AN INVESTIGATION OF THE REGULATION IN TWO GENETIC REGIONS HARBOURING ANTISENSE RNA IN *STREPTOMYCES COELICOLOR*

AN INVESTIGATION OF THE REGULATION IN TWO GENETIC REGIONS HARBOURING ANTISENSE RNA IN *STREPTOMYCES COELICOLOR*

By

HINDRA, B.Sc., M.Sc.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy

McMaster University

McMaster University DOCTOR OF PHILOSOPHY (2012) Hamilton, Ontario (Biology)

TITLE: An Investigation of the Regulation in Two Genetic Regions Harbouring Antisense RNA in *Streptomyces coelicolor*

AUTHOR: Hindra, B.Sc. (University of Pelita Harapan), M.Sc. (University of Saskatchewan)

SUPERVISOR: Dr. Marie A. Elliot

NUMBER OF PAGES: xiv, 151

ABSTRACT

Bacterial small RNAs have emerged as a class of molecules having important regulatory roles. Accumulating numbers of *cis*-encoded sRNAs (antisense RNAs) have been recently discovered to be transcribed from the chromosomal DNA of many bacterial species, including the streptomycetes. Here, we investigate potential regulatory roles for two *S. coelicolor* antisense RNAs, scr4677 and α-abeA.

The scr4677 antisense RNA is transcribed from the intergenic region between *SCO4676* (a gene encoding a conserved protein of unknown function) and *SCO4677*, encoding a regulatory protein with proposed anti-sigma factor activity. Transcription profiling revealed that scr4677 may not only interact with *SCO4676* mRNA but also with *SCO4677-4676* read-through transcripts. Our study suggested that scr4677 functioned to destabilize *SCO4676* mRNA, at the same time that it stabilized the *SCO4677-4676* read-through transcript. The potential role for scr4677 in destabilizing *SCO4676* mRNA was not mediated by the double stranded ribonuclease RNase III. Genetic analysis showed *scr4677* transcription was affected by SCO4677, and the transcription was apparently dependent on an unknown protein binding to the *SCO4676* coding sequence.

A second independent study focused on investigating the regulation of a previously uncharacterized genetic region, *SCO3287-3290*, since renamed *abeABCD*. This region contains an antisense RNA (α-abeA)-encoding gene, and

is adjacent to the downstream *SCO3291* (*abeR*) gene, which encodes a putative regulatory protein. Genetic analysis revealed that overexpression of *abeR* or *abeABCD* stimulated the production of the blue-pigmented antibiotic actinorhodin, and deletion of *abeR* impaired actinorhodin production. Transcription analysis revealed the *abe* genes (including α-*abeA*) to be subject to multiple levels of regulation. We found an internal promoter within the *abeA* coding sequence and that required AbeR for expression. Furthermore, biochemical experiments demonstrated that AbeR regulated *abeBCD* directly, by binding to four heptameric repeats in its promoter region. The expression of α-*abeA* and other *abe* genes were differentially affected by RNase III.

ACKNOWLEDGEMENTS

I would like to express my gratitude to all individuals who have made the completion of this thesis possible, especially to Dr. Marie Elliot, for taking me as her student. She always provides me with invaluable guidance, trust, support, help and inspiration. I also would like to thank Dr. Elliot for devoting her time to supply feedback throughout the writing process of this thesis and advice for me considering future career options.

I would like to extend my gratitude to the members of my supervisory committee, Dr. Justin Nodwell and Dr. Turlough Finan whose questions, comments and suggestions have stimulated me to understand deeper my research study. I also would like to express my appreciation to Dr. Nodwell and Dr. Finan for helping and supporting me in many aspects over the years. I would like to acknowledge with much appreciation my thesis examining committee, Dr. Eric Massé for taking time to evaluate this thesis, and Dr. Douglas Davidson for serving as a chair during the oral examination and providing constructive comments on this thesis.

I am very grateful that, during my study, I have come to know past and present members of the Elliot laboratory. Each individual has always shared unique characters to create a great working atmosphere. I would like to thank: Julia Swiercz and Dr. Jan Bobek for working together to complete the small RNA paper, Patricia Pak for her work supporting the completion of the *abe* gene cluster

paper, Dr. David Capstick and Andrew Duong for kindly providing their RNA samples, Drs. David Capstick and Henry Haiser for providing inspiration to write a thesis. Other thanks go to Matt Moody, Renee St-Onge, Drs. Chan Gao, Mary Yousef and Rahat Zaheer for sharing their scientific expertise, and the members of Nodwell, Finan, Schellhorn, Daniel and Igdoura laboratories. I also would like to thank Marg Biggs, Karen Haines, Luce Lavigne, Patricia Hayward, A. Seong Cheong, Marvin Gunderman, Klaus Schultes, Christina Di Berardo and Barb Reuter for their hospitality, administrative and technical supports. Lastly, special thanks go to Nolan D'Souza, Andrew Duong, Dr. Marie Elliot, Julia Swiercz and Patricia Pak, for reminding me to run my other role in my family.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

^{32}P	phosphorus-32 isotope
A	adenine
aa	amino acid
ATP	adenosine triphosphate
bp	base pairs
C	carboxyl or Celsius or cytosine
CDA	calcium-dependent antibiotic
cDNA	complementary DNA
DNA	deoxyribonucleic acid
dATP	2′-deoxyadenosine-5′-triphosphate
ddG/A/T/C	2′,3′-dideoxynucleoside
dGTP	2′-deoxyguanosine 5′-triphosphate
dNTP	deoxyribonucleotide
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
fMet	N-formylmethionine
G	guanine
g	grams or gravity
h	hours
His6	hexahistidine
IPTG	isopropyl β-D-1-thiogalactopyranoside
KOH	potassium hydroxide
kV	kilovolts
L	liter
M	molar
Mbp	mega basepairs
mg	milligrams
min	minutes

mL	milliliter
mm	millimeters
mM	millimolar
mRNA	messenger RNA
N	amino
NaCl	sodium chloride
NaOH	sodium hydroxide
nM	nanomolar
nt	nucleotides
OD	optical density
PCR	polymerase chain reaction
pmol	picamoles
RNA	ribonucleic acid
RPM	revolutions per minute
rRNA	ribosomal RNA
s	seconds
SDS	sodium dodecyl sulphate
SSC	saline sodium citrate
T	thymine
tRNA	transfer RNA
TES	N-tris(hydroxymethyl)methyl-2-aminoethanesulfonic acid
Tris	tris(hydroxymethyl)aminomethane
U	enzyme units or uracil
UV	ultraviolet
v/v	volume per volume
W	Watt
w/v	weight per volume
WT	wild-type
µg	micrograms
µL	microliter

CHAPTER 1

GENERAL INTRODUCTION

1.1. *Streptomyces*

Streptomyces is the largest genus within the actinomycetes, a prominent group of Gram-positive bacteria with a high GC content (National Center for Biotechnology Information [NCBI] taxonomic database). This genus is commonly found in the soil, where they have been proposed to contribute to carbon recycling (Bentley *et al*., 2002). This notion has been supported by increasing lines of evidence, including the isolation of this genus from municipal solid waste compost (Mokni-Tlili *et al*., 2010) and phytopathogen-suppressive soil (Hjort *et al*., 2010). Streptomycetes have also been isolated from insects, mangrove and marine habitats [Haeder *et al*. (2009); Xu *et al*. (2009), Hodges *et al*. (2012), Tian *et al*. (2012), Khan *et al*. (2011)]. These more broadly heterogeneous sites of isolation suggest that some streptomycetes live in symbiotic association with other organisms (fungi, plants, and animals as reviewed in Seipke *et al*., 2011), in addition to dwelling in the soil as free-living bacteria. Therefore, the streptomycetes play an appreciable - and diverse - role in the ecosystem.

Streptomyces grow multicellularly through several growth stages, culminating in spore formation (reviewed in Flärdh and Buttner, 2009; Chater,

1993). Germination of spores begins with spore swelling, followed by the subsequent emergence of germ tubes. These tubes represent the earliest stages in the development of vegetative or substrate hyphae. Hyphal elongation and branching ultimately result in the formation of a hyphal network called the vegetative or substrate mycelium. Following the establishment of the vegetative mycelium, a second hyphal type can then be found growing upwards, away from the medium surface; these upwardly growing hyphae are termed aerial hyphae. At later stages of growth, aerial hyphae are then segmented through a synchronous cell division event and are transformed into chains of spore compartments (reviewed in Flärdh and Buttner, 2009).

In addition to undergoing complex morphological differentiation, another interesting attribute that has stimulated interest in *Streptomyces* biology is their secondary metabolism (Chater, 2006). Many secondary metabolites – compounds that do not appear to be essential for growth - produced by the streptomycetes have found use as medically important-compounds, including antibiotics, immunosuppressants, antifungal and antiparasitic agents, antitumor agents, enzyme inhibitors, and hypocholesterolemic agents (reviewed in Challis and Hopwood, 2003; Williamson *et al.*, 2006; Saudagar *et al.*, 2008, Manzoni and Rollini, 2002). Several *Streptomyces* metabolites have also been commercially applied as herbicides, and many of their metabolites are being reviewed as potential herbicide candidates (Dayan *et al*., 2011).

1.2. *Streptomyces coelicolor* as a model organism

The complexity of *Streptomyces* morphological development and their invaluable secondary metabolite production have led scientists to further explore this genus at genomic, transcriptomic, proteomic and metabolomic levels. *Streptomyces coelicolor* was the first *Streptomyces* species to have its genome completely sequenced (Bentley *et al*., 2002). Since then, complete genomes of many *Streptomyces* species (more than 40 species as of May 2012; NCBI genome database) have become available to the public. In addition to the completion of the genome sequence for *S. coelicolor*, other genetic tools have been developed/established for use in this organism, including cosmid and transposon libraries, gene cloning techniques, a wide variety of characterized mutant strains and whole-genome microarrays. Thus, this species has been used as the model organism for studying the genetics and morphological development of the streptomycetes. *Streptomyces coelicolor* is also often used to study antibiotic production because this species secretes the blue and red pigmented antibiotics, actinorhodin (Act) (Wright and Hopwood, 1976a) and undecylprodigiosin (Red) (Rudd and Hopwood, 1980), respectively. These pigmented metabolites simplify the identification and relative quantification of antibiotic production. Another *S. coelicolor* antibiotic, the calcium-dependent antibiotic (CDA) (Hopwood and Wright, 1983; Lakey *et al*., 1983), can be detected using a straightforward plate-based bioassay. The biosynthetic gene clusters for Act, Red, and CDA are located on the *S. coelicolor* chromosome, whereas the methylenomycin (another

antibiotic produced by *S. coelicolor*) biosynthetic gene cluster is found on the *S. coelicolor* linear plasmid SCP1 (Kirby *et al*., 1975; Wright and Hopwood, 1976b).

1.3. Genetic regulation of secondary metabolism in *Streptomyces*

Secondary metabolism, as discussed here, focusses primarily on metabolites that exert antibacterial/antibiotic activities (hereafter referred to as simply 'antibiotics'). The metabolism is thought to occur as a response to various biological and environmental cues. How these stimuli are sensed by *Streptomyces* is not yet thoroughly understood, although these factors likely stimulate the onset of secondary metabolism through a tightly controlled process with multiple levels of regulation.

1.3.1. Multilevel regulation of secondary metabolism

The onset of secondary metabolism is typically associated with a reduction in growth rate or growth cessation (Bibb, 2005). Genes involved in antibiotic biosynthetic pathways are typically clustered together on the chromosome, and each biosynthetic gene cluster generally has its own pathway-specific regulator (Bibb, 1996; Martin and Liras, 2010). Classical examples of pathway-specific regulators include DnrI in daunorubicin production in *S. peucetius*, ActII-ORF4 in actinorhodin production in *S. coelicolor*, RedD for undecylprodigiosin production in *S. coelicolor* and CcaR for clavulanic acid in *S. clavuligerus* (Sheldon *et al*., 2002; Arias *et al*., 1999; Narva and Feitelson, 1990; Santamarta *et al*., 2002).

These pathway-specific regulators all belong to a protein family known as the *Streptomyces* antibiotic regulatory proteins (SARPs), and contain an OmpR-like DNA binding domain that recognizes direct repeat sequences within the promoter regions of their target genes (Wietzorrek & Bibb, 1997). While many SARPs function as pathway-specific regulators, there are instances where they also act as global regulators, governing the biosynthesis of multiple antibiotics. AfsR is the best studied example of a globally-acting SARP (Tanaka *et al*., 2007), in that it impacts the production of Act, Red, and CDA in *S. coelicolor*, through the control of transcription of the adjacent *afsS* gene (Floriano & Bibb, 1996; Lee *et al.*, 2002). AfsS encodes a sigma factor-like protein, but exactly how it exerts its effects on antibiotic production is currently unknown. Global antibiotic production is also impacted by non-SARP regulators such as RNase III – an RNA processing enzyme. RNase III is known to contribute to secondary metabolism based on observations that its mutants are defective in synthesizing the four *S. coelicolor* antibiotics (Price *et al.*, 1999; Adamidis and Champness, 1992). Interestingly, cross-talk amongst pathway specific regulators and between global regulators has been reported by Huang *et al*. (2005) and Santos-Beneit *et al*. (2009), respectively.

Adding to the complexity, a number of regulatory factors impact not only secondary metabolism but also morphological development. These include gene products of the *bld* genes, such as *bldA* (reviewed in Chater, 2006), *bldB* (Pope *et al*., 1998), *bldD* (Elliot *et al*., 1998; den Hengst *et al*., 2010). The *bldA* gene

encodes a leucyl tRNA (Lawlor *et al.,* 1987), which is required for the translation of UUA codons. While UUA codons occur rarely in *S. coelicolor* genes (Leskiw *et al*., 1991), this codon is found in pathway-specific antibiotic regulators, such as *actII-ORF4* and *redZ* (encoding a response regulator-like protein required for *redD* transcription) and in determinants for methylenomycin production (*mmyB* and *mmfL*, O'Rourke *et al*., 2008). A TTA codon is also found in *bldH* – a gene that is also known as *adpA*, and it is the inability to translate this codon that renders *bldA* mutants unable to raise aerial hyphae (Takano *et al*., 2003; Nguyen *et al*., 2003). More recently, BldD has been demonstrated to directly bind downstream of *bldA* and upstream of *bldH*, and products of these three *bld* genes have appeared to have common regulons (den Hengst *et al*., 2010). Both antibiotic production and morphological differentiation are also affected by a small protein encoded by *bldB* (Pope *et al*., 1998; Eccleston *et al*., 2002). The mechanism underlying pleiotropic effects of BldB is currently unknown, but it has been suggested that BldB interacts with additional subcellular components in exerting its phenotypic effects (Eccleston *et al*., 2006).

1.3.2. Secondary metabolism is a possible cellular response to diverse extracellular signals

To effectively thrive as a free living organism, *Streptomyces* possess a multitude of 'tools', which include the ability to sense variable environmental conditions. One of these 'tools' would be the two-component signal transduction systems (TCS), which constitute a sensor kinase and response regulator pair.

Sensor kinases are typically membrane proteins that have an extracellular 'sensor' domain, a transmembrane domain and an intracellular kinase domain, whereas response regulators are cytoplasmic proteins with a phosphorylation pocket and a DNA binding domain. *S. coelicolor* encodes more than 60 pairs of sensor kinase and response regulator-encoding genes (Hutchings *et al*., 2004), in addition to having a number of 'orphaned' (not paired) sensor kinases and response regulators. Although a mechanistic understanding of how sensor kinase activity can be triggered by environmental factors has not been established for many of these proteins, some TCSs have been identified as having a role in mediating secondary metabolism.

One of the best studied examples of this is the sensor kinase AbsA1 and its cognate response regulator AbsA2, which play an important role in regulating secondary metabolism in *S. coelicolor* (Anderson *et al*., 2001; Ryding *et al*., 2002; McKenzie and Nodwell, 2007). The signal transduction activities of AbsA1/A2 have been biochemically validated (Sheeler *et al*., 2005), with the phosphorylated form of AbsA2 binding to the promoter region of multiple pathway-specific regulator-encoding genes (e.g. *actII-ORF4*, *redZ, cdaR*), repressing their transcription activities (McKenzie and Nodwell, 2007). A second well-studied example is the PhoR/PhoP sensor kinase/response regulator pair. This TCS functions in sensing the availability of phosphate in the environment and mediating phosphate uptake and utilization (Sola-Landa *et al*. 2003). However, PhoP also binds to the promoter region of *afsS*, competing for binding

with the SARP-like global regulator AfsR (Santos-Beneit *et al*., 2009), which is an activator of *afsS* expression (Lee *et al.*, 2002). Therefore, in addition to its role in phosphate control, PhoP also functions to repress *afsS* transcription, and given that AfsS is a global activator of antibiotic production, PhoP thus acts as a global repressor of antibiotic biosynthesis in *S. coelicolor*.

By and large, *Streptomyces* secondary metabolism is believed to be a cellular response to nutrient limitation (Ochi, 2007), including the availability of carbon (reviewed in Ruiz *et al*., 2010), nitrogen (Voelker and Altaba, 2001) and phosphate (Martin, 2004**)**. The effect of carbon source on secondary metabolism appears to be closely tied to carbon catabolite repression, with streptomycin (*S. griseus*) and actinorhodin (*S. coelicolor)* production being inhibited when glucose is supplied at high concentrations (2-10%) in the growth media (Komatsu *et al*., 2003 and Takano *et al*., 2011). However, in the streptomycetes, the molecular mechanism underlying carbon catabolite repression is not fully understood (van Wezel *et al*., 2007). This is also the case for nitrogen, with the regulatory pathway mediating the effect of nitrogen source on secondary metabolism not yet completely elucidated. One compound that appears to be a key to both pathways, however, is N-acetylglucosamine (a component of both the microbial cell wall and the polymer chitin).

N-acetylglucosamine appears to have a signalling role, particularly during nutrient starvation, and it affects antibiotic production in diverse *Streptomyces* species (Rigali *et al*., 2008). In *S. coelicolor*, it has been determined that N-

acetylglucosamine is imported from the environment, through the N-acetylglucosamine major permease NagE2 to the cytoplasm, where it is phosphorylated by the phosphotransferase system. The phosphotransferase system comprises several proteins that transfer a phosphoryl group from phosphoenolpyruvate to N-acetylglucosamine. The phosphororylated N-acetylglucosamine is then deacetylated by NagA, giving glucosamine-6-phosphate, which interacts with DasR, and inhibits the DNA binding activity of DasR. DasR is a member of the GntR-family of transcription regulators, and it represses the expression of both *actII-ORF4* (Rigali *et al*., 2008) and *nagE2* (Rigali *et al*., 2006). Interestingly, the regulatory effects of DasR are counteracted by the TetR-like transcription factor AtrA, which activates the expression of both *actII-orf4* and *nagE2* (Uguru *et al*., 2005; Nothaft *et al*., 2010). Therefore, the presence of N-acetylglucosamine leads to the activation of antibiotic production, although interestingly, this stimulatory effect is only observed in the absence of glucose (i.e. not during growth on rich medium) (Rigali *et al.,* 2008).

In addition to responding to N-acetylglucosamine in the environment, the streptomycetes also produce and secrete small signalling molecules in a growth dependent manner. These molecules include both γ-butyrolactones and furans, and they have regulatory roles in triggering the production of such antibiotics as streptomycin (A-factor, γ-butyrolactone), actinorhodin and undecylprodigiosin (*S. coelicolor* butanolide 1 [SCB1], γ-butyrolactone), virginiamycin (virginiae

butanolides, γ-butyrolactone), methylenomycin (methylenomycin furan) (reviewed in Takano, 2006; Corre *et al*., 2008). The best understood γ-butyrolactone is A-factor or 'autoregulatory factor' from *S. griseus*. A-factor is synthesized by AfsA (Kato *et al*., 2007), and A-factor associates with ArpA, a repressor protein that binds upstream of the global regulatory gene *adpA*. Once A-factor binds ArpA, it changes the conformation of ArpA, relieving the repression of *adpA* expression (Ohnishi *et al*., 1999). AdpA subsequently activates the transcription of *strR*, a pathway-specific regulator for streptomycin biosynthesis (Ohnishi *et al*., 1999 and Tomono, *et al*., 2005), along with a number of other genes whose products play important roles in morphological differentiation.

A second well-studied γ-butyrolactone is SCB1. The SCB1 receptor protein, ScbR, binds upstream of of itself (*scbR*), as well as *scbA* (SCB1-encoding gene, Hsiao *et al*., 2007), repressing its own expression, but activating that of *scbA* (Takano *et al*., 2001). The DNA binding capacity of ScbR is eliminated following association with SCB1 (Takano *et al*., 2001). Other regulatory targets of ScbR include the upstream region of *kasO*, which encodes a pathway-specific regulator for the cryptic type I polyketide biosynthetic gene cluster (Takano *et al*., 2005). ScbR binding represses *kasO* expression, and its binding with SCB1 relieves *kasO* repression. Expression of *scbA* and *scbR* also affects actinorhodin and undecylprodigiosin production, but the mechanism underlying this control is not yet understood (Takano, 2006). A-factor and SCB1 represent two of the best

studied γ-butyrolactone variants in the streptomycetes; however, these organisms also produce many other γ-butyrolactones whose functions have yet to be discerned (Hsiao *et al*., 2009).

Secondary metabolism has also been reported to be triggered by other environmental factors, such as temperature and pH (Deeble, *et al*., 1995; Kiviharju *et al*., 2004). However, again, we are still far from understanding the mechanisms that connect these environmental factors to secondary metabolism. Clearly, there is much that remains to be discovered about the regulatory cascades governing secondary metabolism. Gene expression is affected by the efficiency of transcription, which is controlled by the many regulatory proteins discussed above. Moreover, gene expression can also be affected by the stability of the resulting messenger RNA (mRNA), efficiency of protein translation and post-translational modification/processing of the ensuing protein product. All of these different regulatory levels can be mediated by distinct regulators, which may be protein-based or RNA-mediated.

