box”. Science, 242 (4881): 1040–2

Tedin, K. and Bremer, H. (1992) Toxic effects of high levels of ppGpp in Escherichia coli are relieved by rpoB mutations. The Journal of biological chemistry, 267 (4): 2337–44

Tehranchi, A.K., Blankschien, M.D., Zhang, Y., et al. (2010) The Transcription Factor DksA Prevents Conflicts between DNA Replication and Transcription Machinery. Cell, 141 (4): 595–605

Tortora, G., Funke, B. and Case, C. (2012) Microbiology: An Introduction (11th Edition). 11th ed. Benjamin Cummings

Trautinger, B.W., Jaktaji, R.P., Rusakova, E., et al. (2005) RNA polymerase modulators and DNA repair activities resolve conflicts between DNA replication and transcription. Molecular Cell, 19 (2): 247–258

Travers, A.A. (1980) Promoter sequence for stringent control of bacterial ribonucleic acid synthesis. Journal of Bacteriology, 141 (2): 973–976

Ventura, M., Canchaya, C., Tauch, A., et al. (2007) Genomics of Actinobacteria: tracing the evolutionary history of an ancient phylum. Microbiology and molecular biology reviews : MMBR, 71 (3): 495–548

Vinella, D., Albrecht, C., Cashel, M., et al. (2005) Iron limitation induces SpoT-dependent accumulation of ppGpp in Escherichia coli. Molecular Microbiology, 56 (4): 958–970

Vogel, U. and Jensen, K.F. (1994) The RNA chain elongation rate in Escherichia coli depends on the growth rate. Journal of Bacteriology, 176 (10): 2807–13

Vrentas, C.E., Gaal, T., Berkmen, M.B., et al. (2008) Still looking for the magic spot: the crystallographically defined binding site for ppGpp on RNA polymerase is unlikely to be responsible for rRNA transcription regulation. Journal of Molecular Biology, 377 (2): 551–564

Vrentas, C.E., Gaal, T., Ross, W., et al. (2005) Response of RNA polymerase to ppGpp: Requirement for the ω subunit and relief of this requirement by DksA. Genes and Development, 19 (19): 2378–2387

Vuthoori, S., Bowers, C.W., McCracken, A., et al. (2001) Domain 1.1 of the sigma(70) subunit of Escherichia coli RNA polymerase modulates the formation of stable polymerase/promoter complexes. Journal of molecular biology, 309 (3): 561–572

Wagner, R. (2002) Regulation of ribosomal RNA synthesis in E. coli: effects of the global regulator guanosine tetraphosphate (ppGpp). Journal of molecular microbiology and biotechnology, 4 (3): 331–40

Waksman, S.A. and Lechevalier, H.A. (1951) The principle of screening antibiotic-producing organisms. Antibiotics & chemotherapy, 1 (2): 125–32

Waksman, S.A. and Woodruff, H.B. (1941) Actinomyces antibioticus, a New Soil Organism Antagonistic to Pathogenic and Non-pathogenic Bacteria 1. Journal of Bacteriology, 42 (2): 231–249

Weber, H., Polen, T., Heuveling, J., et al. (2005) Genome-Wide Analysis of the General Stress Response Network in Escherichia coli: S-Dependent Genes, Promoters, and Sigma Factor Selectivity. Journal of Bacteriology, 187 (5): 1591–1603

Weiss, L. a., Harrison, P.G., Nickels, B.E., et al. (2012) Interaction of CarD with RNA polymerase mediates Mycobacterium tuberculosis viability, rifampin resistance, and pathogenesis. Journal of Bacteriology, 194 (20): 5621–5631

Wendrich, T.M., Blaha, G., Wilson, D.N., et al. (2002) Dissection of the mechanism for the

stringent factor RelA. Molecular Cell, 10 (4): 779–788

Westblade, L.F., Campbell, E.A., Pukhrambam, C., et al. (2010) Structural basis for the bacterial transcription-repair coupling factor/RNA polymerase interaction. Nucleic Acids Research, 38 (22): 8357–8369

Van Wezel, G.P. and Bibb, M.J. (1996) A novel plasmid vector that uses the glucose kinase gene (glkA) for the positive selection of stable gene disruptants in Streptomyces. Gene, 182 (1-2): 229–230

Willey, J., Santamaria, R., Guijarro, J., et al. (1991) Extracellular complementation of a developmental mutation implicates a small sporulation protein in aerial mycelium formation by S. coelicolor. Cell, 65 (4): 641–50

Wright, L. and Hopwood, D. (1976a) Identification of the antibiotic determined by the SCP1 plasmid of Streptomyces coelicolor A3 (2). Journal of general microbiology

Wright, L.F. and Hopwood, D.A. (1976b) Actinorhodin is a chromosomally-determined antibiotic in Streptomyces coelicolar A3(2). Journal of general microbiology, 96 (2): 289–97

Xiao, H., Kalman, M., Ikehara, K., et al. (1991) Residual guanosine 3???,5???-bispyrophosphate synthetic activity of relA null mutants can be eliminated by spoT null mutations. Journal of Biological Chemistry, 266 (9): 5980–5990

Young, B.A., Anthony, L.C., Gruber, T.M., et al. (2001) A coiled-coil from the RNA polymerase beta' subunit allosterically induces selective nontemplate strand binding by sigma(70). Cell, 105 (7): 935–44

Yuan, A.H., Gregory, B.D., Sharp, J.S., et al. (2008) Rsd family proteins make simultaneous interactions with regions 2 and 4 of the primary sigma factor. Molecular Microbiology, 70 (5): 1136–51

Yuzenkova, Y., Tadigotla, V.R., Severinov, K., et al. (2011) A new basal promoter element recognized by RNA polymerase core enzyme. The EMBO journal, 30 (18): 3766–3775

Yuzenkova, Y. and Zenkin, N. (2010) Central role of the RNA polymerase trigger loop in intrinsic RNA hydrolysis. Proceedings of the National Academy of Sciences of the United States of America, 107 (24): 10878–83

Zalacain, M., González, A., Guerrero, M.C., et al. (1986) Nucleotide sequence of the hygromycin B phosphotransferase gene from Streptomyces hygroscopicus. Nucleic acids research, 14 (4): 1565–81

Zenkin, N. and Yuzenkova, Y. (2015) New Insights into the Functions of Transcription Factors that Bind the RNA Polymerase Secondary Channel. Biomolecules, 5 (3): 1195–1209

Zhang, G., Campbell, E.A., Minakhin, L., et al. (1999) Crystal structure of Thermus aquaticus core RNA polymerase at 3.3 A resolution. Cell, 98 (6): 811–24

Zhang, Y., Liu, T., Meyer, C.A., et al. (2008) Model-based analysis of ChIP-Seq (MACS). Genome biology, 9 (9): R137

Zhang, Y., Mooney, R.A., Grass, J.A., et al. (2014) DksA guards elongating RNA polymerase against ribosome-stalling-induced arrest. Molecular cell, 53 (5): 766–78

Zhou, Y.N. and Jin, D.J. (1998) The rpoB mutants destabilizing initiation complexes at

stringently controlled promoters behave like “stringent” RNA polymerases in Escherichia coli. Proceedings of the National Academy of Sciences of the United States of America, 95 (6): 2908–2913

Zumla, A., Nahid, P. and Cole, S.T. (2013) Advances in the development of new tuberculosis drugs and treatment regimens. Nature reviews. Drug discovery, 12 (5): 388–404

Zuo, Y. and Steitz, T.A. (2015) Crystal Structures of the E. coli Transcription Initiation Complexes with a Complete Bubble. Molecular Cell, 58 (3): 534–540

Zuo, Y., 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. Molecular Cell, 50 (3): 430–436

9 Appendix

Sequence of synthesised *rbpA*-DAS+4 fragment

ACTAGTAAGCTTGGCCCGGATCGGGAGTCGGAACGGGAATCTTTACCGCCGCCCGGACGTTGACCGGATG
ACGACGACAGCGACACCTGTCCTGTGGGCGACAAGCCCGGGAGGCACGATTCATGAGTGAGCGAGCTCTT
CGCGGCACGCGCCTCGTGGTGACCAGCTACGAGACGGACCGCGGCATCGACCTGGCCCCGCGCCAGGCCG
TGGAGTACGCATGCGAGAAGGGGCATCGATTCGAGATGCCCTTCTCGGTCGAGGCGGAGATCCCGCCGGA
GTGGGAGTGCAAGGTCTGCGGGGCCCAGGCACTCCTGGTTGACGGCGACGGCCCTGAGGAGAAGAAGGCC
AAGCCCGCGCGTACGCACTGGGACATGCTGATGGAGCGACGCACCCGCGAGGAACTCGAAGAGGTCCTCG
AGGAGCGGCTGGCCGTTCTGCGCTCCGGCGCGATGAACATCGCGGTCCATCCGCGAGACAGCCGCAAGAG
TGCG**AAGCTT**GCCGCCAACGACGAGAACTACTCCGAGAACTACGCCGACGCCTCCTGA**GGATCC**

Sequence of *S. coelicolor* codon optimised *sspB*

TCTAGAGGCAGCGTGGACGGCGTCAGAGAAGGGAGCGGACATATGGACTTGTCACAGCTCACACCGAGGC
GTCCCTACCTGCTGCGTGCATTCTATGAGTGGTTGCTGGATAACCAGCTCACGCCGCACCTGGTGGTGGA
TGTGACGCTCCCTGGCGTGCAGGTACCTATGGAATATGCGCGTGACGGGCAAATCGTACTCAACATCGCG
CCGCGTGCTGTCGGCAACCTGGAACTGGCGAACGATGAGGTGCGCTTCAACGCGCGCTTCGGTGGCATCC
CGCGTCAGGTTTCGGTGCCGCTGGCTGCCGTGCTGGCTATCTACGCCCGTGAAAACGGCGCAGGCACGAT
GTTCGAGCCTGAAGCTGCCTACGATGAAGATACCAGCATCATGAACGATGAAGAGGCATCGGCAGACAAC
GAAACCGTTATGTCGGTGATCGATGGCGACAAGCCAGATCACGATGATGACACTCATCCTGACGATGAAC
CTCCGCAGCCACCACGCGGTGGTCGACCGGCACTGCGCGTTGTGAAGTGA**TCTAGA**

10 Publications

Hubin, E.A., Tabib-Salazar, A., Humphrey, L.J., Flack, J.E., Olinares, P.D.B., Darst, S.A., Campbell, E.A., Paget, M.S (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 112:7171–7176.