1.4. Small RNA (sRNA)

Bacterial RNA molecules are now recognized as playing an important role in regulating gene expression. One class of these regulatory RNA molecules is the small RNAs (sRNAs), which, at least in *Escherichia coli*, are considered to range in size from 50 to 300 nt (reviewed in Storz *et al*., 2011). It has been established that sRNAs exert their regulatory roles either by base pairing with target mRNAs, affecting their stability or translatability, or by binding to proteins

and modulating the activity of those proteins. Small RNAs are generally not translated into proteins; however, a well-studied sRNA, SgrS, has a "dual function", in that it interacts with target mRNAs and encodes a small regulatory peptide (Wadler and Vanderpool, 2007).

Small RNAs can be transcribed either in *cis* or in *trans* relative to their target mRNAs. *cis*-encoded sRNAs are transcribed as antisense RNA strands, at the same genetic location as their target RNAs, whereas *trans*-encoded sRNAs are transcribed elsewhere in the genome. While *trans*-encoded sRNAs are currently the most represented chromosomally-encoded sRNAs, *cis*-encoded sRNAs are commonly found to be encoded by plasmids, phages and transposons (Thomason and Storz, 2010). When associating with their cognate mRNAs, *trans*-encoded sRNAs generally exhibit limited sequence complementarity with their target mRNAs, while *cis*-encoded sRNAs and their targets share extended complementarity (reviewed in Waters and Storz, 2009).

It is widely accepted, particularly in Gram-negative bacteria, that interaction between *trans*-encoded sRNAs and their target mRNAs typically requires the activity of the RNA chaperone Hfq (Storz, 2011). Hfq homologues are, however, not found in all bacteria. Streptomycetes lack an Hfq homologue (Sun *et al*., 2002; Swiercz *et al.,* 2008), suggesting that other proteins may help to stabilize sRNAs and facilitate their interaction with cognate mRNAs. With respect to *cis*-encoded sRNAs, however, Hfq is thought to be dispensable for the interaction of these RNAs and their target mRNA (Thomason and Storz, 2010).

For the sake of clarity, I will refer to *cis*-encoded sRNAs as "antisense RNAs" throughout the remainder of my thesis.

1.5. Antisense RNAs

Antisense RNAs were initially discovered through studies of bacteriophages, plasmids, and transposons (reviewed in Thomason and Storz, 2010). These studies suggested that antisense RNAs were involved in the control of plasmid replication, the lysogen to lytic switch in phages, and transposition regulation (reviewed in Brantl, 2007). More recently, antisense RNAs have begun to be recognized as elements encoded throughout the chromosomes of numerous bacteria, including *E. coli, Bacillus subtilis*, *Helicobacter pylori*, *Listeria monocytogenes*, *Synechocystis* sp. PCC6803, and *Mycoplasma pneumonia* (Dornenburg *et al*., 2010; Selinger *et al*., 2000; Nicolas *et al*., 2012; Sharma *et al*., 2010; Mraheil, *et al*., 2011; Mitschke *et al*., 2011; Güell *et al*., 2009). Much of this recent identification has come through transcriptome (microarray and/or RNA-sequencing) analyses, and it is worth noting that antisense RNA transcripts detected using these genome-wide technologies may contain artifacts, which can be introduced during cDNA synthesis (Perocchi *et al*., 2007). Nonetheless, this widespread antisense RNA population does, however, raise questions regarding the biological relevance of these molecules. Thus far, antisense RNAs have been shown to exert their roles in many cellular processes from modulating the expression of transcription regulators and metabolic

enzymes, to repressing the production of virulence proteins and bacterial toxins (reviewed in Thomason and Storz, 2010).

1.6. Mechanisms of gene regulation mediated by antisense RNAs

As mentioned above, antisense RNAs usually target their sense mRNAs via a base-pairing mechanism. This base-pairing can occur within the coding region or within the 5′ or 3′ untranslated regions. However, the extent of complementarity required for an interaction is not yet fully understood, and can be impacted by the structure of both the antisense and mRNAs. Currently, the base-pairing steps involved in the interaction between these two RNAs are best understood in plasmid/transposon systems; whether the mechanisms are similar for chromosomally-encoded antisense RNAs remain to be determined. By and large, there are two possible mechanisms by which base-pairing can occur (reviewed in Brantl, 2007): (1) an initial contact between RNA strands takes place, followed by a spontaneous extension of base-pairing; (2) initial base-pairing occurs through the formation of a so-called "kissing complex" that is recognized by stabilizing proteins, and from this stabilized complex, additional base-pairing follows. Such base-pairing may alter the secondary structure of both interacting transcripts, and this may ultimately impact either the stability or translation of the target mRNA.

A classic example of post-transcriptional regulation by an antisense RNA is the *isiA*-IsrR sense-antisense pair in *Synechocystis* PCC6803 (Dühring *et al*.,

2006). IsrR is transcribed as an antisense RNA strand within the coding region of *isiA,* at such a position where it is unlikely to interact with the ribosome binding site (RBS) of *isiA*. Under conditions of iron starvation, oxidative stress, and high intensities of light, the expression profiles of IsrR and *isiA* are completely inverted. Modulating IsrR levels, positively by overexpression and negatively by overexpressing its complementary sequence, results in opposing effects on the transcript levels of *isiA*. From these results, it was inferred that *isiA* mRNA was destabilized in the presence of IsrR; however, the mechanism by which this occurs has not yet been elucidated.

Additional insights into sRNA-mediated regulation at a post-transcriptional level can be seen in the *sodB*-RyhB mRNA-sRNA system in *E. coli* (Massé *et al*., 2003). Under conditions of iron limitation, RyhB expression is induced and *sodB* transcript levels are correspondingly reduced. The effect of RyhB appears to be independent of RNase III, but depleting RNase E alleviates the RyhB-dependent effect on *sodB*. This suggests that RyhB-mediated *sodB* degradation requires RNase E. The involvement of RNases in antisense RNA-mediated effects is described in more detail below. Interestingly, as RhyB binds at a site overlapping the RBS of *sodB,* RyhB also exerts its regulatory role by blocking *sodB* translation (Morita et al., 2006; Prévost *et al.*, 2011).

For those sRNAs/antisense RNAs acting at translational levels, this usually involves base pairing within the 5′ end (including or excluding Shine-Dalgarno sequences) of their sense mRNAs (reviewed in Georg and Hess, 2011).

For many *trans*-encoded sRNAs, their interaction with their target mRNAs prevents mRNA binding to the ribosome, leading to decreased stability of the associated mRNA (reviewed in Waters and Storz, 2009). An example of such a mechanism can be seen with the *symE*-SymR sense-antisense pair in *E. coli*, which constitutes a toxin-antitoxin system (Kawano *et al*., 2007). The SymR antisense transcript overlaps the 5′ end of the *symE* transcript at a site that includes its Shine-Dalgarno sequence. Consistent with a role in translational regulation, overexpression of SymR negatively affected SymE expression. In addition, it also simultaneously affected the stability of *symE* mRNA, and like RyhB, did so in an RNase III-independent manner. It is worth noting that the translational repression of a target gene can be sufficient to impact expression levels without any corresponding mRNA destabilization, at least based on observations for a *trans*-encoded sRNA-mRNA pair in *E. coli* (SgrS-*ptsG*) in the absence of RNase E (Morita *et al*., 2006).

In another toxin-antitoxin system (*tisAB*-IstR-1) reported by Darfeuille *et al*. (2007), the sRNA IstR-1 is transcribed upstream of the *tisAB* transcriptional start site (and therefore is not an antisense RNA, as defined in this thesis). The full-length *tisAB* mRNA is translationally inactive, but it is processed into two forms, termed the +42 and +106 transcripts (corresponding to the first nucleotide of the *tisAB* transcript). While *tisA* is not translated, *tisB* is translated exclusively from the +42 mRNA transcript. IstR-1 binds to the +42 mRNA transcript, upstream of the *tisB* ribosomal binding site, and not only effectively prevents

translation (as determined through *in vitro* toeprinting experiments), but also stimulates RNase III cleavage of the *tisAB* mRNA in this overlapping region (when IstR-1 is present *in vitro*), leading to the generation of the +106 mRNA transcript, from which *tisB* cannot be translated. In the case of antisense RNA-mRNAs interaction that occurs downstream of the ribosomal binding site, the base-paired structure is thought to be sufficient to slow down the elongation of translation (Georg and Hess, 2011).

As described above, the antisense RNA-sense mRNA interaction is often associated with a corresponding negative regulation of gene expression (Dühring *et al*., 2006; Kawano *et al*., 2007; Han *et al*., 2010). There are a few examples of sRNAs that enhance target gene expression (Opdyke *et al*., 2004; Zemanová *et al*., 2008; Majdalani *et al*., 2002). *rpoS* translation, for example, is positively affected by two independent sRNAs (RprA and DasR) that are expressed under different conditions (Majdalani *et al*., 2002; Majdalani *et al*., 1998) The leader region of *rpoS* forms a stem-loop structure that occludes the ribosome binding site, thus preventing *rpoS* translation. As determined by compensatory mutation analysis, DasR and RprA can bind to the inhibitory part of the leader region, leaving the ribosome binding site accessible for ribosomes to initiate *rpoS* translation.

A further mechanism of antisense RNA-mediated gene regulation occurs at the level of transcription termination. This mechanism has been observed for the virulence plasmid pJM1 in *Vibrio anguillarum* (Stork *et al*., 2007). This

plasmid encodes an antisense RNA, RNAβ, on the strand opposite the *fatDCBAangRT* operon, where the RNAβ transcript overlaps with the 5′ end of *angR* and the 3′ end of *fatA*. Binding occurs through stem-loops predicted to be present on both transcripts and serves to terminate transcription within the *fatA-angR* intergenic region. Notably, none of the stem loop structures form a Rho-independent-like terminator. Although the mechanism by which the stem loop interaction can terminate the operonic transcription is unknown, this work implies that base-pairing formation mediates the regulatory effect of RNAβ on the expression of *fatDCBAangRT*.

As can be seen from the examples above, most of the characterized antisense RNAs act through base-pairing with their target mRNAs. Additionally, they can also affect the expression of their targets through transcription interference. This mechanism is best exemplified with the three gene-operon *ubiGmccBA* in *Clostridium acetobutylicum* (André *et al*., 2008). A number of antisense RNAs are expressed from a transcriptional start site positioned downstream of the 3′ end of the *ubiGmccBA* operon, resulting in up to 700 nt-overlapping sequence shared between the antisense RNAs and *mccA* mRNA. Interestingly, when these antisense RNAs are expressed in *trans* (from a plasmid), they fail to affect the expression of *ubiGmccBA*, but negatively affect the operon expression when they are expressed in *cis*. While an indepth mechanistic understanding has not yet been achieved, it has been proposed that transcription machineries of the antisense RNA and the sense polycistronic mRNA *ubiGmccBA*

interfere with one another by colliding within the overlapping region. Alternatively, transcription of the antisense RNAs may introduce positive supercoiling into the downstream region, ahead of the transcribing RNA polymerase, and this altered DNA structure may inhibit mRNA transcription (Moulin *et al*., 2005; El Hanafi and Bossi, 2000).

Overall, antisense RNAs can affect the expression of their partner mRNAs at transcriptional, post-transcriptional, and translational levels, and this regulation may have positive or negative effects.

1.7. The contribution of ribonucleases (RNases) to antisense RNA-mediated gene regulation

Messenger RNA stability – or lack thereof - is often associated with RNase activity, and the accessibility of an mRNA for particular RNases can be either promoted or obstructed by the base-pairing of an antisense RNA to its sense mRNA. It is widely accepted that, based on studies of *trans*-encoded sRNAs, RNase E makes a remarkable contribution to the functioning of sRNAs in affecting mRNA stability (reviewed in Storz, 2011). In *E. coli,* RNase E is an integral part of the RNA degradosome (reviewed in Carpousis, 2007). To date, however, RNase E has been shown to have a role in antisense RNA-dependent mRNA stability in only a few instances. One such example centres on the AmgR antisense RNA in *Salmonella enterica* (Lee and Groisman, 2010). AmgR overlaps with the *mgtC* portion of the *mgtCBR* polycistronic mRNA, and appears

to promote the destabilization of *mgtC* in an RNase E-, but not RNase III-, dependent manner.

RNase E has a demonstrated preference for RNA substrates having a 5′ monophosphorylated end, at least in *E. coli* (Celesnik *et al.*, 2007), while RNase III targets double-stranded RNAs. Such double-stranded RNA substrates could be formed following the binding of antisense RNAs to their sense RNAs. In *E. coli,* RNase III is involved in GadY antisense RNA-dependent cleavage of the *gadX-gadW* dicistronic mRNA (Opdyke *et al*., 2011). This cleavage leads to the accumulation of individual *gadX* and *gadW* transcripts. It has been proposed that the *gadX-gadW* cleavage event could lead to "protective" folding of *gadX* and *gadW* transcripts, where they are more resistance to further degradation than the full-length *gadX-gadW* transcript. A second example can be seen for the novel RyeA-RyeB pair in *E. coli* (Vogel *et al*., 2003). Both RyeA and RyeB are sRNAs, and when they are bound to each other, they are substrates for RNase III, which appears to cleave both RNA strands simultaneously, resulting in smaller stable products of each.

These examples suggest that RNase III may exert its activity by processing "antisense RNA-sense RNA" double stranded molecules, but the situation may vary depending upon the organism. Recently, RNase III has been shown, through high resolution tiling arrays of RNase depletion strains, to have a small role in cleaving antisense RNAs in *B. subtilis*, where instead, a more dominant role is observed for RNase Y (Durand *et al*., 2012). RNase Y is an

endoribonuclease that resembles the *E. coli* RNase E in terms of its preference of processing 5′ monophosphorylated substrates and its substantial role in mRNA decay (Shahbabian *et al*., 2009). It is worth noting that *B. subtilis* does not encode an RNase E homologue, and that RNase Y is not found in *E. coli.*

It appears that *S. coelicolor* does not possess an RNase Y orthologue, but it does have an RNase E orthologue encoded by *SCO2599*. This protein, known as RNase ES, can suppress the lethal effect of an RNase E mutation in *E. coli*, suggesting that RNase ES has similar functional activities as RNase E in *E. coli* (Inagawa *et al.*, 2003). However, they exhibit some differences in their enzymatic activities. While they have been shown to cleave the same rRNA substrate (9S rRNA), the resulting products (5S rRNA) differ in size in each organism (Lee and Cohen, 2003). Moreover, RNase E and RNase ES have different arrangements of their functional domains (Lee and Cohen, 2003). In *S. lividans,* RNase ES expression (or activity) increases during cellular growth, and this increase requires the function of the *bldA* gene product, the $tRNA_{Leu}$-specifying the UUA anticodon (Hagège and Cohen, 1997). While there is a report describing the knockout of the RNase ES-encoding gene in *S. coelicolor* (Lee and Cohen, 2003), attempts to disrupt the RNase E-encoding gene in *S. coelicolor* in our lab have so far been unsuccessful, suggesting that RNase E may be essential for *S. coelicolor* growth.

In contrast to RNase E, RNase III is not essential in *S. coelicolor*. This enzyme cleaves double-stranded RNA (Chang *et al.*, 2005), and mutants (in both

E. coli and *S. coelicolor*) accumulate 30S rRNA precursors (Babitzke *et al.*, 1993; Price *et al.*, 1999), suggesting that RNase III plays a role in converting 30S rRNA to 23S and 16S rRNA in both bacteria. RNase III also affects the abundance of ~200 mRNA transcripts (Huang *et al.*, 2005), including its own mRNA transcript (Xu *et al.*, 2008) and transcripts of genes in the characterized antibiotic biosynthetic clusters (Huang *et al.*, 2005**)**. Although it is well established that *S. coelicolor* strains lacking a functional RNase III are impaired in antibiotic production (Adamidis and Champness, 1992; Gravenbeek and Jones, 2008; Sello and Buttner, 2008), the target(s) of RNase III that contribute to the control of antibiotic production have yet to be identified.

The regulatory effects of RNase III can, in some instances, be mediated by AdpA (Xu *et al*., 2010), a pleiotropic regulatory protein controlling morphological development and antibiotic production in *S. griseus* (Ohnishi *et al*., 2005). Transcripts of *adpA* mRNA were more abundant in a strain lacking functional RNAse III (*absB*/*rnc* mutant) and it appears to be cleaved by RNase III (Xu *et al*., 2010). In *S. coelicolor,* AdpA is encoded by the *bldH/SCO2792* gene, and this protein is a determinant for aerial hyphae formation (Takano *et al.*, 2003). Recently, RNA-Seq and RNA immunoprecipitation analyses have revealed 777 potential RNase III target-genes, with 79 transcripts showing expression altered (more than 2 fold) in an *rnc* deletion mutant (Gatewood *et al*., 2012). Their work also revealed an *rnc* mutation to affect the expression of two genes encoding previously confirmed-sRNAs, scr2101 (Swiercz *et al*., 2008) and scr6925

(Vockenhuber *et al*., 2011). These findings suggest that RNase III is an important player in RNA metabolism in *S. coelicolor*.

1.8. The goal of this thesis

Given the increasing number of newly identified antisense RNAs in bacteria, including in the streptomycetes, and the fact that our understanding of the physiological relevance of these RNA molecules is still in its infancy, this thesis is aimed at dissecting the biological roles of two antisense RNAs (α-abeA and scr4677) and their associated targets in *S. coelicolor*. Chapter 3 describes a genetic analysis of the *SCO4677-4675* locus, which harbours the antisense RNA-encoding gene *scr4677*. This analysis identified a role for scr4677 in modulating the gene expression of *SCO4676*, at least in part through its potential contribution to destabilizing the *SCO4676* transcript. Chapter 4 outlines a comprehensive genetic and biochemical analysis of the *SCO3287-3291* (*abeABCDR*) genetic region containing the α-abeA-encoding gene. These analyses revealed that these genes were differentially regulated by RNase III and a previously uncharacterized SARP-like protein encoded by *abeR,* and that the products of this *abeABCDR* locus affect secondary metabolism.

CHAPTER 2

MATERIALS AND METHODS

2.1. Bacterial strains and their cultivation

Streptomyces coelicolor and *Escherichia coli* strains used in this study are summarized in Table 2.1. *Streptomyces* strains were grown at 30°C on solid agar and in liquid media. The composition of these media is described below. During *Streptomyces* liquid culturing, strains were grown in "universal bottles" or in baffled flasks with a stainless steel spring. Unless otherwise indicated, *E. coli* strains were grown at 37°C in Luria-Bertani (LB) or SOB medium (Sambrook and Russell, 2001). Liquid cultures were shaken at 200 RPM. For strain selection or maintaining plasmid selection, antibiotics were added into growth media at the final concentration indicated (μg/mL): ampicillin, 100; apramycin, 50; chloramphenicol, 25; hygromycin B, 50; kanamycin, 50; nalidixic acid, 25; thiostrepton, 50; and viomycin, 50, unless otherwise indicated.

2.2. DNA introduction into *E. coli* and *S. coelicolor*

2.2.1. DNA introduction into E. coli

Plasmids or cosmids used in this study are summarized in Table 2.2. Competent *E. coli* strain DH5α (unless specified differently) was prepared by a series of magnesium chloride and calcium chloride washes, or purchased (subcloning-efficiency DH5α from Invitrogen), and plasmids were introduced

through heat shock at 42 or 37°C, respectively. Cosmids were introduced into *E. coli* strains by electroporation (MicroPulser™, Bio-Rad) at ~2.48 kV in ice-cold cuvettes. Protocols for preparing competent cells, transformation using heat shock and electroporation are all described in Sambrook and Russell (2001). The transformants were screened by colony PCR, in which a small fraction of a colony was used as template for amplification of the DNA fragment of interest. Positive clones were then grown up in liquid culture, and plasmids were isolated using alkaline lysis with phenol-chloroform extraction or using a silica-based purification kit (PureLink™ Quick Plasmid Miniprep kit).

2.2.2. DNA introduction into S. coelicolor

Streptomyces coelicolor possesses a methyl-specific restriction system (González-Cerón *et al.*, 2009); hence, the plasmids/cosmids isolated from methylating *E. coli* strains (such as DH5α) were first introduced into a non-methylating *E. coli* strain (ET12567) carrying pUZ8002, a plasmid that facilitates the conjugal transfer of *oriT*-containing plasmids/cosmids from *E. coli* to *Streptomyces* strains (Kieser *et al*., 2000). For plasmids/cosmids lacking an *oriT,* DNA isolated from *E. coli* ET12567 was introduced into *S. coelicolor* strains through protoplast transformation. Both conjugation and protoplast transformation procedures were conducted as described in Kieser *et al*. (2000).

2.3. Construction of null mutations in *S. coelicolor*

Null mutant strains listed in Table 2.1 were constructed using Redirect© technology (Gust *et al*., 2003). Using cosmids (Table 2.2) from an *S. coelicolor* cosmid DNA library (Redenbach *et al*., 1996), single genes or operons were individually replaced by an apramycin resistance cassette (*aac(3)IV*) through double-crossover homologous recombination. The one exception to this was when a viomycin resistance (*vph*) cassette was used to replace *adpA* (*SCO2792*) in an RNase III (*rnc*) mutant that already contained an apramycin resistance cassette (Sello and Buttner, 2008). Resistance cassettes, which were flanked by target gene-specific sequences, were generated by PCR amplification using iProof™ high fidelity-DNA polymerase (Bio-Rad). All oligonucleotides used for PCR amplification are summarized in Table 2.3 (Appendix).

To increase the likelihood of DNA recombination, wild type cosmids of interest were introduced through electroporation into *E. coli* BW25113 (Datsenko and Wanner, 2000), which carries the L-arabinose-inducible *λ*RED genes on pIJ790 (Table 2.2). Once the cosmid was confirmed in *E. coli* BW25113, a subsequent electroporation was conducted to introduce the PCR amplified resistance cassette into this strain, with the aim being to replace the gene/region of interest. Successful mutations were verified by PCR analysis and use of diagnostic restriction enzymes. For example, to confirm that *SCO4677* had been disrupted, both wild type and mutant cosmids were digested using *Bam*HI and *Xho*I, and their resulting digestion profiles were compared following agarose gel

electrophoresis. In addition, PCR analyses were conducted using an oligonucleotide pair 4676HOT (anneals within the *SCO4676* coding region) and KO4677-1 (a knockout primer) and another oligonucleotide pair 4677-1IN (anneals within the *SCO4677* coding region) and 4676HOT (please see Table 2.3 in Appendix for oligonucleotide sequences and description of their use). At this stage, a single colony may possess a mixed population of both wild type and mutant cosmids; however, only *E. coli* cells carrying at least one copy of the mutant cosmid will survive in the presence of appropriate antibiotic selection (apramycin or viomycin).

The resistance cassettes described above have been engineered to include an origin of transfer (*oriT*), and thus the mutant cosmids can be transferred into *S. coelicolor* through conjugation from *E. coli*, followed by selection for a double-crossover recombination event. To confirm gene replacement in the chromosome with the resistance cassette, chromosomal DNA (or a small portion of colony biomass) was isolated, and used as template in PCR checks following the strategies described above for verifying the mutant cosmid.

In instances where gene replacement might affect the transcription of downstream genes (polar effects), the resistance cassette, which is flanked by FRT sites, was excised from the genome (Fedoryshyn *et al*., 2008). Briefly, *E. coli* ET12567 carrying pUWLFLP (a plasmid bearing a Flp recombinase-encoding gene and an *oriT* sequence) was conjugated with *S. coelicolor* mutant strains (*e.g. SCO4676* deletion strain). The resulting exconjugants were spread on

selective MS plates, followed by screening for sensitivity to nalidixic acid and apramycin. To verify that the resistance cassette was no longer present in the chromosome, chromosomal DNA was isolated from colonies that appeared to be 'unmarked', and was analyzed by PCR (*e.g.* for the *SCO4676* deletion, this involved using oligonucleotides KO4676-1IN and 4675-2IN [Appendix, Table 2.3]).

2.4. Complementation of null mutant strains

2.4.1. Generating an abeR *complementation strain*

To generate an *abeR* (*SCO3291*) complementation construct, wild type *abeR* was excised from pMC118 (Table 2.2) using *Kpn*I and *Hin*dIII, gel purified using a MinElute® gel extraction kit (Qiagen), and ligated into pMS82, an integrating plasmid vector (Gregory *et al.*, 2003). The resulting clone was designated pMC122. This plasmid, along with pMS82 alone, was passaged through the non-methylating *E. coli* ET12567/pUZ8002 strain prior to being introduced into the *abeR* deletion mutant E304 (and wild type strain M145) through conjugation. To select for exconjugants of interest, "conjugation" plates were overlaid with 1 mL of sterile water containing nalidixic acid and hygromycin (antibiotic concentrations are listed in section 2.1), followed by gentle shaking to spread the solution evenly on the surface, and air drying. After 3-4 days incubation, the exconjugants were streaked on selective media to confirm the antibiotic resistance of the strains.

2.4.2. Generating a SCO4676 *complementation strain*

A similar approach to that described in section 2.4.1 was undertaken to complement the *SCO4676* deletion mutant (strain E320, Table 2.1). Wild type *SCO4676*, together with 261 bp upstream and 211 bp downstream sequence, was PCR amplified using phosphorylated oligonucleotides 4677-3E and 4676-2IN (Appendix, Table 2.3). The amplified fragment was gel extracted and ligated with pIJ2925 digested with *Sma*I. The resulting plasmid clone (pMC145) was sequenced to confirm the integrity of *SCO4676* using either M13F or M13R (Appendix, Table 2.3). *SCO4676* was then excised from pMC145 using *Bgl*II and introduced into *Bam*HI-digested pIJ82 (integrating plasmid) (gift from H. Kieser). The resulting construct (pMC146) was introduced into *E. coli* ET12567, and conjugated into the *SCO4676* deletion mutant, followed by selection for plasmid integration into the chromosome.

2.5. Construction of *Streptomyces* overexpression strains.

DNA sequences destined for overexpression were PCR amplified using *Pfu* DNA polymerase (Stratagene) or iProof™ high fidelity-DNA polymerase (Bio-Rad). For the *abe* gene cluster, however, we took advantage of unique flanking restriction enzyme sites in the *abe*-containing cosmid, to generate an appropriate overexpression fragment. These fragments were all initially cloned into pIJ2925 (Table 2.2), excised and cloned into pWHM3, a high-copy-number plasmid in both *E. coli* and *Streptomyces* (Vara *et al*., 1989). These clones were

introduced into *E. coli* ET12567/pUZ8002, and the plasmid was introduced into wild type or mutant *S. coelicolor* strains by protoplast transformation.

2.5.1. Generating abeABCDR, abeABCD, *and* abeR *overexpression strains*

For *abeABCD* and *abeR* (*SCO3287-3290* and *SCO3291*, respectively) overexpression constructs, cosmid E15 (Table 2.2) was digested using *Kpn*I, to liberate an ~6.5 kb DNA fragment containing the *abeABCDR* locus with ~550 bp flanking regions, which was then introduced into pIJ2925, generating pMC123. The insert was then excised from pMC123 using *Bgl*II, and cloned into the *Bam*HI site of pWHM3, giving pMC120 – a multicopy *abeABCDR*-containing construct. To create an *abeABCD* overexpression construct, pMC123 was digested with *Bam*HI and the resulting ~4.5 kb fragment was ligated together with *Bam*HI-digested pWHM3, generating pMC116. Construction of the *abeR* overexpression construct involved PCR amplification of the wild type *abeR* gene (together with 231 bp and 219 bp upstream and downstream flanking regions, respectively) using oligonucleotides OE3291-1 and OE3291-2 (Appendix, Table 2.3), followed by blunt-end cloning into pIJ2925 digested with *Sma*I, to give pMC124. The *abeR* fragment was excised from pMC124 using *Bgl*II and *Eco*RI, gel-purified, and inserted into *Bam*HI and *Eco*RI sites of pWHM3, generating pMC118.

2.5.2. Generating antisense RNA-overexpression strains

To create overexpression constructs for α –3287 and scr4677 antisense RNAs, their corresponding genes were PCR-amplified using pairs of oligonucleotides [OEα 3287-1 and OEα 3287-2, and 4677-3E and 4676HOT, respectively (Appendix, Table 2.3)]. The resulting DNA fragments were gel-purified and inserted into the *Sma*I site of pIJ2925, giving pMC147 (α –*3287*) and pMC148 (*scr4677*). After the insert sequences were verified by sequencing, the recombinant plasmids were digested using *Bam*HI and *Eco*RI (or *Bgl*II for *scr4677*) to liberate the sRNA genes, before being introduced into pWHM3, generating pMC117 and pMC149 as α –*3287* and *scr4677* overexpression plasmids, respectively.

2.6. Phenotypic comparisons of wild type, mutant and overexpression strains of *S. coelicolor*

Two independent spores stocks of all strains were used for all phenotypic analyses; spore stocks were prepared as described in Kieser *et al*. (2000).

2.6.1. Comparison of antibiotic production and morphological development

Approximately 10^6 spores were streaked out on a variety of solid media, including (unless otherwise specified): minimal medium with glucose or mannitol as a carbon source, rich medium (R2YE with glucose or mannitol), soy flour medium with glucose (SFG) or mannitol (SFM), supplemented minimal medium solid (SMMS), and LB (Sambrook and Russell, 2001). The composition of solid

media (other than LB) used for phenotypic comparison is described in Kieser *et al*. (2000) with carbon source variation as follows: (1) minimal medium containing 1% (w/v) glucose or 0.5% (w/v) mannitol; (2) R2YE rich medium containing 10.12 g glucose or mannitol per 0.8 liter of distilled water; (3) soy flour medium containing 2% (w/v) mannitol or glucose. Petri plates were incubated at 30°C. Pigmented antibiotic production and morphological development were compared among strains over a week-long time course. These experiments were conducted at least three times. Morphological development was examined by light microscopy, observing coverslip impressions of *Streptomyces* colonies.

Phenotypic comparisons of liquid-grown cultures (specifically, differences in pigmentation/antibiotic production) were conducted using cultures grown in "universal bottles". Cultures of 7.5-10 mL, inoculated with ~10^7 spores/mL, were incubated at 30°C and shaken at 200 RPM. The liquid media used included: minimal liquid medium (NMMP) and rich medium (R5), both supplemented with glucose or mannitol; yeast extract-malt extract (YEME); tryptone soy broth (TSB); and supplemented liquid minimal medium (SMM). The liquid media were supplemented with different carbon sources, as follows: (1) NMMP supplemented with 0.5% (w/v) glucose or mannitol; (2) R5 supplemented with 1% (w/v) glucose or 0.5% (w/v) mannitol.

2.6.2. Quantification of actinorhodin and undecylprodigiosin titers in liquid media

Prior to assessing the production of actinorhodin and undecylprodigiosin during liquid culturing, *S. coelicolor* spores were pre-germinated (Kieser *et al.*, 2000). Briefly, approximately 10^8 spores were inoculated into 2.5 mL TES buffer (0.05 M, pH 8.0) and heated at 50°C for 10 min. Into this spore suspension, 2.5 mL of double strength germination medium (1% [w/v] yeast extract, 1% [w/v] casaminoacids, and 10 mM calcium chloride) was added, followed by shaking at 37°C for 3 h with a rotation speed of 200 RPM. Germinated spores were then inoculated into R5 medium to an optical density at 450 nm of 0.04 to 0.08. Cultures (25 or 50 mL) in 250 mL baffled flasks were then incubated at 30°C in a shaker. At specified time points, duplicate 0.5 mL samples were spun down to separate cells (for assessing wet weight and undecylprodigiosin levels) and culture supernatant (for actinorhodin assays). Actinorhodin and undecylprodigiosin were quantified using visible light spectroscopy as described in Kang *et al.* (1998) and Kieser *et al.* (2000), respectively. To quantify total actinorhodin, an equal volume of 2 M KOH (or NaOH) was added to the culture supernatant, to give a final concentration of 1 M, followed by absorbance measurement of the resulting solution at 633 nm. For undecylprodigiosin quantification, undecylprodigiosin was extracted by shaking cell pellets in acidified methanol overnight. After centrifugation, the absorbance of supernatant was measured at 530 nm.

2.6.3. Calcium-dependent antibiotic (CDA) bioassay

The evaluation of CDA production followed the method described in Kieser *et al.* (2000), with minor adjustments. Approximately, 10^7 to 10^8 spores were spread on nutrient agar plates, leading to confluent growth after 24 or 48 h incubation at 30°C. From these confluent areas, agar plugs were excised and transferred into empty plugs (equivalent in sizes to the agar plugs) on fresh nutrient agar plates with and without 12 mM calcium chloride, to determine whether any zones of growth inhibition (ZI) required calcium supplementation. These plates were overlaid with soft nutrient agar (1.15% [w/v] nutrient agar, 0.4% [w/v] nutrient broth) that had been previously inoculated with 1/100 volume of overnight *S. aureus* ATCC 29213 cultures. ZIs were measured after overnight incubation of the assay plates at 37°C.

2.6.4. Assessing sensitivity of S. coelicolor *strains to various stresses*

Spores of different *S. coelicolor* strains were subjected to various stress conditions. This included exposure to heat (60°C, Haiser *et al*., 2009) for a period of time up to 15 min, 1% (v/v) SDS, 50 mM EDTA, 30 mg/mL lysozyme, 16 mg/mL vancomycin, 0.1 M dipyridyl, 0.05 M diamide, and 0.1 and 0.25% (v/v) hydrogen peroxide. Briefly, soft nutrient agar (0.4 g nutrient broth and 1.15 g nutrient agar per 100 mL medium) was inoculated with approximately 10^6 spores. This mixture was overlaid onto nutrient agar plates. Onto the solidified soft agar, paper disks saturated with 20 µl of each compound were placed onto the medium

surface, followed by incubation at 30°C for 1 and 2 days. The diameter of any ZI was measured on both days.

2.7. Lux reporter assays

To measure promoter activities of *actII-ORF4* and *actIII*, these promoter regions were fused to a luciferase gene cluster (kindly provided by Y. Xu and J. Nodwell) for use as a transcriptional reporter system. These constructs, including a promoterless reporter construct (pMU1, Craney *et al*., 2007), were introduced by conjugation into wild type strains carrying the high-copy plasmid pWHM3, pWHM3 + *abeR*, and pWHM3 + *abeABCD*. Spores of these strains (10^5 to 10^6) were spotted onto 0.2 mL R2YE (with thiostrepton) agar plugs in 96-flat bottom well plates (Microfluor white plate, Thermo Scientific).

Luminescence levels were assessed during their growth at 30°C (time points are indicated in Fig. 3.3) after exposing the cultures to 1% decanal, which is a luciferase substrate. For each time point, 12 readings (corresponding to 12 wells per strain) were taken, and background levels (from the promoterless pMU1-carrying strains) were subtracted from the readings obtained from the *act* promoter-fused pMU1-carrying strains. Three independent time courses were conducted, and the results were analyzed using the nonparametric Mann-Whitney test in the SPSS v13.0 statistical analysis software. This analysis was designed to test whether luminescence levels, which represent *actII-ORF4* or *actIII* transcript levels, were significantly different in the overexpression strains versus the empty pWHM3 carrying-control strain.

2.8. RNA isolation from *S. coelicolor*

RNA isolation was performed using a protocol described in Kieser *et al*. (2000), with some modification. Briefly, between 5×10^6 to 4×10^7 spores were inoculated onto different solid (R2YE with glucose, SFG, and minimal medium with mannitol) media overlaid with cellophane discs. Cells were collected at various time points, followed by adding 5 mL modified Kirby mix (1% [w/v] sodium N-lauroylsarcosinate, 6% [w/v] sodium-4-amino salycilate, and 6% [v/v] saturated phenol adjusted to pH 7.9 using 50 mM Tris-HCl pH 8.3) and 6-8 glass beads (4 mm diameter). The cells were mechanically disrupted by vigorous vortexing for 2 min, followed by two rounds of 5 mL phenol-chloroform (saturated phenol : chloroform : isoamyl alcohol = 50 : 50 : 1) extraction in which the suspensions were vortexed for a further 25 seconds. Total nucleic acid and phenol-chloroform mixtures were separated by centrifugation at 6,793 ×g for 5 min.

Total nucleic acid was precipitated by adding 0.1 volume 3 M sodium acetate (pH 6) and 1 volume isopropanol. The resulting pellets were re-suspended in double autoclaved distilled water and subjected to DNase I digestion (3 rounds of 15-20 min digestion using 10 U per round). This was followed by another phenol-chloroform extraction, after which RNA samples were precipitated in isopropanol and re-suspended in double autoclaved milli-Q water. To assess the quantity and quality of RNA samples, UV spectroscopy and agarose gel electrophoresis were conducted, respectively.

2.9. Northern blotting

Northern blot analyses were conducted as outlined in Haiser *et al*. (2008) with a few modifications. Briefly, total RNA (20-30 µg) was loaded into TBE (Tris-Borate-EDTA) or MOPS (3-N-morpholinopropanesulfonic acid)-based polyacrylamide gels, and run using 1 × TBE buffer or 1 × MOPS buffer pH 7.0 (20 mM MOPS, 5 mM sodium acetate, 1 mM EDTA, and 1 mM EGTA), respectively. The size fractionated RNA was then transferred to Zeta probe (charged) nylon membranes (Bio-Rad) or uncharged nylon membranes (Amersham Biosciences) using a semi-dry transfer apparatus (Trans-Blot® Semi-Dry, Bio-Rad) at 15 V for 40 min. RNA was crosslinked to the membranes using a chemical crosslinking method developed by Pall and Hamilton (2008). Crosslinking required two sheets of 3MM chromatography filter papers (Whatman) to be saturated with freshly made crosslinking solution (1-ethyl-3-3-dimethylaminopropyl carbodiimide or EDC). Making this solution involved mixing 0.753 g EDC, 245 µL 1-methylimidazole, 300 µL 1N HCl, and 24 mL distilled water. The transfer membranes, with the 'RNA-side' facing up, were placed onto the saturated filter papers, and these layers were wrapped with plastic wrap. This transfer process was carried out at 55°C for 2 h, after which the membranes were rinsed using 1 × MOPS buffer or distilled water.

To generate appropriate hybridization probes, oligonucleotides were phosphorylated using 10 U T4 polynucleotide kinase (Invitrogen) as per the manufacturer's instructions, with 4.76 pmol [γ-^{32}P]ATP (Perkin Elmer) in 20 µL-

reactions, followed by purification using the MinElute® Reaction Cleanup kit (Qiagen). Either 5′ end-radiolabeled locked nucleic acid (LNA) oligonucleotides or regular oligonucleotides were hybridized overnight with the crosslinked membranes in hybridization buffer (Ultrahyb®-oligo, Ambion) at 45 or 42°C, respectively. Prior to autoradiography or exposure to phosphor screens, the membranes were washed twice using low stringency buffer (2 x SSC and 0.1% [v/v] SDS) for 5 min each, followed by two rounds of washes using high stringency buffer (0.2 × SSC and 0.1% [v/v] SDS) for 7 min each. Signals (bands) were captured using a phosphorimager (Molecular dynamics). Bound probes were stripped by washing the membranes twice using the high stringency buffer for 1 h each round. If an LNA-based probe was used, the membranes were subjected to washing with previously boiled hot buffer (high stringency) for 5 min. As a control for RNA levels and integrity, 5S rRNA transcript levels were assessed after stripping previously bound probes.

2.10. Mapping 5′ and 3′-ends of RNA transcripts

S1 nuclease mapping and rapid amplification of cDNA ends (RACE) were carried out to map 5′-ends of mRNA transcripts and both 5′ and 3′ ends of antisense RNA transcripts, respectively. S1 nuclease mapping assays were carried out as described in Elliot *et al*. (2003). Probes were generated by PCR amplification with one radiolabeled oligonucleotide. For example, the *abeB* (*SCO3288*) probe was amplified using primers M13F and radiolabeled 3288-R (Appendix, Table 2.3), together with pMC125 (Table 2.2) as template DNA, after

which the PCR products were purified using a MinElute® Reaction Cleanup kit (Qiagen). Forty micrograms of total RNA, isolated from the *abeABCD* (*SCO3287-3290*) overexpression strain grown on R2YE plates, was incubated with the purified *abeB* probe at 70°C for 20-25 min, followed by hybridization at 45°C for 4 h. S1 nuclease (100 U) was used to digest unprotected single stranded nucleic acids in S1 digestion buffer (Kieser *et al*., 2000) at 37°C for 40 min, followed by isopropanol precipitation. The resulting pellets were re-suspended in S1 loading buffer (80% [v/v] formamide, 10 mM NaOH, 1 mM EDTA, 0.1% [w/v] xylene cyanol, 0.1% [w/v] bromophenol blue]. The DNA suspensions were then denatured by heating at 95°C for 5 min and separated on 6% denaturing polyacrylamide gels (Gene Gel®, Bioshop).

To map the ends of antisense RNA transcripts, the FirstChoice® RLM-RACE kit (Ambion) was used with a few modifications from what was described in Swiercz *et al*. (2008). For 5′-RACE, briefly, 10 µg of total RNA harvested from the α*–3287* overexpression strain at 40 and 67 h was treated with calf intestine alkaline phosphatase and tobacco acid phosphatase and ligated to a 5′-RACE adapter using T4 RNA ligase. The ligation products served as templates for reverse transcription using SuperScript® III reverse transcriptase (Invitrogen), and the resulting cDNA was used for subsequent PCR amplification using gene-specific and 5′ adapter-specific primers (Table 2.2). The PCR products served as templates for a subsequent nested PCR amplification using other 5′ adaptor-specific and gene-specific primers (Table 2.2). Products of the latter PCR

amplification were separated on 2% high resolution agarose gels (Bioshop), and the resolved bands that were absent from the 'no template' controls were extracted using MinElute® gel extraction kit (Qiagen). The gel-purified DNA was inserted into pCR®2.1-TOPO vector (Invitrogen) and subsequently introduced into One Shot® DH5α -TOP10F′ competent *E. coli* cells (Invitrogen). To firstly verify the presence of insert in plasmid clones, single colonies of *E. coli* transformants were subjected to colony PCR using M13F and M13R primers (Appendix, Table 2.3), after which plasmid DNA was isolated using QIAprep® Spin Miniprep kit (Qiagen) and sequenced.

As described for the 5′-RACE, a similar approach was used to map the 3′ ends of antisense RNA transcripts, with several exceptions. Twenty micrograms of total RNA were first treated with 5 U of poly(A) polymerase (Ambion) at 37°C for 1 h. The polyadenylated RNA samples were then reverse transcribed at 42°C for 50 min using M-MLV reverse transcriptase (Ambion) and the 3′-RACE specific primers (Appendix, Table 2.3).

2.11. Semi quantitative reverse transcription – polymerase chain reaction (RT-PCR)

Each RT reaction contained 2 pmol reverse primers (if two primers were used for two different transcripts, 1 pmol was used for each primer to minimize non-specific amplification in the subsequent PCR), dNTPs (10 nmol of each nucleotide), 5 µg total RNA (or 3 µg for qPCR experiments), and RNase-free water, in a total volume of 12 µL. RT samples were heated at 95°C for 10 min,

followed by immediate cooling on ice. First strand buffer, DTT, RNase OUT and SuperScript® III reverse transcriptase (Invitrogen) were added to RT samples as per the manufacturer's instructions and was incubated at 42°C for 50 min, before being stopped by heating to 70°C for 15 min. The RT products were used as templates for PCR amplification. In general, PCR programs comprised an initial denaturation step (95°C, 7 min); 15-32 amplification cycles of: 95°C for 45 sec, 58-62°C for 30 sec (depending on primer annealing temperatures), 72°C for 30-45 sec (depending on product sizes); and a final extension step (72°C, 5 min).

To ensure that the original RNA template was free of any residual genomic DNA, a "No RT" control (without reverse transcriptase) for each RNA sample was included in the experiment. For this purpose, the number of amplification cycles used should yield a specific product in the positive control (chromosomal DNA as a template), but no product in the "No template" control. If any amplicon was found, these RNA samples were re-treated with DNase I (RNase free) and subsequently checked again using PCR. The PCR products were run on 2% high resolution-agarose gels for transcriptional profile analysis. Quantification of bands was performed using the "Analyze-Measure" menu in the ImageJ 1.42q software package (Abramoff *et al*., 2004).

2.12. Real time quantitative polymerase chain reaction (RT-qPCR)

Reverse transcription was conducted as described in section 2.11. A pool of cDNA samples (*e.g.* cDNA samples of *scr4677* overexpression and control strains) was used to identify an optimal annealing temperature for each pair of

primers used in qPCR. Amplification specificity was determined by separating the amplicons on agarose gels. For each pair of primers, seven dilutions of the pooled cDNA sample (triplicate for each dilution) were used to generate a standard curve for determining amplification efficiency and linearity. Established criteria for "acceptable" standard curves, require an amplification efficiency of 90-110%, and linearity (coefficient of determination) higher than 0.980 (Taylor *et al*., 2010). Using the established standard curves, a dilution factor lying in the middle of the regression line was chosen for use in diluting each cDNA sample.

In the subsequent qPCR, 2 µL of diluted cDNA was added to 8 µL of qPCR cocktail (containing 5 µL of SsoFast™ Evagreen® supermix, 2.8 µL of nuclease-free water, and 5 pmol of each primer) in triplicate. A single replicate of "standard" cDNA (dilution of cDNA samples used previously for generating standard curves), duplicate of "No template" control and a positive control (plasmid/chromosomal DNA template) were included in each qPCR. The qPCR program was assigned based on optimized cycling conditions for the SsoFast™ Evagreen® supermix (Bio-Rad) as per the manufacturer's instructions, which comprised an enzyme activation step (95°C, 30 sec); amplification steps (40 cycles of: 95°C for 5 sec, 60-62.8°C for 5 sec); and melting step (60 to 95°C, 5 sec/step, 0.5°C/step). Fluorescence signals were recorded after each amplification and melting step. Cq or Ct values (PCR cycle at which fluorescence signal crosses a threshold set within the exponential phase of amplification) were determined using Bio-Rad CFX Manager 2.1.

A mathematical model (comparative Ct method) described in Pfaffl (2001) was chosen for relative quantification. Briefly, for each gene, Cq values (average of triplicates) for the test samples were subtracted from the values of appropriate control samples, giving a ΔCq value (control-test). These ΔCq values were then normalized using ΔCq values of *rpoB* (reference gene), yielding ΔΔCq. The expression ratio of a gene in the overexpression strain relative to its plasmid-alone control strain was $2^{\Delta\Delta Cq}$. Standard errors were computed using the expression ratios obtained from two biological replicates, with two RT replicates for each biological replicate.

2.13. Overexpression of His_6-AbeR in *E. coli* and its purification

To overexpress AbeR in *E. coli* and purify the resulting protein, the *abeR* coding sequence was first amplified by PCR using oligonucleotides OE3291-3 and OE3291-4 (Appendix, Table 2.3), and the amplicon was inserted into pIJ2925 to generate pMC121. From this plasmid, the coding sequence was excised using *Nde*I and *Bgl*II, and was inserted into the same sites of pET15b (Novagen). After the integrity of the *abeR* coding sequence and T7 promoter in pET15b were verified by sequencing, the plasmid clone (pMC119) was then introduced into *E. coli* strain Rosetta 2 (Novagen), which is a strain that produces additional rare tRNAs – associated with uncommon codons in *E. coli*.

The *E. coli* strain carrying pMC119 was grown overnight at room temperature in 100 mL of LB with ampicillin and chloramphenicol (without IPTG induction). After the cells were collected by centrifugation, the resulting pellet

was re-suspended in lysis buffer (20 mM Tris-HCl, pH 7.5, 250 mM NaCl, 10 mM imidazole). Lysozyme (750 µg/mL) and protease inhibitor (Roche), as per the manufacturer's instructions, were added into the cell suspension and this was incubated on ice for 30 min. To help in lysing the cells, the suspension was subjected to sonication (Branson Sonifier® 350). RNase A (80 µg/mL) and DNase I (40 U/mL) were added into the cell lysate, followed by 15 min incubation on ice and a subsequent centrifugation at 16,000 ×g (4°C) for 30 min. The soluble fraction was mixed with Ni-nitrilotriacetic acid agarose resin (Qiagen) as per the manufacturer's instructions, after which the resin was washed three times using wash buffers containing 20 mM Tris-HCl pH 7.5, and 250 mM NaCl, with increasing concentrations of imidazole (50-125 mM). His_6-AbeR was eluted in the wash buffer containing 250 mM imidazole, and the eluate was concentrated and de-salted using Nanosep centrifugal devices (Pall) with a 30 kDa cutoff. The concentration of purified His_6-AbeR was determined using a Bradford assay (Bradford, 1976), with pre-made Bradford reagent from Bio-Rad.

2.14. Crude *S. coelicolor* lysate preparation and protein purification

The *S. coelicolor* strains were grown on R2YE agar plates overlaid with cellophane discs. Confluent growth was obtained by spreading 5 µL of spores (approximately 2 x 10^7 spores) on each plate, followed by a 48 h-incubation at 30°C. Biomass was scraped from plates and weighed to obtain ~ 100 mg per sample (a small-scale crude cell lysate preparation). The cells were then re-suspended in lysis buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA, 1 mM DTT,

10% [v/v] glycerol, and 150 mM NaCl) with a ratio of 0.9 mL lysis buffer per 100 mg biomass. Into this suspension, protease inhibitor (Roche) was added, as per the manufacturer's instructions, and lysozyme (2.5 mg/mL lysis buffer) was added. The resulting mixture was placed on ice for 45 min, followed by mechanical cell disruption (Branson Sonifier® 350 set at "microtip limit" with three cycles of 10 sec with 10 sec rest). The soluble fraction was separated by centrifugation at 4°C (16,000 ×g for 20 min). The AbeR binding activity in this fraction was retained after being stored at -20°C for at least two weeks. Protein concentrations were determined using Bradford assays.

In the case of larger scale crude lysate preparation, biomass was obtained by growing *S. coelicolor* on cellophane discs overlaid on 250 or 350 mL R2YE agar medium in 1.9 or 3 L-casserole dishes (Libbey), respectively. Between 12-20 g biomass (per harvest) was mixed in the lysis buffer supplemented with protease inhibitor and lysozyme using the ratio described in the small scale-crude lysate preparation, after which this mixture was placed on ice for 45 min. The mixture was homogenized using a glass tissue grinder, followed by passaging through a French® press (Thermo Scientific) three times. A soluble fraction was separated by centrifugation at 16,000 ×g (4°C) for 20 min. The resulting supernatant was then subjected to ammonium sulphate (AS) precipitation.

The AS saturation was initially brought to 50%, *e.g.* by adding 24 g ammonium sulphate to ~80 mL lysate. Those proteins precipitated by centrifugation at 16,000 ×g for 20 min were designated the 0-50% AS saturation-

protein fraction, and the resulting pellet was re-suspended in TED buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA, and 1 mM DTT). The resulting supernatant was subjected to higher levels of AS saturation, until achieving close to 100% saturation (AS was unable to be further dissolved), followed by centrifugation to yield a 50-100% AS saturation protein fraction-pellet. This pellet was then also re-suspended in TED buffer. Both 0-50 and 50-100% AS saturation fraction were dialyzed twice against 1 L of dialysis buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA, 1 mM DTT, and 10 mM NaCl) for 16-17 h in total.

2.15. Electrophoresis mobility shift assays (EMSAs)

Probes used for EMSAs were prepared by labeling either PCR-amplified DNA or two annealed complementary oligonucleotides with [γ-^{32}P]ATP at their 5′ ends. Annealing of the complementary oligonucleotides involved incubation of the oligonucleotides at 95°C for 10 min in a buffer containing 10 mM Tris-HCl pH 7.5, 1 mM EDTA, and 60 mM NaCl, before slowly being cooled to room temperature. To radiolabel these DNA fragments, 30 pmol annealed DNA fragments or ~0.6 pmol of gel-purified PCR products were phosphorylated as described in section 2.9.

Each 20 µL-binding reaction consisted of 10 mM Tris-HCl (pH 7.5), 1 mM DTT, 1 µg poly dI/dC, 10% (v/v) glycerol, 0.5 to 25 nM [γ-^{32}P]-labelled DNA fragment, and varying amounts of purified AbeR, fresh crude lysate, or different AS protein fractions. After being incubated at 30°C (or room temperature in the case of AbeR) for 25-30 min, samples were separated in non-

denaturing 5-8% polyacrylamide gels using a Mini-PROTEAN® II electrophoresis cell (Bio-Rad). Electrophoresis was performed in 1 × TBE buffer at 100 V for 50-60 min. Gels were placed on layers of filter paper and wrapped using plastic wrap prior to exposure to phosphorscreens (Amersham Biosciences) or to Kodak Biomax XAR films at room temperature for 1 h.

2.16. DNase I footprinting assays

To generate probes in which template or nontemplate (coding) strands were specifically end-labelled, PCR amplification was conducted using primers 3287-2 and 3287MSAF (Appendix, Table 2.3), with one or the other being first labeled using [γ-^{32}P]ATP (as described in section 2.9). After the resulting PCR products were purified using a MinElute® Reaction Cleanup kit (Qiagen), approximately 6 nM probe was mixed with purified AbeR (0 to 1.6 μM) using the same AbeR-binding condition described in the EMSA section. DNase I (2.5 U, Roche) was then added into the mixture and incubated for 1 min. The reaction was stopped by adding PN buffer (Qiagen), and the mixture was applied to a MinElute® column (Qiagen), followed by elution in 30 μL distilled water. After the eluates were vacuum-dried for 30 to 45 min, the DNase I-cleaved products were re-suspended in 13 μL loading dye (80% [v/v] formamide; 1 mM EDTA, pH 8.0; 10 mM NaOH; 0.1% [w/v] bromophenol blue; 0.1% [w/v] xylene cyanol FF).

To prepare sequencing ladders, the Sequenase v2.0 DNA sequencing kit (USB Corporation) was used as per the manufacturer's instructions, with slight modification. Briefly, 10 μL-reactions comprising an alkaline-denatured template

(pMC123 in Table 2.2) and 5 μM appropriate primers (3287-2 or 3287MSAF) were boiled for 3 min and immediately placed on ice for 5 min. Into this template-primer mixture, Sequenase buffer, DTT, [α-^{33}P]dATP (Perkin Elmer), and Sequenase were added, as per the Sequenase manufacturer's instructions. The resulting mixture was then divided into four tubes (pre-warmed at 37°C), one of which contained the dGTP labeling mix and the ddG/A/T/C termination reagent. These tubes were then incubated at 37°C for 2 min, followed by the addition of a stopping reagent (provided in the kit).

Prior to loading the DNase I-cleaved products and associated sequencing ladders onto a 6% denaturing polyacrylamide gel (Gene Gel®, Bioshop), all samples were heated at 90°C for 3 to 5 min. Electrophoresis was run using constant power at 45 W for 45 or 75 min (for template or nontemplate/coding strand, respectively).

2.17. Tables

Table 2.1. Bacterial strains and plasmids used in this study

Strains	Genotype, description or use	Reference
Streptomyces coelicolor		
M145	SCP1$^-$ SCP2$^-$	(Kieser *et al.,* 2000)
M600	SCP1$^-$ SCP2$^-$	(Chakraburtty & Bibb, 1997)
J3410	M600 *rnc*::*apr*	(Sello & Buttner, 2008)
Δ*rnc*	M145 *rnc*::*apr*	(Sello & Buttner, 2008)

Table 2.1. Bacterial strains and plasmids used in this study (continued)

Strains	Genotype, description or use	Reference
Streptomyces coelicolor		
E301	M600 *SCO3287-3290::aac(3)IV (ΔabeABCD)*	(Hindra *et al*., 2010)
E304	M600 *SCO3291::aac(3)IV (ΔabeR)*	(Hindra *et al*., 2010)
E305	M600 *rnc*:: *aac(3)IV adpA*::*vph*	(Hindra *et al*., 2010)
E320	M145 *SCO4676::aac(3)IV*	This work
E321	M145 *SCO4676::scar* (Δ*SCO4676* unmarked)	This work
E322	M145 *SCO4677::aac(3)IV*	This work
E323	M145 pIJ82	This work
E324	M145 pWHM3	This work
E325	M600 pMS82	(Hindra *et al*., 2010)
E326	M600 pWHM3	(Hindra *et al*., 2010)
E327	M600 pMC116	(Hindra *et al*., 2010)
E328	M145 pMC149	This work
Escherichia coli		
DH5α	Plasmid construction and general subcloning	Invitrogen
ET12567/pUZ8002	Generation of methylation-free plasmid DNA	(MacNeil *et al*., 1992)
BW25113	Construction of cosmid-based knockouts	(Gust *et al*., 2003)
Rosetta 2(DE3)pLysS	Protein overexpression	Novagen

Table 2.2. Cosmids/plasmids used in this study

Plasmid/Cosmid	Description or use	Reference
pIJ790	Temperature sensitive plasmid carrying λ-RED genes	(Gust *et al*., 2003)
pIJ773	Plasmid carrying *aac(3)IV, oriT*	(Gust *et al*., 2003)
pIJ780	Plasmid carrying *vph, oriT*	(Gust *et al*., 2003)
pIJ2925	General cloning vector	(Janssen & Bibb, 1993)
pMC500	General cloning vector containing *ermE** promoter	the Elliot lab
pWHM3	High copy number *E. coli/ Streptomyces* plasmid	(Vara *et al*., 1989)
pMS82	Complementation of mutant strains	(Gregory *et al*., 2003)
pIJ82	pSET152 *aac(3)IV*::hyg^R	Gift from H. Kieser
pMU1	Transcriptional reporter (luminescence)	(Craney *et al*., 2007)
pUWLFLP	Plasmid carrying a Flp recombinase gene	(Fedoryshyn *et al*., 2008)
pMC116	pWHM3 + *abeA-D* (*SCO3297-3290*)	(Hindra *et al*., 2010)
pMC117	pWHM3 + α−*abeA (α-3287)*	(Hindra *et al*., 2010)
pMC118	pWHM3 + *abeR* (*SCO3291*)	(Hindra *et al*., 2010)
pMC119	pET15b + *abeR (SCO3291*)	(Hindra *et al*., 2010)
pMC120	pWHM3 + *abeA-D, R* (*SCO3287-3291*)	(Hindra *et al*., 2010)
pMC121	pIJ2925 + *abeR* for protein purification	(Hindra *et al*., 2010)
pMC122	pMS82 + *abeR* (*SCO3291*)	(Hindra *et al*., 2010)
pMC123	pIJ2925 + *abeA-D, R* (*SCO3287-3291*)	(Hindra *et al*., 2010)
pMC124	pIJ2925 + *abeR* (*SCO3291*)	(Hindra *et al*., 2010)
pMC125	pIJ2925 + *abeB* (*SCO3288*)	(Hindra *et al*., 2010)
pET15b	Overexpression of 6×His tagged proteins	Novagen

Table 2.2. Cosmids/plasmids used in this study (continued)

Plasmid/Cosmid	Description or use	Reference
pMC145	pIJ2925 + *SCO4676*	This work
pMC146	pIJ82 + *SCO4676*	This work
pMC147	pIJ2925 + *α-3287*	This work
pMC148	pIJ2925 + *scr4677*	This work
pMC149	pWHM3 + *scr4677*	This work
pMC150	pIJ82 + FLAG-*SCO4676*	This work
pMC151	pIJ82 + *ermE*-SCO4676*-FLAG	This work
pMC152	pRT801 + *SCO4676*-FLAG	This work
pMU1 + p*actIII*	Reporter plasmid for *actIII* expression	Xu & Nodwell
pMU1 + p*actII-4*	Reporter plasmid for *actII-orf4* expression	Xu & Nodwell
StE15	Cosmid for *abe* gene knockouts and PCR amplification	(Redenbach *et al*., 1996)
2StC13	Cosmid for *adpA* knockout	(Redenbach *et al*., 1996)
StD31	Cosmid for *SCO4676* and *SCO4677* knockouts and PCR amplification	(Redenbach *et al*., 1996)

CHAPTER 3

ANALYSIS OF THE *SCO4677-4675* LOCUS HARBOURING THE SCR4677 ANTISENSE RNA-ENCODING GENE

Preface

Our initial work on identifying *S. coelicolor* small RNAs (sRNAs) and profiling their transcription has been published in Swiercz *et al*. (2008). A small-scale cDNA cloning experiment was performed by Henry Haiser, Ph.D, a former graduate student in the Elliot laboratory. Two of the identified sRNAs, scr4677 and α-abeA, were selected for further investigation. This chapter outlines our investigation of scr4677. Mapping of scr4677 transcript ends and initial transcription profiles of scr4677 were performed by Jan Bobek, Ph.D, a former postdoctoral fellow in the laboratory.

3.1. Introduction

The *Streptomyces* genus is a group of filamentous bacteria that can be isolated from variable environments as free-living bacteria or as symbionts with fungi, plants, and animals (Seipke *et al*., 2011). These bacteria exhibit an intricate life cycle that starts with spore germination, following which they develop different morphological structures, starting with the vegetative/substrate hyphae, then raising aerial hyphae, and lastly converting the aerial hyphae into chains of exospores (reviewed in Flärdh and Buttner, 2009). In addition, the streptomycetes

also possess higher numbers of secondary metabolic gene clusters (ranging from 23 to 37 clusters in the currently sequenced genomes) than most other bacteria (Nett *et al*., 2009), which enable them to produce a myriad of secondary metabolites. Collectively, these numeruous genes contribute to a large genome size (~7-12 Mbp), which also enables the cells to adapt to the environmental stresses imposed by their "free-living" lifestyle (Bentley and Parkhill, 2004).

Those bacterial species having genomes larger than 6 Mbp are typically found to harbour a higher proportion of transcription regulators than those having smaller genomes (Cases *et al*., 2003). The model organism for studying the genetics of the filamentous *Streptomyces* bacteria is *Streptomyces coelicolor,* which is predicted to devote more than 12% of its proteins to regulation (Bentley *et al*., 2002). The regulatory elements encoded within a bacterial genome comprise not only protein-coding genes but also important non-coding elements (Christen *et al*., 2011), and as such, the *S. coelicolor* genome is expected to contain additional regulatory elements in the form of non-coding regulatory RNAs.

Bacterial non-coding RNAs, or small RNAs (sRNAs) as they are often called, exert their regulatory effects either through base-pairing with target mRNAs, affecting their stability or translatability, or through binding target proteins, modulating protein activities (reviewed in Storz, 2011). The base pairing of antisense RNAs (*cis*-encoded sRNAs) with their cognate sense mRNAs leads to the formation of structures that could be recognized by the double strand-

specific ribonuclease RNase III (Resch *et al*., 2008; Opdyke *et al*., 2011). In *S. coelicolor*, RNase III is a global regulator of antibiotic biosynthesis (Adamidis and Champness, 1992; Aceti and Champness, 1998), and it affects the abundance of anywhere between 79 and 200 transcripts, including that of two sRNAs (Huang *et al*., 2005; Gatewood, *et al*., 2012). Although it is well-established that RNase III acts as a pleiotropic regulator in *S. coelicolor*, the regulatory connection between RNase III and antisense RNAs, as well as RNase III and antibiotic production, has not been explored experimentally.

Given the advanced technologies that can now be used in discovering sRNAs, such as high-resolution microarrays and RNA-Seq, an increasing number of chromosomally-encoded antisense RNAs has been detected directly from genome-wide studies in many bacteria (Selinger *et al*., 2000; Sharma *et al*., 2010; Vockenhuber *et al*., 2011). In the streptomycetes, the identification of antisense RNAs has been reported in several studies (D'Alia *et al*., 2010; Vockenhuber *et al*., 2011), including several from the Elliot laboratory (Swiercz *et al*., 2008). With the representation of antisense RNAs in the bacterial transcriptome continuing to expand, these newly discovered antisense RNAs raise obvious questions regarding their physiological function in bacterial cells.

In our study, we identified several antisense RNA candidates using a cDNA cloning approach (Swiercz *et al.,* 2008). These included scr4677, an antisense RNA transcribed in the intergenic region between *SCO4676* and *SCO4677*. In this chapter, I focus on exploring the potential for antisense RNA-

mediated genetic regulation of uncharacterized target genes, and probing the mechanism of antisense RNA activity and the contributions made by these molecules to *Streptomyces* biology. Through transcriptional profiling, I show that *SCO4677-4675* comprises a three gene-operon, and reveal that the expression of *SCO4676* and *SCO4675* is differentially affected by nutritional status. I propose that the scr4677 antisense RNA may contribute to stabilizing the *SCO4677-4676* read-through transcript, and at the same time, destabilizing the *SCO4676* transcript. Furthermore, I show that RNase III affects the expression of *SCO4677-4675,* but not *scr4677*. Finally, phenotypic analyses reveal a role for *SCO4676* and *SCO4677* in enhancing actinorhodin production during growth on soy flour agar with glucose.

3.2. Results

3.2.1. Identification and transcription profiling of scr4677

The scr4677 antisense RNA was detected using northern blotting, in RNA samples isolated during *S. coelicolor* growth on either R2YE or minimal media (MM) (containing glucose and mannitol, respectively) agar plates (Fig. 3.1a). The expression of *scr4677* was readily detectable at all stages of growth (vegetative, aerial hyphae, and sporulation) on both media types. The *scr4677* transcript was more abundant during growth on R2YE than on MM. RACE experiments led to the successful mapping of two 5′ ends and one 3′ end for the scr4677 transcripts, which corresponded to transcripts of either 34 nt or ~60 nt in length. The smaller transcript likely represented the original cloned cDNA

fragment (Swiercz *et al.,* 2008), and may represent a processed version of the larger transcript. Notably, the 34 nt scr4677 transcript was detected in the R2YE RNA samples, but was not observed in the MM samples. This could be due to: (1) overall differences in the nutritional composition of R2YE and MM (rich versus minimal medium), or (2) specific effects stemming from carbon source usage (glucose versus mannitol). To differentiate between these possibilities, scr4677 transcript levels were examined during growth on soy flour agar media containing mannitol or glucose as a carbon source (SFM or SFG, respectively). The smaller scr4677 transcripts were observed exclusively in RNA samples isolated during growth on SFG (Fig. 3.1b), suggesting the appearance of the 34 nt-scr4677 transcripts was glucose-dependent.

3.2.2. Bioinformatic analysis of the scr4677*-containing genetic region*

The scr4677 antisense RNA was transcribed from the intergenic region between *SCO4676* and *SCO4677* (on the opposite strand; Fig. 3.2a). The RACE experiments described above revealed *scr4677* transcription to begin 14 nt upstream of the *SCO4676* translation start site and its transcription to terminate after ~60 nt. *SCO4676* encodes a protein of unknown function, possessing a helix-turn-helix motif that was predicted through protein sequence and structural comparison using CDD, HHpred and Phyre databases (Marchler-Bauer *et al*., 2011; Söding *et al*., 2005; and Kelley and Sternberg, 2009; respectively), while *SCO4677* encodes a putative regulator containing a histidine kinase-like ATPase domain (Marchler-Bauer *et al*., 2011). SCO4677 has been shown to interact with

SigF, a sigma factor required during the later stages of sporulation (Kelemen *et al*., 1996), and with two other *S. coelicolor* proteins that resemble anti-sigma factor antagonists, suggesting that SCO4677 may function as an anti-sigma factor (Kim *et al*., 2008). Its interaction with SigF suggests that SCO4677 may function in modulating *Streptomyces* development. This proposal is further supported by the fact that *SCO4677* is a regulatory target of BldD (den Hengst *et al*, 2010), a protein with an essential role in both morphological differentiation and antibiotic production (Elliot *et al*., 1998 and 2001).

Bioinformatic analyses using BLAST programs (Altschul *et al*., 1997) exploring the National Center for Biotechnology Information (NCBI) and Broad Institute (Harvard and MIT) databases (up until May 2012), revealed SCO4676 homologues (having more than 50% identity and 59% similarity) in 14 out of 43 *Streptomyces* species. These 14 species (listed in Table 3.1) were then interrogated with the *scr4677* sequence, along with its putative promoter region, and it was discovered that 12 of the 14 species appeared to have a genetic locus similar to that of *scr4677* and *SCO4676* (Fig. 2b). Similar analyses also suggested that SCO4677 was widely conserved in the streptomycetes, with SCO4677 and SCO4676 homologue-encoding genes positioned next each other in eight *Streptomyces* species. Of these eight species, seven appeared to possess the *scr4677* sequence in the intergenic region separating *SCO4677* and *SCO4676* (Table 3.1).

Although it is not known whether SCO4677 and SCO4676 have a physiological connection, *in silico* analysis of available *S. coelicolor* transcriptome data has suggested that *SCO4677* is co-transcribed with *SCO4676,* along with the immediate downstream gene *SCO4675* (Castro-Melchor *et al*., 2010). This computer prediction used the following criteria: (1) functional similarity of proteins encoded by a gene and a pair of its adjacent genes, based on protein classification schemes at the Welcome Trust Sanger Institute and Gene Ontology; (2) gene order conservation across a number of bacterial genomes; and (3) gene expression similarity, based on correlation between pairs of adjacent genes. Like *SCO4676*, *SCO4675* encodes a conserved hypothetical protein; six other *Streptomyces* species encode SCO4675 homologues (proteins sharing more than 50% identity and 67% similarity). Within these species, only three have *SCO4675* positioned immediately downstream of *SCO4676*-like genes. Based on the bioinformatic data described above, scr4677*,* SCO4676 and/or SCO4677, but not always SCO4675, appear to be genetically linked, and may contribute to a shared physiological role in *S. coelicolor*.

3.2.3. SCO4677-4675 *are co-transcribed, but exhibit intra-operonic differences in expression in response to nutrient conditions*

Given the conservation of the *scr4677* sequence within the intergenic region of *SCO4676* and *SCO4677*, we postulated that scr4677 may function in regulating the expression of either *SCO4676* or *SCO4677.* We also considered the possibility that the scr4677 antisense RNA may influence the stability of the

SCO4677-4675 polycistronic mRNA, given that these genes are predicted to be transcribed as an operon. Therefore, as an initial step towards testing these possibilities, we examined the transcription profiles of *SCO4677, SCO4676* and *SCO4675* using semi-quantitiative RT-PCR.

Under nutrient-limited growth conditions (MM agar+mannitol), we found all genes (*SCO4675, SCO4676* and *SCO4677*) were constitutively expressed (Fig. 3.3). In contrast, there were differences in the expression of each gene during growth on R2YE (a rich medium containing glucose) agar: *SCO4677* and *SCO4675* were constitutively expressed, but the expression of *SCO4676* was reduced at the final time point examined (48 h). Moreover, through comparison of fluorescence intensities of PCR products separated on the same agarose gel, it appears that *SCO4676* and *SCO4675* transcripts were more abundant during growth on MM+mannitol, relative to R2YE (Fig 3.3). We were also able to detect read-through transcription of *SCO4676* into *SCO4675*, suggesting that they are co-transcribed (Fig. 3.3). Like *SCO4676* and *SCO4675,* the *SCO4676-SCO4675* read-through transcript was also apparently more abundant during growth on MM+mannitol, relative to R2YE. Moreover, the differences in the expression of *SCO4676* and *SCO4675* during growth on R2YE agar suggested that *SCO4675* may also have its own promoter, despite being separated from *SCO4676* by a mere 17 bp.

Unlike *SCO4676* and *SCO4675,* growth conditions did not appear to affect *SCO4677* transcription, except at 40 h during growth on R2YE, where we also

observed reduced levels of 16S rRNA (Fig. 3.3). Given the *in silico* prediction by Castro-Melchor *et al*. (2010), we tested the possibility that *SCO4677-4676* were also co-transcribed (despite being separated by ~250 bp) using RT-PCR. We found there to be read-through transcription of *SCO4677* into *SCO4676* during growth on minimal medium, but interestingly, read-through levels were reduced during growth on R2YE (Figure 3.3). We wondered whether this difference in co-transcription on the different media types was in any way connected to *scr4677*, and/or whether the difference was resulted from interaction between scr4677 and the *SCO4677-4675* polycistronic mRNA.

3.2.4. scr4677 *overexpression affects the levels of* SCO4676 *and* SCO4677-4676 *read-through mRNA transcripts, but not* SCO4677 *transcript levels*

In addition to potential interactions between scr4677 and the *SCO4677-4675* polycistronic mRNA, scr4677 could also base-pair with the untranslated region of *SCO4676* mRNA, which we determined to be approximately 50 nt-long using S1 nuclease mapping experiments (Fig. 3.4a). This suggested that *SCO4676* transcription was driven by at least two promoters (illustrated in Fig. 3.4b): one upstream of *SCO4677*, and one upstream of *SCO4676*. To investigate scr4677 function in modulating *SCO4676/4677* expression, scr4677 was overexpressed by cloning its gene (together with all necessary upstream regulatory elements) into a high-copy number plasmid, and the effect of its overexpression on the transcript levels of *SCO4676* and the *SCO4677-4676* read-through were assessed during growth on R2YE.

We first confirmed the overexpression of *scr4677* using northern blotting, and showed that scr4677 transcript levels were indeed increased in the *scr4677* overexpression strain (relative to a control strain) over a 72 h-time course (Fig. 3.4c). We found the highest expression ratio to be at 42 h, where scr4677 levels were ~3 fold higher than in the control strain (Fig. 3.4c). Similar levels of overexpression were reported in Vockenhuber and Suess (2011), where they were examining another sRNA (scr5239) whose gene they had cloned at high copy under control of the strong SF14 promoter (Labes *et al*., 1997).

Transcript levels of *SCO4676* and *SCO4677-4676* read-through were then assessed in the *scr4677* overexpression and control strains using real time-quantitative PCR (RT-qPCR) and semi-quantitative RT-PCR, respectively. The levels of the *SCO4677-4676* read-through transcript were examined using semi-quantitative RT-PCR because the amplicon length (284 bp) exceeded the upper limit of optimal amplicon size for qPCR (200 bp, Wang *et al*., 2006). We found that *SCO4677-4676* read-through transcript levels in the overexpression strain were higher than that in the control plasmid-carrying strain over the entire 72 h-time course (Figure 3.4d). This suggested that scr4677 might act to stabilize the *SCO4677-4676* read-through transcript. We asked whether this stabilizing effect also impacted the *SCO4677* transcript, but found that this did not appear to be the case, as transcript levels of *SCO4677* were similar in the *scr4677* overexpression and its plasmid control-carrying strains (Fig. 3.4e).

Overexpressing *scr4677* did, however, lead to differential effects on the levels of *SCO4676* mRNA (Fig. 3.4d). At both 42 and 56 h, *SCO4676* transcript levels were decreased by almost 2 fold in the overexpression strain, suggesting that *scr4677* might negatively affect stability of *SCO4676* transcripts. However, at 72 h, increased levels of *SCO4676* transcripts were observed in the overexpression strain. This increased expression may have been resulted from increased levels of *SCO4677-4676* read-through transcript (Fig. 3.4d), which was examined using the same product of reverse-transcription product but separate PCR amplification (the relative location of primers is depicted in Fig. 3.4b).

3.2.5. Probing the biological role(s) of scr4677 and SCO4676

Given that *SCO4676* and *scr4677* are conserved amongst a number of *Streptomyces* species, we wondered about the functional role of these gene products in the cell. We generated Δ*SCO4676* strains, including an unmarked Δ*SCO4676* mutant to avoid any polar effects on the downstream gene *SCO4675*. Deleting *scr4677* alone would not be straightforward, as such a disruption would also likely disrupt *SCO4676* expression (given that its promoter lay within the *SCO4676* coding sequence), and therefore, only the *scr4677* overexpression strain (strain E328, Table 2.1) was used for phenotypic analyses.

We compared the phenotype of Δ*SCO4676* and wild type strains on a number of different solid media types [rich medium (R2YE+glucose/mannitol), minimal medium+glucose/mannitol, soy flour medium (SF+glucose/mannitol), maltose-yeast extract-malt extract (MYM), supplemented minimal medium

(SMM with casaminoacids and mannitol) and LB], and found the mutant strain resembled its wild-type parent during growth on all types of solid media tested, except for soy flour glucose (SFG). On this particular solid medium, both marked and unmarked Δ*SCO4676* strains precociously produced the blue pigmented antibiotic actinorhodin (Fig. 3.5a), and exhibited enhanced actinorhodin production over the course of 6 days growth (Fig. 3.5b). To ensure that this phenotypic change was due to the deletion of *SCO4676*, a single copy of wild-type *SCO4676* (including *scr4677*) was re-introduced into Δ*SCO4676* on an integrating plasmid vector. The resulting strain produced actinorhodin at wild-type levels (Fig. 3.5c), when compared with a plasmid alone-containing wild type strain. We then tested whether this *SCO4676*-specific effect on actinorhodin production could be observed during growth in liquid media, focusing specifically on R5 (rich) and NMMP (minimal) supplemented with either glucose or mannitol as a carbon source. The actinorhodin production of the *SCO4676* mutant resembled that of wild-type strain in all liquid media. We also examined the phenotype of the *scr4677* overexpression strain during growth on SFG. This strain did not exhibit any phenotypic change relative to its control plasmid-carrying strain during growth on SFG, or under any other growth condition tested.

Having established that *SCO4676* was co-transcribed with *SCO4677,* a deletion strain of *SCO4677* was also generated, and its phenotype was examined. This mutant strain exhibited a similar phenotype to that of Δ*SCO4676* strains during growth on SFG agar plates (Fig. 3.5b), in that it overproduced

actinorhodin. This suggested that the gene products of both genes may contribute to the same pathways in the cell. However, Δ*SCO4677* showed additional phenotypic changes on SFM (soy flour with mannitol) agar (Fig. 3.5d), namely: (1) enhanced actinorhodin biosynthesis, and (2) enhanced grey pigmentation during sporulation. Similar phenotypic effects have been previously reported for a *SCO4677* deletion mutant (Kim *et al*., 2008).

Enhanced grey pigmentation observed in *ΔSCO4677* may suggest altered spore wall integrity, as the grey pigment is thought to be deposited in the coat of mature spores (Chater, 1998). This rationale and the fact that SCO4677 has been shown to interact with the sporulation sigma factor SigF (which controls the production of the grey spore pigment; Kelemen *et al*., 1998) suggested that SCO4677 may function in modulating spore development. Therefore, we wondered whether genetic perturbation of *SCO4676* and *scr4677* would impact spore integrity. We could not detect any morphological defects or sporulation abnormalities for *ΔSCO4676* or *scr4677* overexpression strains, as examined using coverslip impressions (a technique used to attach spore chains to microscope coverslips), coupled with light microscopy. To test whether *SCO4676* deletion or *scr4677* overexpression affected spore sensitivity to heat, spores of these strains, along with appropriate control strains, were exposed to heat (60°C) for a period of time before being inoculated onto solid agar plates; however, these spores were neither more sensitive nor resistant to heat than their control spores.

We also examined whether deleting *SCO4676* or *SCO4677* and overexpressing *scr4677* had any influence on *S. coelicolor* growth following exposure to a panel of environmental stresses. Spores of the deletion and overexpression strains were inoculated onto growth medium, after which it was overlaid with filter discs saturated with different chemical stressors, to induce cell envelope stress (SDS, EDTA, lysozyme, and vancomycin), oxidative and thiol stresses (hydrogen peroxide and diamide, respectively), and iron starvation (dipyridyl). The sensitivity of the mutant and overexpression strains to these stressors did not differ when compared to control strains (Appendix, Table 6.1 and 6.2). Therefore, based on our phenotypic analyses, *SCO4677* disruption appears to affect both antibiotic production and sporulation on several media types, *SCO4676* disruption impacts actinorhodin production, although only on SFG agar medium, and modulating *scr4677* levels does not appear to have any striking phenotypic consequences.

3.2.6. Identification of a transcription activator of scr4677 *expression*

In addition to investigating the function of scr4677*,* we also asked whether *scr4677* transcription was controlled by either of its flanking protein-encoding genes (*SCO4676* and *SCO4677)*. Given that SCO4676 is predicted to have a helix-turn-helix motif, suggesting that it may function as a DNA-binding protein, scr4677 expression was examined in a Δ*SCO4676* mutant strain. The majority of the *SCO4676* coding sequence in this strain had been replaced with an apramycin resistance cassette (in-frame deletion), although the first 21 aa (63 bp) of the N-

terminus had been intentionally retained to ensure that the *scr4677* promoter was fully intact (Fig. 3.6a). Interestingly, we found the transcript levels of *scr4677* to be almost completely abolished in the *SCO4676* mutant (Fig. 3.6b). This could be due to either the disruption of a regulatory element required for *scr4677* transcription in the *SCO4676* mutant, or the requirement of a functional SCO4676 protein for *scr4677* expression.

To differentiate between these possibilities, we took advantage of our *scr4677* overexpression strain. We set about assessing whether a Δ*SCO4676* mutant strain carrying the *scr4677* overexpression plasmid (where *scr4677* had all the necessary elements for effective transcription) was impaired in its expression of *scr4677*. In this strain, we found that *scr4677* was effectively transcribed (Fig. 3.6c), indicating that its transcription did not depend on a functional SCO4676 protein. This instead implied that *scr4677* expression was dependent upon a regulatory element located 80-160 bp upstream of *scr4677* start site, reflecting the sequence that was present in the overexpression construct, but missing in the Δ*SCO4676* mutant.

We also found *scr4677* transcript levels to be less abundant in the Δ*SCO4677* mutant than its wild-type parent strain (Fig. 3.6d), suggesting that SCO4677 (a putative regulatory protein) might be required for *scr4677* transcription. To assess whether SCO4677 directly bound and activated *scr4677* transcription, the 80-160 bp upstream region was used as a probe for EMSA experiments (illustrated in Fig. 3.6a). Crude lysates from wild-type, Δ*SCO4676*,

and Δ*SCO4677* strains were tested for their ability to shift this probe using EMSAs. We found that neither SCO4676 nor SCO4677 were required for binding and shifting the probe, and instead showed that an unknown protein(s) specifically bound the probe (Fig. 3.6e). We have been working towards identifying this unknown protein.

As a first approach to identifying the transcription activator controlling *scr4677* expression, we created an *S. coelicolor* genomic "overexpression" library, together with a reporter system based on kanamycin resistance/sensitivity. This reporter construct contained an upstream region of *scr4677* fused to a promoterless kanamycin resistance gene, and this construct was integrated into the *S. coelicolor* genome. The overexpression library comprised chromosomal DNA fragments cloned downstream of a strong promoter (*ermE**, Bibb *et al*. [1994]) on another integrating plasmid, which was integrated into a different site on the *S. coelicolor* genome. This library was introduced into the reporter strain, with the aim being to screen clones that were more resistant to kanamycin than the reporter strain alone. However, identifying positive clones was hampered by a high rate of spontaneous kanamycin resistance observed for *S. coelicolor*.

A more direct biochemical approach was then undertaken using protein isolation and fractionation, via ammonium sulphate (AS) precipitation, fast protein liquid chromatography using a heparin column (which has a high binding capacity for DNA-binding proteins), and finally, biotinylated probe-streptavidin affinity chromatography. Protein fractions yielded from each step were examined

using EMSAs to identify those fractions containing specific binding activity for the 80-160 bp upstream region of *scr4677*. Several attempts have been undertaken using this strategy; however, specific binding activity has only ever been reproducibly observed for fractions following AS precipitation.

3.2.7. RNase III indirectly affects the expression of SCO4676 *and the levels of* SCO4677-4676 *read-through transcript*

We had determined that overexpressing *scr4677* appeared to negatively affect the stability of *SCO4676* transcripts, and therefore, we wondered whether RNase III may play a role in this system. In other organisms, RNase III has been reported to contribute to antisense RNA-mediated control of target gene expression (Resch *et al*., 2008; Opdyke *et al*., 2011). A common mechanism of antisense RNA activity involves base pairing with its complementary mRNA, forming an extended region of double stranded RNA, and such a structure is predicted to be readily targeted for degradation by RNase III.

To address whether RNase III was involved in processing a potential scr4677 – 5′ UTR *SCO4676* duplex, we examined the expression of these two genes in both a wild-type strain and its isogenic RNase III null mutant (Δ*rnc,* listed in Table 2.1). We found that scr4677 transcript levels were relatively unchanged in the *rnc* mutant relative to the wild-type parent strain, suggesting that the scr4677 antisense RNA is not a target for RNase III (Fig. 3.7a). The additional scr4677 transcripts (smaller transcripts) that were detected exclusively during growth on glucose-containing media (section 3.2.1) were also still present

in the *rnc* mutant, indicating that these transcripts were not the product of RNase III digestion.

Interestingly, *SCO4676* transcripts were less abundant in the *rnc* mutant than the wild-type strain (Fig. 3.7b), suggesting that RNase III indirectly affected *SCO4676* transcription and/or transcript stability. A similar effect was observed for the *SCO4677-4676* read-through transcript, where its levels were also reduced in an *rnc* mutant (Fig. 3.7c). Taken together, it appears that RNase III does not have a role in the *scr4677*-mediated effects on *SCO4676* and *SCO4677-4676* read-through transcripts: however, it does have an indirect, and apparently *scr4677*-independent, effect on *SCO4676* and *SCO4677-4676* readthrough transcript levels.

3.3. Discussion

In this study, we have determined that increasing transcript levels of scr4677 antisense RNA affected the levels of *SCO4676* and *SCO4677-4676* read-through transcripts. We also observed that *scr4677* expression differed, depending on nutritional status. scr4677 was more abundant when *S. coelicolor* was grown on R2YE (rich medium) than on minimal medium, and the processed (smaller) *scr4677* transcripts were observed only during growth on glucose-containing media like R2YE or soy flour agar supplemented with glucose (section 3.2.1). The presence of both full length scr4677 transcript and putative processed forms is reminiscent of another antisense RNA-intergenic sense transcript system: the GadY and *gadX-gadW* system in *E. coli* (Opdyke *et al.*, 2004). Here, full-

length GadY sRNA and two additional processed forms of the sRNA have been reported, where the two processed forms could act in place of the full-length transcript (Opdyke *et al.*, 2011). It is worth noting, however, that we also often observed additional (smaller) transcripts when probing for the 5S rRNA transcript via northern blotting, when examining R2YE-isolated RNA samples (and not MM RNA samples), irrespective of the strain from which these RNAs were isolated (Fig. 3.1a). Therefore, we speculate that the smaller scr4677 transcripts may be the result of generally increased nuclease activity during growth on glucose-containing media.

The role of antisense RNAs in modulating the stability of their sense counterparts has been studied previously in *E. coli*; however, there are only a few examples where antisense RNAs overlap within intergenic sequences of polycistronic mRNAs (Desnoyers *et al.*, 2009; Opdyke *et al.*, 2004). One of the first documented examples was RyhB, an antisense RNA encoded in the intergenic region between *iscR* and *iscS*. This antisense RNA stimulated the degradation of the *iscSUA* polycistronic mRNA, but at the same time, promoted the stabilization and accumulation of *iscR* transcripts (the first gene of the operon) during growth under iron limiting conditions (Desnoyers *et al.*, 2009). It has been suggested that a strong secondary structure forms within the *iscRS* intergenic transcript upon RyhB binding. Intriguingly, degradation of *iscSUA* transcript does not take place within the *iscRS* intergenic region, but rather it occurs

downstream of the RyhB interaction site, within the *iscS* cooding sequence (Prévost *et al.*, 2011).

A second example of antisense-mediated polycistronic RNA control has been seen with GadY, an antisense RNA transcribed within the intergenic region of *gadX-gadW* mRNA (Opdyke *et al.*, 2004), described briefly above. This sRNA guides the processing of the *gadX-gadW* mRNA, where both *gadX* and *gadW* encode transcription regulators that respond to acid stress. Intra-operonic processing occurs within the sense-antisense complementary region (Opdyke *et al.*, 2011), and results in accumulation of the processed *gadX* and *gadW* mRNA fragments. Therefore, our finding that *scr4677* acts on the *SCO4677-4676* polycistronic transcript is not unprecedented. It is worth noting, however, that *scr4677* appeared to have little effect on the level of *SCO4677* mRNA, unlike the situation with RyhB. It is possible that the 3′ untranslated region of the *SCO4677* mRNA may readily adopt a secondary structure that helps to stabilize the upstream mRNA, irrespective of whether scr4677 is bound within the intergenic region.

Here, we have shown that *SCO4676* transcription is driven by two promoters, one positioned immediately upstream of *SCO4676,* and the second, located upstream of *SCO4677*. Thus, overall levels of *SCO4676* transcripts are influenced by both *SCO4677-4676* read-through and *SCO4676*-specific transcription, as well as by the effect of scr4677 on both these transcripts. We determined that overexpressing *scr4677* increased the levels of *SCO4677-4676*

read-through transcript, and thus we expected that this would be accompanied by a subsequent increase in overall *SCO4676* levels. However, our data suggest the opposite, as we observed the overall levels of *SCO4676* transcripts were reduced upon *scr4677* overexpression, at least at 42 and 56 h. This suggested that scr4677 may promote the degradation of the *SCO4676*-specific transcript. The scr4677 antisense RNA may exert its function on the 5′ UTR *SCO4676* transcript and the *SCO4677-4676* read-through transcripts through different mechanisms, although how this works at a mechanistic level has not yet been fully elucidated. One possibility is that the *SCO4677-4676* read-through transcript may be prone to degradation during growth on R2YE, but upon scr4677 binding, the RNA within this intergenic region adopts an alternative conformation that is more resistant to nuclease activity.

In the case of the *SCO4676*-specific transcript, its stability may be negatively affected by scr4677; however, we cannot yet exclude the possibility that scr4677 exerts its function through a transcription interference mechanism. We hypothesized that this effect could be mediated by the catalytic action of RNase III, since the double stranded RNA formed as a result of base-pairing between scr4677 and the *SCO4676* transcript would seem to be an outstanding candidate for RNase III-mediated degradation (as reviewed in Nicholson, 2011). If this were the case, we would expect to observe more abundant scr4677 and *SCO4676* transcripts in an *rnc* mutant strain. However, our data did not support such a scenario. RNase III had a little effect on the transcript levels of scr4677, as

determined by comparing scr4677 levels in a wild type and *rnc* mutant strain. This would be consistent with a very recent report showing that RNase III has a little effect on antisense transcript levels in *B. subtilis* (Durand *et al.*, 2012). Furthermore, in the *rnc* mutant, transcript levels of *SCO4676* and *SCO4677-4676* read-through were both reduced relative to that observed in a wild type strain. Taken together, it appears that RNase III is not involved in regulating the expression of *SCO4676* through base-pairing with scr4677.

In considering other alternatives, we wondered whether scr4677 could be exerting its regulatory effects on *SCO4676* through a non-catalytic mechanism. The scr4677 binding to the untranslated region of the *SCO4676* transcript may promote conformational changes in the *SCO4676* transcript, perhaps altering the accessibility of the ribosome binding site, thereby impacting *SCO4676* translation. Preventing ribosome binding to the *SCO4676* mRNA could leave the mRNA transcript unprotected, leading to reduced transcript levels of *SCO4676.* This is consistent with our data. In such a scenario, the relative abundance of antisense RNA versus mRNA would be critical for affecting mRNA translatability (Waters and Storz, 2009). Such a mechanism is well documented for sRNAs in general, and has also been observed for a number of antisense RNAs. One such example, outlined in chapter 1, is the SymR antisense RNA, which overlaps with 5′ terminus of the *symE* mRNA, encompassing its Shine-Dalgarno sequence (Kawano *et al*., 2007). In wild-type *E. coli* strain, SymR transcript levels are approximately 10 fold higher than that of *symE.* When SymR

is no longer transcribed, SymE protein levels are elevated, along with increased levels of its mRNA transcript. When SymR is transcribed *in trans*, it affects the translation of *symE* mRNA and impacts the stability of *symE* mRNA. We actively pursued testing this possibility for scr4677 and *SCO4676*; however, we were unable to draw any conclusions, as we were never able to detect tagged variants of SCO4676 (either N-terminally or C-terminally 3×FLAG tagged) by immunoblotting, at any point during growth on minimal medium, soy flour (with mannitol), or R2YE media.

Both Δ*SCO4676* and Δ*SCO4677* strains precociously produced actinorhodin, and exhibited greater actinorhodin production relative to their wild-type parents. While these phenotypic consequences were observed for both mutants during growth on soy flour agar supplemented with 2% (w/v) glucose, only Δ*SCO4677* mutant strain showed enhanced actinorhodin production during growth on soy flour agar supplemented with mannitol (SFM). The *SCO4677* deletion has previously been shown to affect both morphological development and antibiotic production (Kim *et al*., 2008). In their work, SCO4677 was also shown to interact *in vitro* with the sporulation sigma factor SigF, as well as with the putative anti-sigma factor antagonist SCO0869, suggesting that SCO4677 may function as an anti-sigma factor. Interestingly, however, mutants lacking *sigF* or *SCO0869* do not exhibit any defects in antibiotic production (Kim *et al*., 2008), suggesting that SCO4677 has additional functions/potential partners in the cell. SCO4676 may be a potential interacting partner for SCO4677, although this is not

something that we have yet tested, in part due to the fact that there is nothing about the SCO4676 sequence to suggest that it has sigma factor or anti-sigma factor properties. Collectively, our phenotypic observations suggest that SCO4677 function may be connected with that of SCO4676, but that this may be restricted on the basis of nutritional status. SCO4677 may therefore act through additional regulatory pathways which are currently elusive, in affecting antibiotic production in *S. coelicolor*.

Given that previous work by Kim *et al*. (2008) showed that SCO4677 may function as an anti-sigma factor for SigF, we considered the possibility that SCO4677 regulated *scr4677* expression through its interaction with SigF. This sigma factor is one of the nine *S. coelicolor* sigma factors that are similar to the general stress response-sigma factor SigB in *B subtilis* (Dalton *et al*., 2007). As these sigma factors bind a relatively well-defined promoter sequence, we searched for such a sequence in and around the *scr4677* promoter region, and found hexameric motifs that were similar to the *B. subtilis* SigB binding sequences, upstream of *scr4677*. If *scr4677* was a member of the SigF regulon, and SCO4677 was an antagonist of SigF activity, then we would expect to see increased levels of scr4677 transcript in the Δ*SCO4677* mutant strain. Instead, we observed the reduced transcript levels for *scr4677*. Therefore, the connection between SCO4677 and *scr4677* is not yet resolved. We have, however, established that *scr4677* expression depends on an unknown transcription activator(s). Initial attempts have not been successful in detecting *scr4677*

specific-binding activity after protein fractionation using a heparin column in the FPLC system. As we reproducibly observed specific binding activity following ammonium sulphate (AS) precipitation, we are considering the use of alternative columns for FPLC fractionation. We are also working towards modifying the biotinylated probe-streptavidin affinity chromatography step by, for instance, introducing protein fractions firstly to the biotinylated probe, then to the streptavidin beads (to avoid any steric hindrance that might be introduced when the biotinylated probe is already bound to the streptavidin beads).

Despite the fact that transcription analyses have suggested that scr4677 negatively affected *SCO4676* transcript levels, and Δ*SCO4676* mutants precociously produced actinorhodin during growth on SFG agar plates, we did not observe any phenotypic effects resulting from overexpression of scr4677. For one, overexpressing scr4677 did not completely eliminate *SCO4676* expression. This may support a more subtle role for scr4677 in modulating the expression of its adjacent genes, which is consistent with the emerging view that in bacteria, non-coding RNAs often act to 'fine-tune' gene expression (reviewed in Gottesman and Storz, 2011). A classic example of this can be seen in *E. coli*, where the stationary phase sigma factor SigS-encoding gene *rpoS* is differentially affected by multiple sRNAs (DsrA, RprA, ArcZ, and OxyS), and the expression of each sRNA is triggered by a different environmental cue (reviewed in Gottesman and Storz, 2011). While the oxidative stress inducible-OxyS sRNA represses *rpoS* translation (Zhang et al., 1998), DsrA and RprA, expression of

which increased by low temperature and cell surface stress, respectively, positively regulate *rpoS* translation (Sledjeski *et al*., 1996; Majdalani *et al*., 2002). Moreover, the ArcZ sRNA, which can be repressed by anaerobic growth conditions, also upregulates *rpoS* translation (Mandin and Gottesman, 2010). Deletion of all three sRNA (DsrA, RprA, and ArcZ)-encoding genes failed to completely abolish translation of *rpoS* (Mandin and Gottesman, 2010).

In our phenotypic analyses, we also subjected *scr4677* overexpression and *SCO4676* deletion strains to various chemical stressors; however, we could not detect any phenotypic changes (relative to their respective control strains) under any conditions tested. Notably, in similar studies undertaken in *S. griseus,* where 12 different sRNAs were deleted individually, no significant phenotypic changes were observed for any of the mutant strain relative to the wild type strain (Tezuka et al., 2009). More pronounced phenotypic changes, however, have been detected following the modulation of expression levels of other sRNAs in *S. coelicolor* (D'Alia *et al*., 2010; Vockenhuber and Suess, 2011).

In D'Alia *et al*. (2010), an antisense RNAs (*cnc2198.1*) was identified using RT-qPCR coupled with tagged-oligonucleotides, transcribed opposite the glutamine synthase-encoding *glnA* gene (*SCO2198*). The authors reported that the induction of expression of this antisense RNA reduced the levels of glutamine synthase (as determined by immunoblotting), as well as undecylprodigiosin and overall growth rate. A more recent study showed the scr5239 sRNA to be transcribed from the intergenic region between *SCO5238* and *SCO5239*

(Vockenhuber and Suess, 2011). scr5239 transcript levels were increased or decreased by expressing either its transcript or a complementary transcript, respectively, under the control of a strong promoter on a high copy-number plasmid. The level of overexpression reported for scr5239 (2.5 fold) is similar to that seen for our *scr4677* overexpression strain (~ 3 fold). scr5239 overexpression abolished the ability of *S. coelicolor* to utilize agar as a carbon source, while the scr5239 depletion strain retained this ability. Changes in growth rate and actinorhodin production were observed for both the *scr5239* depletion and overexpression strains relative to the wild-type. It was not clear, however, as to whether the "wild type" strain referred to here contains an empty-plasmid vector, and so it is possible that the high-copy plasmid alone could mediate the observed phenotypic changes. Although we did not detect any phenotypic changes as a result of scr4677 overexpression, we observed evidence for the impact of scr4677 on transcript levels of both *SCO4676* and *SCO4677-4676* read-through mRNAs.

3.4. Figures and Table

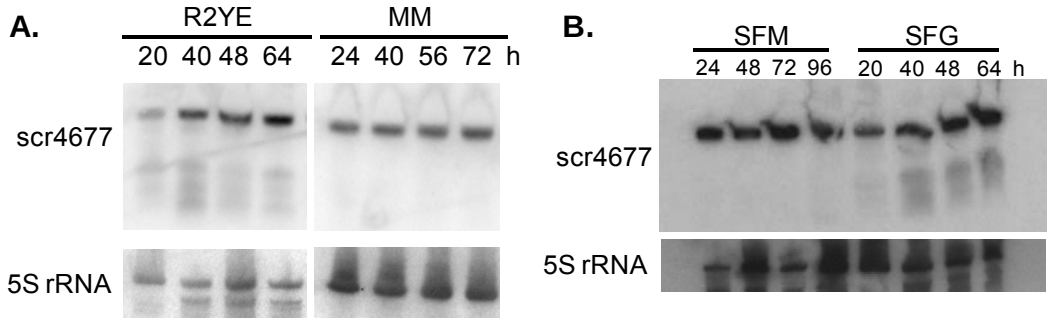


Figure 3.1. Transcription profiles of the scr4677 antisense RNA in *S. coelicolor* wild-type strain M145. (A) Northern blot analysis of *scr4677* expression during growth on R2YE (glucose-containing rich medium) and MM (minimal medium with mannitol) agar plates. RNA samples were isolated at the indicated time points. 5S rRNA served as a control for RNA levels and integrity. (B) *scr4677* expression during growth on soy flour agar media, containing glucose (SFG) or mannitol (SFM).

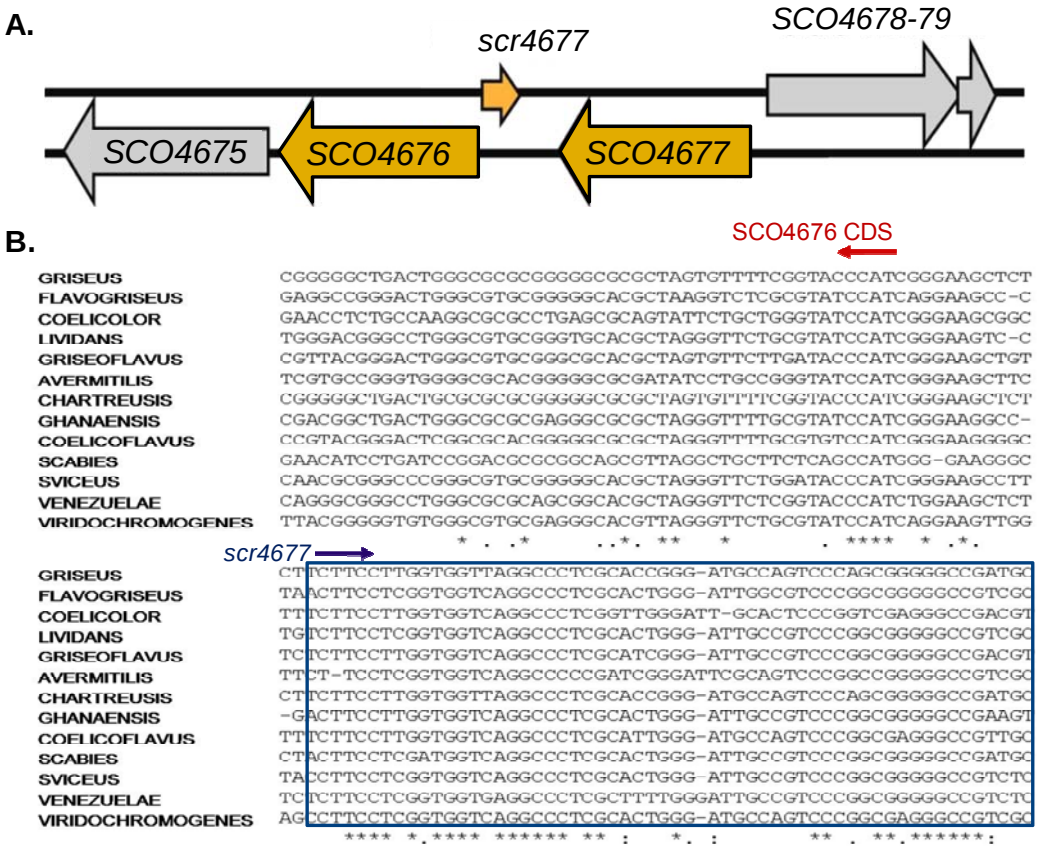


Figure 3.2. Bioinformatic analysis of the *scr4677*- and *SCO4676*-containing genetic region. (A) Location of *scr4677*, relative to the surrounding protein-coding genes. (B) Nucleotide alignment of *scr4677* (blue box) and the start of the *SCO4676* coding sequence (red bracket) among 13 divergent *Streptomyces* species. CDS = coding sequence; asterisks under the alignment indicate nucleotides conserved between all aligned species; periods or dots indicate conservation of purines or pyrimidines; and colons indicate conservation between A/T or G/C.

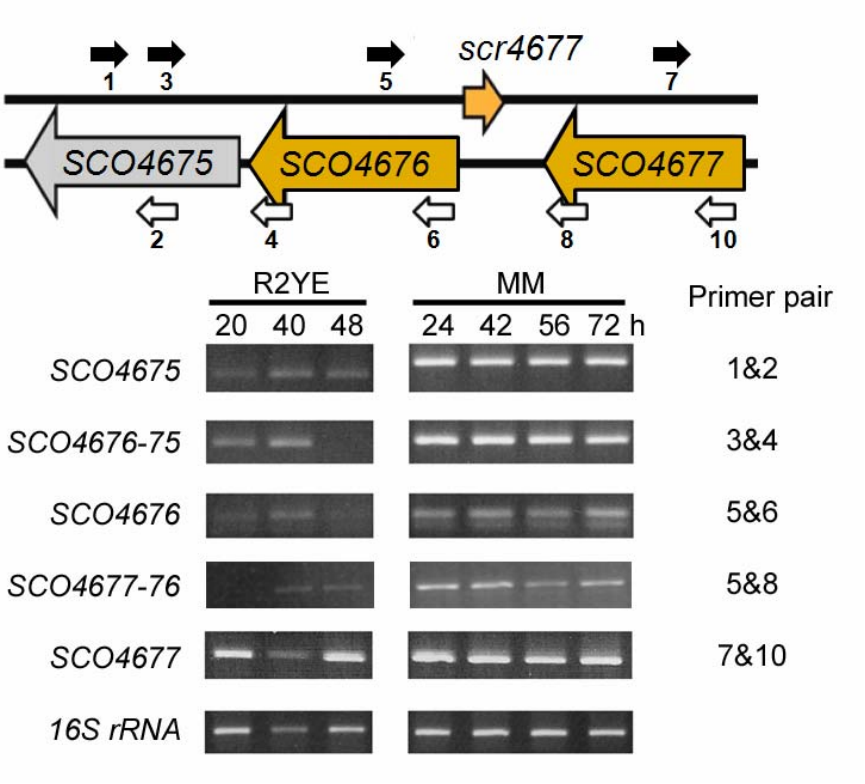


Figure 3.3. Expression profiles for *SCO4677-4675* genes in *S. coelicolor* wild-type strain M145 during growth on R2YE (glucose-containing rich medium) and MM (minimal medium with mannitol) agar media. Semi-quantitative RT-PCR was conducted using two biological replicates of RNA samples (at least three technical replicates), with the number of amplification cycles optimized for each transcript (*SCO4675, SCO4676-75, SCO4676, SCO4677*: 28 cycles; *SCO4677-76*: 32 cycles; and 16S rRNA gene: 15 cycles). The above schematic diagram depicts the relative locations of primers used for amplification, which are depicted as numbered black or white arrows. 16S rRNA served as a positive control for the RT-PCR and a control for RNA levels and integrity.

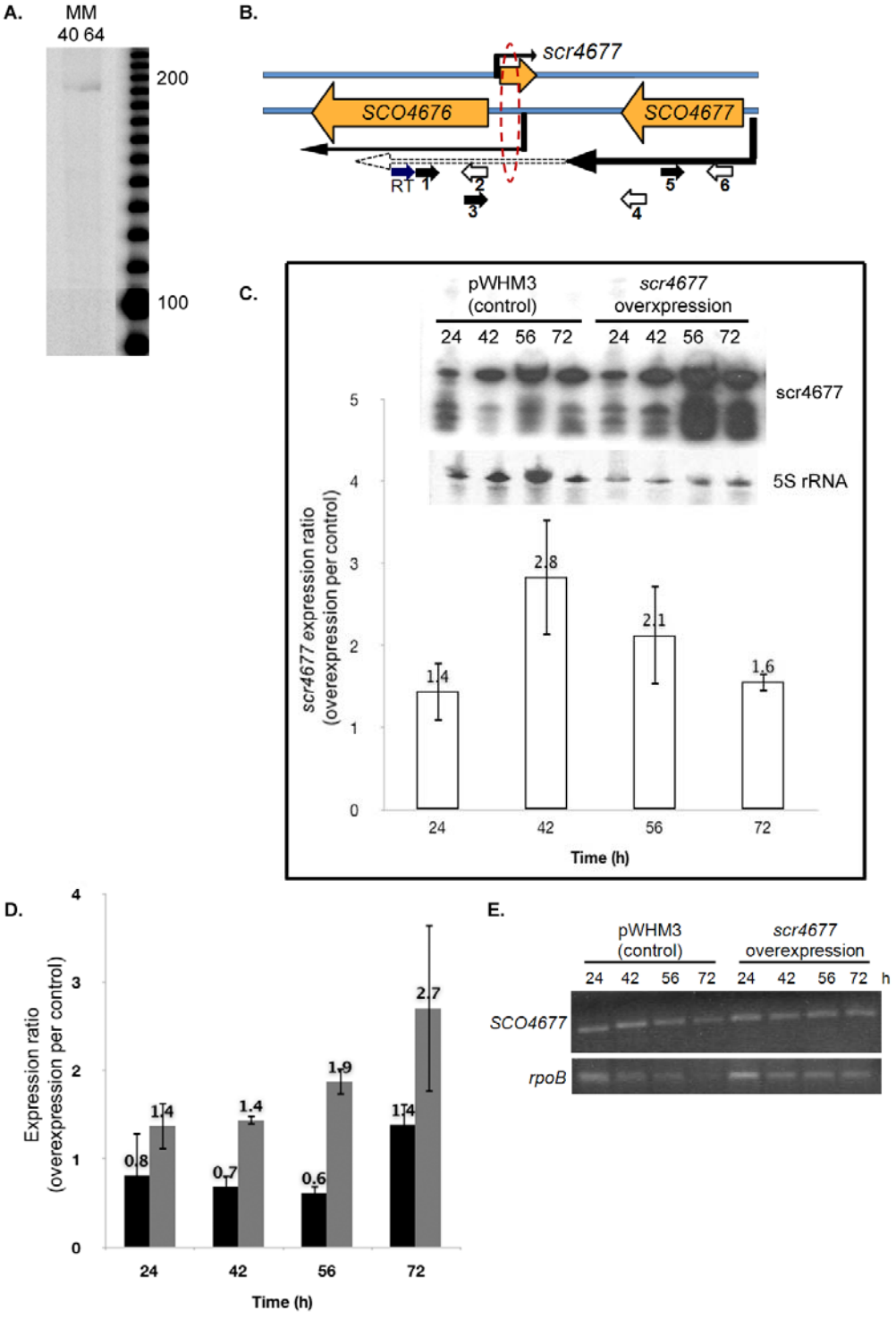


Figure 3.4. Expression of *scr4677*, *SCO4676*, *SCO4677-4676*, and *SCO4677*. (A) S1 nuclease mapping of the 5' end of the *SCO4676* transcript. RNA samples

were harvested from *S. coelicolor* wild-type strain after 40 and 64 h of growth on minimal medium+mannitol, and were run adjacent to a 10 bp ladder. (B) Schematic diagram depicting the approximate locations of promoters and corresponding transcripts, in the *SCO4676-4677* genetic region. A potential region of interaction between the scr4677 antisense RNA and mRNAs *SCO4676* and *SCO4677-4676* is circled with a broken red line. Approximate transcription start sites are indicated by vertical lines, with solid and broken lines (and associated arrows) reflecting transcription and read-through transcription, respectively. The relative locations of primers used in the RT-PCR experiments are indicated by numbered and coloured arrows (primers 1&4 were used for creating the *scr4677* overexpression construct; RT [blue]: a primer used for reverse transcription; primers 1&2: qPCR; and primers 3-6: semi-quantitative PCR [primer 5: RT of *SCO4677* transcript]). (C) Northern blot analysis of *scr4677* expression in the *scr4677* overexpression strain relative to the plasmid control-carrying strain during a 72h-time course on R2YE (rich) agar plates. The difference in expression of *scr4677* in the overexpression versus control strain is illustrated in the graph. Expression levels were determined by quantifying *scr4677* levels using ImageJ 1.42q (Abramoff *et al*., 2004), followed by normalizing these levels to 5S rRNA (RNA loading control). The data presented are the averages (± standard errors) from four experiments (two experimental replicates of two independent RNA time courses). (D) The difference in transcript levels of *SCO4676* (black bars) and *SCO4677-4676* read-through (grey

bars) between the *scr4677* overexpression and control strains. The transcript levels were normalized to *rpoB* levels (a control for RNA levels and integrity). For the semi-quantitative PCR used to assess the levels of *SCO4677-4676* read-through transcript, the intensity of fluorescent bands was quantified using ImageJ 1.42q. (E) Expression levels of *SCO4677* in the *scr4677* overexpression and control strains. The number of amplification cycles optimized for *SCO4677* and *rpoB* was 24 cycles.

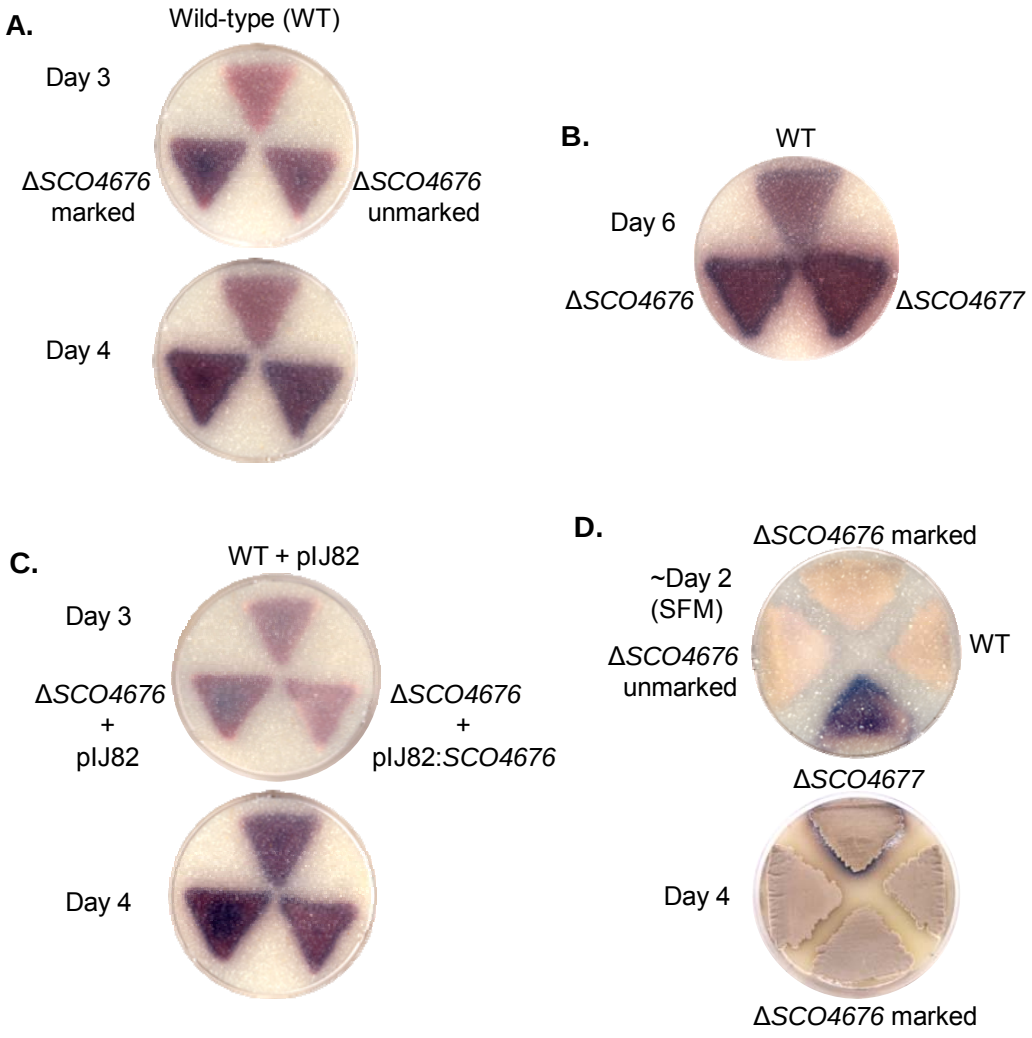


Figure 3.5. Phenotypic effects of *SCO4676* or *SCO4677* deletion during growth on soy flour agar media, supplemented with glucose (A-C) or mannitol (D). (A) Comparison between *S. coelicolor* wild-type (M145), *SCO4676* deletion (the gene was replaced by an apramycin resistance-marker), and marker-less *SCO4676* deletion strains. (B) Wild-type, *SCO4676* deletion, and *SCO4677* deletion strains. (C) Wild-type and *SCO4676* deletion strains containing the integrating plasmid pIJ82 and *SCO4676* deletion strain containing pIJ82 with a wild-type copy of

SCO4676. (D) Wild-type, *SCO4676* deletion, marker-less *SCO4676* deletion, and *SCO4677* deletion strains. The pictures were taken from the underside of the plates, except for the lower plate at panel D, after 2 to 6 days of growth, as indicated for each plate.

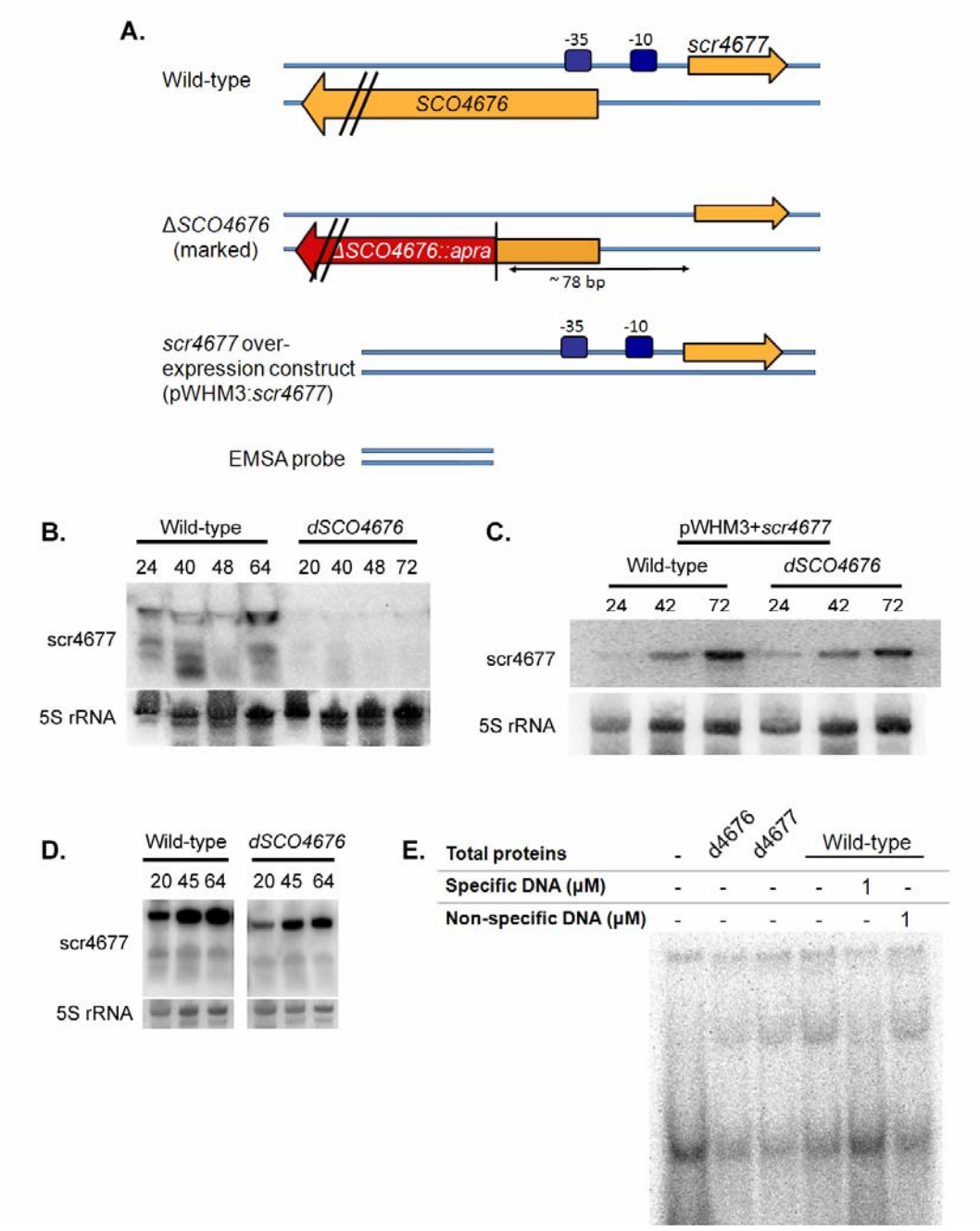


Figure 3.6. Expression of *scr4677* in *S. coelicolor* wild-type (M145) and mutants lacking *SCO4676* or *SCO4677*. (A) The position of an apramycin resistance gene replacing a majority of the *SCO4676* coding sequence (leaving the first 63 bp from the translational start site of SCO4676 intact). The *scr4677* overexpression

construct was cloned into the high copy vector pWHM3. The EMSA probe encompasses a potential regulatory element for *scr4677* transcription that was disrupted by the *SCO4676* deletion. (B) Northern blot analysis of *scr4677* expression in wild-type and the *SCO4676* deletion strains during growth on R2YE (rich) agar plates. The 5S rRNA transcript served as an RNA loading control. (C) Expression of *scr4677* in wild-type and *SCO4676* deletion strains carrying the *scr4677* overexpression plasmid. (D) Same as for (b), only comparing expression in the wild-type strain and its isogenic *SCO4677* deletion strain. (E) Electrophoretic mobility shift assay of [γ-^{32}P]ATP radiolabeled probe (83 bp) containing the *scr4677* upstream region together with crude lysates from *S. coelicolor* strains. The specificity of binding was tested using excess amount of specific (non-radiolabeled probe) and non-specific competitor DNA (α-*abeA* sequence).

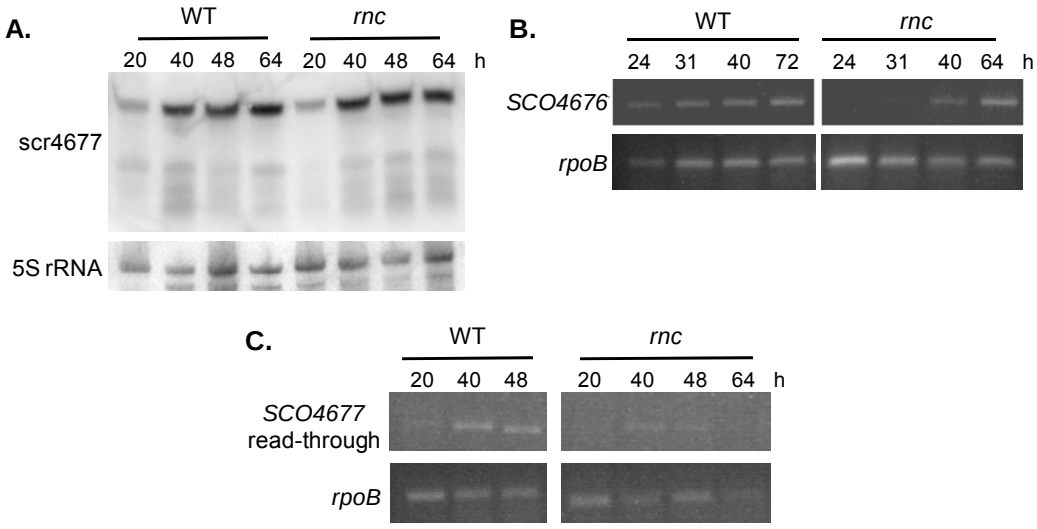


Figure 3.7. Expression levels of *scr4677, SCO4676*, and *SCO4677-4676* read-through in *S. coelicolor* wild-type and mutant strain lacking RNase III (*rnc*) during growth on R2YE (rich) agar plates. (A) Northern blot analysis of *scr4677* expression. The 5S rRNA transcript served as an RNA loading control. (B) Expression levels of *SCO4676*, determined by semi-quantitative PCR, in wild-type strain M600 and its isogenic *rnc* mutant strain. The *rpoB* gene served as an internal control. The numbers of amplification cycles optimized for *SCO4676* and *rpoB* were 28 and 24 cycles, respectively. (C) Same as (b), only comparing the levels of *SCO4677-4676* read-through in wild-type strain M145 and its isogenic *rnc* mutant strain. The number of amplification cycles used for *SCO4677-4676* read-through was 30 cycles.

Table 3.1. List of *Streptomyces* species possessing SCO4675, SCO4676, SCO4677 homologues- and scr4677-encoding genes

Species	*SCO4676*	*scr4677*	*SCO4677***	*SCO4675*
S. avermitilis	✓	✓	✓	
S. griseus	✓	✓	✓	
S. scabies	✓	✓		
S. venezuelae	✓	✓	✓*	
S. lividans	✓	✓	✓	✓
S. flavogriseus	✓	✓	✓*	✓
S. sviceus	✓	✓	✓*	
S. griseoflavus	✓	✓	✓	✓
S. viridochromogenes	✓	✓	✓*	
S. chartreusis	✓	✓	✓	
S. zinciresistens	✓		✓	
S. ghanaensis	✓	✓	✓	
S. himastatinicus	✓		✓*	
S. coelicoflavus	✓	✓	✓	
S. sp. S4				✓
S. albus				✓
S. griseoaurantiacus			✓*	✓
S. sp. e14			✓	
S. roseoporus			✓	
S. sp. W007			✓	

*) the *SCO4677* homologous gene is positioned neither next to *SCO4676* nor *SCO4675*-homologue encoding-gene

**) conservation of SCO4677 is beyond the species listed in this table.

Shaded rows indicate a number of species that have a genetic region harboring SCO4676, SCO4677 homologue- and *scr4677*-encoding genes

CHAPTER 4

REGULATION OF A NOVEL GENE CLUSTER INVOLVED IN SECONDARY METABOLITE PRODUCTION IN *STREPTOMYCES COELICOLOR*

Preface

The work presented in this chapter has been published in the Journal of Bacteriology (Hindra *et al*., 2010). The construction of *SCO3291* (*abeR*) deletion and α-*3287* (α-*abeA*) overexpression strains, and several plasmid constructs (pMC118, pMC119 and pMC147) was performed by Patricia Pak, M.Sc., who was an undergraduate student in the Elliot laboratory at the time. Most of the figures are from the published article, as authors of the article are given the right to re-use these figures in their theses.

4.1. Introduction

The secondary metabolism of *Streptomyces* has frequently been co-opted for use in medical and agricultural settings. For example, these organisms produce abundant compounds that have found widespread usage as antibiotics. *Streptomyces coelicolor*, a model system for studying *Streptomyces* genetics, produces calcium-dependent antibiotic (CDA) and methylenomycin, which was the first antibiotic found to be produced by *S. coelicolor* (Wright and Hopwood, 1976b). In addition, this species produces two pigmented antibiotics, the red

pigmented undecylprodigiosin and the blue pigmented actinorhodin, which make qualitative and quantitative assessments of antibiotic production extremely straightforward.

Antibiotic biosynthesis in the streptomycetes is dynamically determined by many regulatory inputs, including nutrient availability (Rigali *et al*., 2008), signaling molecules (Takano *et al*., 2006), and the proposed coupling of antibiotic synthesis and resistance (Tahlan *et al*., 2007; Le *et al*., 2009). Thus, antibiotic biosynthesis is controlled at multiple levels. This regulation involves the activity of several types of transcription factors: (1) pleiotropic regulators that influence both antibiotic biosynthesis and other cellular processes (*e.g.* colony development); (2) global regulators that control the production of multiple antibiotics; and (3) pathway-specific regulators that directly activate or repress the biosynthesis of a single antibiotic.

Genes required for antibiotic synthesis are typically clustered together, and frequently include a pathway-specific regulator (Bibb, 1996; Martin and Liras, 2010). By and large, pathway-specific regulators fall into one of two protein families (Bibb, 2005): the large ATP-binding regulators of the LuxR (LAL) family, which contain an NTP-binding domain within their N-terminal domain, and a DNA binding domain of the LuxR family at their C-terminus; and the *Streptomyces* Antibiotic Regulatory Protein (SARP) family, which has been detailed in chapter 1. Not all SARPs, however, are encoded within antibiotic biosynthetic gene clusters. For example, AfsR (one of the best studied SARPs) is

encoded at a genetic locus that is distant from any antibiotic gene cluster. AfsR activates the transcription of its immediate downstream gene *afsS* (or *afsR2* in *S. lividans*) (Lee *et al*., 2002), which encodes a small protein that affects the production of multiple antibiotics (Vögtli *et al*., 1994). In addition to *afsS*, AfsR also binds to promoter regions of PhoP regulons (*ptsS* and *phoRP*), which mediate phosphate control in *Streptomyces* (Santos-Beneit *et al*., 2009). The binding specificity of AfsR to its targets, shown by its specific interaction with direct heptameric repeats within the promoter regions of target genes, has also been well studied (Tanaka *et al*., 2007).

In this work, I focused on investigating the regulation of a previously uncharacterized genetic region *SCO3287-3291* (*abeABCDR*). This region harbors genes encoding a SARP-like regulatory protein, SCO3291 (AbeR), and a potential regulatory RNA, α-*3287* (α-*abeA*). Combining deletion and overexpression analyses with phenotypic characterization, I found that this genetic region contributes to secondary metabolism in *S. coelicolor*, and in particular, to actinorhodin biosynthesis. Using both genetic and biochemical approaches, I demonstrated that the *abe* genes were differentially regulated, and that AbeR regulated the expression of *abeBCD*. The regulatory role of AbeR was mediated by its DNA binding activity, which resembled the DNA-binding characteristics of other SARP-like proteins.

4.2. Results

4.2.1. Identification and transcriptional profiling of α-3287

Our interest in this genetic region came as a result of the identification of α-3287, an antisense RNA encoded on the opposite strand of the *SCO3287* coding region (Fig. 4.1a). α-3287 was transcribed exclusively during growth on rich (R2YE) agar medium, and its transcript abundance increased from 32 to 96 h, corresponding to a time when wild-type *S. coelicolor* raises aerial hyphae and develops spores (Fig. 4.1b). While this expression profile was reproducibly observed, it was not always possible to detect the α-3287 transcript in RNA isolated from R2YE medium. Rapid amplification of cDNA ends (RACE) experiments reproducibly mapped 56 and 62 nt-length transcripts, which were larger than the cloned cDNA fragment (~30-34 nt; Haiser *et al*., 2008). This suggests that the cloned cDNA fragments may represent a processed form of one of these two larger transcripts.

4.2.2. Bioinformatic analysis of the SCO3287-3291 *genetic region*

The *SCO3287-3291* genetic region contains an apparent four gene-operon, *SCO3287-3290*, together with the downstream *SCO3291,* which is expressed in the same direction (Fig. 4.1a). These genes all encode proteins of unknown function with the exception of *SCO3291*, which encodes a predicted regulatory protein. SCO3287 is an arginine-rich protein unique to *S. coelicolor* and its close relative, *S. lividans*. This protein (172 aa) does not have any obvious functional motifs, in contrast to SCO3288 or SCO3289 (243 and 536 aa, respectively),

which are each predicted to have a single membrane spanning region based on analyses using TMMOD and TMHMM servers (Kahsay *et al*., 2005; Krogh *et al*., 2001, respectively). SCO3290 (249 aa) has predicted nucleotide binding capabilities, and contains a Toll/Interleukin receptor (TIR)-like domain, which mediates recognition of invading pathogens in eukaryotic cells, and initiates a signal transduction cascade that ultimately activates innate immune systems (reviewed in O'Neill and Bowie, 2007).

SCO3288-3291 (but not SCO3287) share sequence identities and similarities (more than 44% and 60%, respectively) with proteins from other species of *Streptomyces* and the actinomycetes. Recently, genome sequencing of *S. hygroscopicus* subsp. jinggangensis (Wu and Bai, unpublished data) has revealed that this species likely possesses a similar genetic region (*SHJG7032-7035)* harboring *SCO3288-3291*-like genes. Two of these genes (*SCO3289-3290* and *SHJG7033-7034*) appear to be co-translated as a membrane-cytoplasmic protein complex containing TIR domain. This may suggest a possible role for these proteins in signal transduction (Martin and Wesche, 2002) or protein-protein interaction (Spear *et al*., 2009).

The N-terminus of SCO3291 has sequence motifs that are similar to the DNA-binding domain of OmpR (Martínez-Hackert and Stock, 1997) and the bacterial transcription activator domain or BTAD (Yeats *et al*., 2003). Unlike its N-terminus, the C-terminus of SCO3291 does not contain any known protein domains. Such an arrangement (N-terminal OmpR-like DNA-binding and BTAD

domains) is commonly seen in the SARP family (Wietzorrek and Bibb, 1997; Yeats *et al*., 2003). Most SARPs control the transcription of neighboring antibiotic gene clusters. However, SCO3291 has no obvious antibiotic-associated target, as it is encoded ~40 kb away from the nearest antibiotic biosynthetic gene cluster (the calcium-dependent antibiotic cluster).

4.2.3. Actinorhodin production increases upon gene/gene cluster overexpression

A genetic approach (gene deletion and gene overexpression) was undertaken to investigate the biological function for SCO3287-3291. Two deletion mutants, Δ*SCO3287-3290* and Δ*SCO3291*, were generated using PCR-targeted gene replacement (Gust *et al*., 2003), and were subjected to phenotypic analyses. During growth on a variety of solid media types, colony development and antibiotic production of the Δ*SCO3287-3290* mutant strain closely resembled that of wild-type *S. coelicolor*, whereas Δ*SCO3291* exhibited decreased actinorhodin production relative to its wild-type parent strain during growth on solid rich medium (R2YE) (Fig. 4.2a). To confirm that this phenotypic difference was due to the disruption of *SCO3291*, a wild-type copy of *SCO3291* on the integrating plasmid pMS82 was introduced into the mutant strain. This construct effectively restored wild type levels of actinorhodin production to the mutant strain (Fig. 4.2a), confirming that the loss of SCO3291 was responsible for the defect in antibiotic production.

Phenotypic comparisons were also conducted to see whether overexpressing *SCO3287-3291* (by inserting these genes into the high copy

plasmid pWHM3, and subsequently introducing them into *S. coelicolor*) had any effect on antibiotic production. Compared to the control pWHM3-carrying strain, the *SCO3287-3291* overexpression strain dramatically increased actinorhodin production on R2YE medium (Fig. 4.2b). To determine whether the effect on actinorhodin production was due to the overexpression of *SCO3287-3290*, α-*3287*, and/or *SCO3291* alone, strains overexpressing *SCO3287-3290,* α-*3287*, and *SCO3291* were created and tested for actinorhodin production during growth on R2YE. Actinorhodin production by the *SCO3291* and *SCO3287-3290* overexpression strains was enhanced relative to that of the control pWHM3-carrying strain (Fig 4.2b), whereas actinorhodin production was not affected by α-*3287* overexpression. This indicates that enhanced production of actinorhodin was due to the overexpression of one or more protein-encoding genes within the *SCO3287-3291* gene cluster.

Enhanced actinorhodin production by the overexpression strains was reproducibly observed on R2YE solid media; however, these strains did not consistently show increased actinorhodin production (relative to the control plasmid-carrying strain) during growth in liquid culture. Based on the contribution of the *SCO3287-3291* genetic region to actinorhodin production during solid culture growth, the five gene *SCO3287-3291* locus was named *abeABCDR*, respectively, with *abe* standing for antibiotic enhancement upon overexpression. The antisense RNA-encoding gene was accordingly named α-*abeA*.

4.2.4. Enhanced actinorhodin production by the SCO3287-3290 (abeABCD) *and* SCO3291 (abeR) *overexpression strains may be due to extended expression of* actII-ORF4

Many research groups (Aceti and Champness, 1998; Okamoto *et al*., 2003; Rigali *et al*., 2008) have observed that an impact on actinorhodin production can be mediated through the transcriptional control of the actinorhodin pathway-specific activator, *actII-ORF4*. To investigate whether increased transcription of *actII-ORF4,* and correspondingly the biosynthetic structural gene *actIII,* led to enhanced actinorhodin production by the *abeABCD* and *abeR* overexpression strains, a luciferase transcriptional reporter system (Craney *et al*., 2007) was used. Reporter constructs carrying *actII-ORF4* and *actIII* promoters fused to the *lux* operon, together with a promoter-less reporter construct (kindly provided by Ye Xu and Justin Nodwell) were introduced into the overexpression and control plasmid-carrying strains. After subtracting the luminescence output of the promoterless pMU1-carrying strains during their growth on R2YE agar plugs, luminescence levels of the *abeABCD* and *abeR* overexpression strains were compared to the pWHM3-carrying control strain (Fig. 4.3). Differences in the level of *actIII* promoter activities were not consistently detected. In contrast, *actII-ORF4* promoter activity in both *abeABCD* and *abeR* overexpression strains was significantly lower than that of the control pWHM3-carrying strain at 32 h, but at 72 h (when enhanced actinorhodin production by the overexpression strains was observed, Fig. 4.2b), *actII-ORF4* promoter activity in the overexpression

strains was significantly higher. This suggested that extended *actII-ORF4* expression may contribute to the enhanced actinorhodin production observed for the overexpression strains.

4.2.5. Transcription analysis of abeABCD *reveals a complex operon expression structure*

To dissect the regulation of the *abeABCD* genetic region during growth on R2YE agar, the transcription profiles of these genes were examined using semiquantitative RT-PCR. As shown in Fig. 4.4a, PCR failed to amplify *abeA* cDNA in any RNA samples isolated over a 3-day time course. This indicated that *abeA* was transcribed at a very low level. Transcripts of *abeB-abeC* and *abeC-abeD,* however, were readily detected during the same time course. The *abeB-abeC* cDNA (558 nt) was successfully amplified from 31 h RNA samples, when both aerial hyphae formation and secondary metabolism had initiated, through until 72 h. A similar transcription profile was also observed for *abeR*. Although *abeA* and *abeB* were initially thought to share the same promoter (as the coding regions were separated by a mere 8 bp), the transcription profiles of *abeA* versus *abeBCD* suggested that *abeBCD* was transcribed independently of *abeA*. The promoter activity of *abeA* is discussed below.

To examine whether *abeBCD* possessed an independent promoter located within the *abeA* coding region, S1 nuclease protection assays were conducted using a probe encompassing the first 64 nt of the *abeB* coding sequence and extending 334 nt upstream of the *abeB* translational start site. As shown in figure

4.4b, protection of the probe from S1 nuclease digestion (S1 nuclease degrades single stranded DNA) by the *abeB* transcript, suggested a transcriptional start site for *abeBCD* positioned ~77 nt upstream from the *abeB* translational start site. These results suggest that an independent promoter drives *abeBCD* transcription.

As shown in Figure 4.4c, we were able to detect *abeA* expression in the *abeABCD* overexpression strain, in which wild-type *S. coelicolor* strain carries the *abeABCD* overexpression plasmid (a high copy plasmid with *abeABCD* under the control of its native promoters). We were also able to detect read-through transcription from *abeA* into *abeB* (Fig 4.4c). Taken together, these findings suggest that *abeABCD* is transcribed constitutively at a low level, and that the upregulation of *abeBCD* transcription occurs coordinately with the initiation of aerial hyphae formation and secondary metabolism.

4.2.6. AbeR activates the expression of abeBCD *but not* α-abeA *transcription*

We observed *abeR* to have a similar expression profile to that of *abeBCD* (Fig. 4. 4a), and found the *abeR* and *abeABCD* overexpression strains had similar phenotypes (Fig 4.2b). Given these findings, and the fact that AbeR is predicted to be a DNA-binding protein, we wondered whether there was a regulatory link between AbeR and the other *abe* genes. To begin addressing this question, transcript levels of the *abeABCD* genes were compared in the *abeR* overexpression and deletion strains, relative to levels in the control plasmid-carrying strain and wild-type strain, respectively. We found the *abeR* overexpression strain had increased levels of *abeBCD* transcription relative to the

control strain (Fig. 4.5a), and that the *abeR* deletion strain expressed *abeBCD* at a low level, significantly less than that observed in the wild-type strain (Fig. 4.5b). Interestingly, transcription initiating upstream of *abeA* (and extending into the *abeBCD* cluster) was unaffected by AbeR deletion/overexpression. Taken together, it appears that *abeBCD* expression depends on AbeR.

EMSAs were conducted to further test whether AbeR regulated *abeBCD* directly by binding to its promoter region within the *abeA* coding region. Using crude lysates from *S. coelicolor* (the *abeR* deletion and *abeR* overexpression strains), potential protein-DNA complexes were observed (Fig. 4.5c). Given this finding, AbeR was overexpressed as a 6×Histidine tagged fusion protein, and purified using Ni-affinity chromatography. The purified His_6-AbeR bound to a DNA probe encompassing the *abeBCD* promoter region (Fig 4.5d). To determine whether this interaction was specific, competition experiments were conducted. A DNA fragment within the *S. coelicolor chpD* coding sequence (*chpD* is associated with aerial hyphae development but not antibiotic production [Elliot *et al*., 2003]) was chosen as a non-specific DNA competitor, and was unable to compete with the *abeBCD* probe to form a binding complex. Conversely, non-radiolabeled probe (specific competitor) effectively competed with the labeled probe for binding by AbeR.

Analysis of the EMSA probe sequence revealed four tandemly arranged direct repeats of 7 nt. To determine whether AbeR bound these repeated sequences, DNase I footprinting assays were conducted to map the regions

protected by AbeR binding. As shown in Fig. 4.5e and 4.5f, the AbeR-protected region contained the predicted four heptameric repeats, which were separated by 4 or 15 nt. Notably, regions of DNase I hypersensitivity were observed within the 15 nt-spacer, suggesting that binding by AbeR affected the DNA conformation in the area as well (Fig. 4.5f).

As shown in Fig. 4.5f, AbeR bound a region located between the promoter regions of *abeBCD* and α-*abeA*. The *abeB* and α-*abeA* transcriptional start sites were separated by ~210 nt. Therefore, it was possible that AbeR also regulated α-*abeA* transcription. To determine whether this might be the case, α-abeA transcripts levels were assessed in the *abeR* overexpression or deletion strains using northern blot analysis. There were no considerable differences in α-*abeA* expression observed, suggesting that AbeR did not influence α-*abeA* transcription. Taken together, these findings demonstrate that AbeR is an activator of *abeB-D* transcription.

4.2.7. RNase III has differential effects on the levels of the α-abeA transcript and its likely interacting partner, abeA-*containing mRNA*

The positioning of α-*abeA* within the *abeA* coding region (illustrated in Fig. 4.1a) and the mapping of α-abeA transcript ends by RACE experiments suggested that α-abeA most likely interacted with *abeA*-containing transcripts. To address whether RNase III was involved in processing a potential α-abeA - *abeA* duplex, we examined the expression of these two genes in both a wild-type strain and its isogenic RNase III null mutant (Δ*rnc*) strain carrying the *abeABCD*

overexpression plasmid (Table 2.2). The reason underlying the use of overexpression strains stemmed from the fact that α-abeA was not consistently detected at native level via northern blotting, and that *abeA* transcripts were challenging to detect when expressed at native levels. We found that α-abeA transcripts were more abundant in the RNase III mutant (relative to the wild-type strain) during a 96 h-time course, with the greatest difference observed at 72 and 96 h (Fig. 4.6a and 4.6b). This suggested that RNase III may target either the α-*abeA* transcript itself or a potential α-abeA - *abeA* duplex for degradation, and its loss serves to stabilize the antisense RNA. To differentiate between these possibilities, *abeA* transcript levels were examined. The ratio of *abeA* expression was determined by comparing its expression levels in the *rnc* mutant with that in the wild-type strain, and was normalized to the 16S rRNA gene and *tsr* (thiostrepton resistance) gene. At all time points, *abeA* transcript levels were lower in the mutant than in the wild type (Fig. 4.6b), suggesting that the transcript was not targeted by RNase III. Similar effects were observed for the *abeBCD* transcript (Fig. 4.6b), suggesting that these transcripts were not degraded together with α-abeA. These results also suggested no obvious correlation between α-abeA and *abeA* transcript levels; therefore, it appears that α-*abeA* has little effect on the stability of the *abeA* transcript.

The increased levels of α-abeA transcript in the *rnc* mutant were reminiscent of the upregulation of *ramR* and *sti1* (*SCO0762*) during the later stages of growth of an *rnc* mutant (Xu *et al*., 2010). Work by Xu *et al*. showed

that these effects were mediated by the pleiotropic regulatory protein AdpA (also known as BldH in *S. coelicolor;* Takano *et al*., 2003) (Ohnishi *et al*., 2005). AdpA is also required for *SCO0762* transcription (Kim *et al*., 2005). The *adpA* mRNA transcript is targeted by RNase III; therefore, an *rnc* mutant shows higher levels of AdpA than a wild-type strain, and therefore exhibits increased expression of AdpA regulon members (Xu *et al*., 2010). The AdpA-binding site has been reasonably well characterized, and a match to this sequence (TGGCCGGCCC versus TGGCSNGWWY, where S is G/C, W is A/T, and Y is T/C [Yamazaki *et al*., 2004]) was found upstream (-134 nt) of α-*abeA* (Fig. 4.5f). To determine whether α-*abeA* upregulation in the *rnc* mutant was mediated by AdpA, expression of α-*abeA* was assessed in an *rnc adpA* double mutant carrying the α-*abeA* overexpression plasmid. The double mutant had a similar α-*abeA* transcription profile when compared with the *rnc* mutant, suggesting that the increased transcript levels observed for α-abeA in the *rnc* mutant were not mediated by AdpA, but by another mechanism that remains to be determined.

4.3. Discussion

The *SCO3287-3291* (*abeABCDR*) genetic region was shown to play a role in activating/stimulating secondary metabolism during growth on rich R2YE solid medium. In particular, the *abeR* deletion strain was unable to produce the blue-pigmented antibiotic actinorhodin, and both *abeR* and the *abeABCD* overexpression strains exhibited enhanced actinorhodin production. Unlike many genes that affect actinorhodin production when cloned at high copy number (*e.g.*

afsR [Floriano and Bibb, 1996] and *metK* [Okamoto *et al*., 2003]), overexpression of *abeR* and *abeABCD* led to enhanced *actII-ORF4* expression at later stages of growth, rather than upregulating expression at all times.

Antibiotic production is regulated at multiple levels by many factors, including by physiological factors [*e.g.* growth stages, ppGpp levels (reviewed in Bibb, 2005) and metabolite precursor availability (reviewed in Stegmann *et al*., 2010)], as well as environmental cues [*e.g.* nutrient levels and signaling molecule availability (reviewed in Bibb, 2005)]. The products of the *abeBCD* genes appear to localize to the cell membrane, in the case of AbeB and AbeC, while AbeD appears to be a cytoplasmic protein. Therefore, it is possible that these proteins collectively act as a system for sensing environmental cues, such as small signalling molecules, and may act to transduce a signal from the extracellular environment to the cytoplasm. *Streptomyces coelicolor* produces at least four extracellular signalling molecules, namely SCB1 (Takano *et al*., 2001), SCB2, SCB3 (Hsiao *et al*., 2009) and MMF (Corre *et al*., 2008; O'Rourke *et al*., 2009). As these extracellular signalling molecules are expected to be readily diffusible across the cytoplasmic membrane, it seems unlikely that AbeBCD would be involved in mediating a cellular response to these molecules. These proteins may, however, have a role in sensing a currently unknown signalling molecule or environmental signal.

Given that the expression of *abeBCD* coincided with the initiation of secondary metabolism, AbeBCD might be responding to nutritional status. The

"enhanced actinorhodin production" phenotype was only observed when the *abe* genes-overexpressing strains were grown on glucose-containing medium (R2YE), suggesting that this phenotypic change is subject to catabolite repression. Although mechanisms of carbon catabolite repression in the streptomycetes are not fully understood (van Wezel *et al*., 2007), recent work has shown a connection between nutrient source and secondary metabolism by the regulatory proteins DasR (Rigali *et al*., 2008) and AtrA (Uguru *et al*., 2005).

DasR functions as a repressor of *actII-ORF4* and *redZ*, which are required for actinorhodin and undecylprodigiosin production, respectively (Rigali *et al*., 2008). DasR-repression is alleviated in the presence of glucosamine-6-phosphate (Rigali *et al*., 2006), a modified form of *N*-acetylglucosamine (GlcNac). DasR also binds to the promoter regions of genes required for importing GlcNac into the cytoplasm (Rigali *et al*., 2008). AtrA, on the other hand, is a direct activator of *actII-ORF4* transcription (Uguru *et al*., 2005), and interestingly, also activates the expression of *nagE2,* which encodes a permease required for importing GlcNac into cytoplasm (Nothaft *et al*., 2010). The *abe* genetic region appears to be unrelated to DasR or AtrA due to two reasons. First, there are no obvious DasR/AtrA binding sites upstream of *abeR,* and second, the upregulation of *actII-*ORF4 in the *abe*-overexpressing strains did not occur at early stages of growth but rather occurred during the later stages of growth.

Both *abeABCD* and *abeR* overexpression strains exhibited enhanced actinorhodin production, but unexpectedly, deletion of these genes did not lead to

similar phenotypes. While the *abeR* deletion strain lacked the ability to produce actinorhodin, the *abeABCD* deletion strain closely resembled the wild-type strain. This suggests that AbeR might have additional regulon members or other possible functions that are facilitated by its uncharacterized C terminus. Through DNase I footprinting assays, AbeR was found to bind four direct heptameric repeats, each of which was separated from the next by 4 or 15 nt, in the upstream region of *abeBCD* (located within the *abeA* coding region). The organization and spacing of these repeated sequences, CGGAA(G/C)C($n_{4/15}$)CGGAA(G/C)C, would effectively position the repeated regions on the same face of the DNA helix. One of the repeated sequences was located between the -35 and -10 sequences of the *abeB-D* promoter region; a specific space (8 nt) was also observed between this repeated sequence and the -10 sequences. Both the spacing relative to the promoter and the position of the repeated sequence in the promoter region are reminiscent of other SARPs binding sites (Wietzorrek and Bibb, 1997; Tanaka *et al*., 2007). Using the CGGAA(G/C)C($n_{4/15}$)CGGAA(G/C)C sequence as a pattern to search both chromosomal and plasmid DNA of *S. coelicolor*, a single hit was only found in the *abeA* coding sequence. This indicates that AbeR, like most other characterized SARPs, locally regulates the expression of its target genes.

Transcription analyses revealed differential expression of genes within the *abeABCDR* genetic region, with *abeA* (the first gene in the apparent four gene-operon) being expressed at very low levels and the downstream genes being expressed at higher levels. Unexpectedly, this was due to an active *abeBCD*

promoter within the *abeA* gene. Such differential expression of *S. coelicolor* genes within an operon has been reported previously in Charaniya *et al*. (2007), where the entire *S. coelicolor* genome was analyzed for its operonic structure. They describe the fact that differential expression can be affected by internal regulation, such as internal promoters, transcription factor binding sites, and transcription terminators. Therefore, *abeABCD* represents an excellent example of an operon exhibiting differential expression of the genes contained within.

Our observations, however, are at odds with analyses done using microarray-derived gene expression-data, which have shown that the first gene of an operon is generally expressed at a higher level than its downstream genes (Laing *et al*., 2006). More recent transcriptome analyses of other bacteria, including *E. coli* (Dornenburg *et al*., 2010), *Helicobacter pylori* (Sharma *et al*., 2010), *Listeria monocytogenes* (Toledo-Arana *et al*., 2009), and *Mycoplasma pneumoniae* (Güell *et al*., 2009), have provided unprecedented and holistic views into bacterial transcription. Our results are consistent with findings in *H. pylori* and *M. pneumoniae,* which revealed highly dynamic expression of different genes within an operon.

The presence of the α-*abeA* antisense RNA gene added another layer of transcriptional complexity in the genetic region. The stability (or expression) of α-*abeA* transcript was affected by RNase III, although this enzyme did not appear to cleave base paired α-abeA and *abeA* mRNA transcripts. At this stage, we cannot exclude the possibility that α-*abeA* regulates *abeA* expression at a

translational level. Abundant antisense RNA transcripts have now been detected in the comprehensive transcriptome analyses of *E. coli, H. pylori*, *L. monocytogenes*, and *M. pneumonia* (Dornenburg *et al*., 2010; Selinger *et al*., 2000; Sharma *et al*., 2010; Mraheil, *et al*., 2011; Güell *et al*., 2009). It seems reasonable to expect that the physiological role of antisense RNAs will be a fascinating research subject for many years to come. Taken together, the transcriptional complexity in the *abeABCDR* genetic region has become another example of non-canonical transcriptional dynamics in the bacterial kingdom.

4.4. Figures

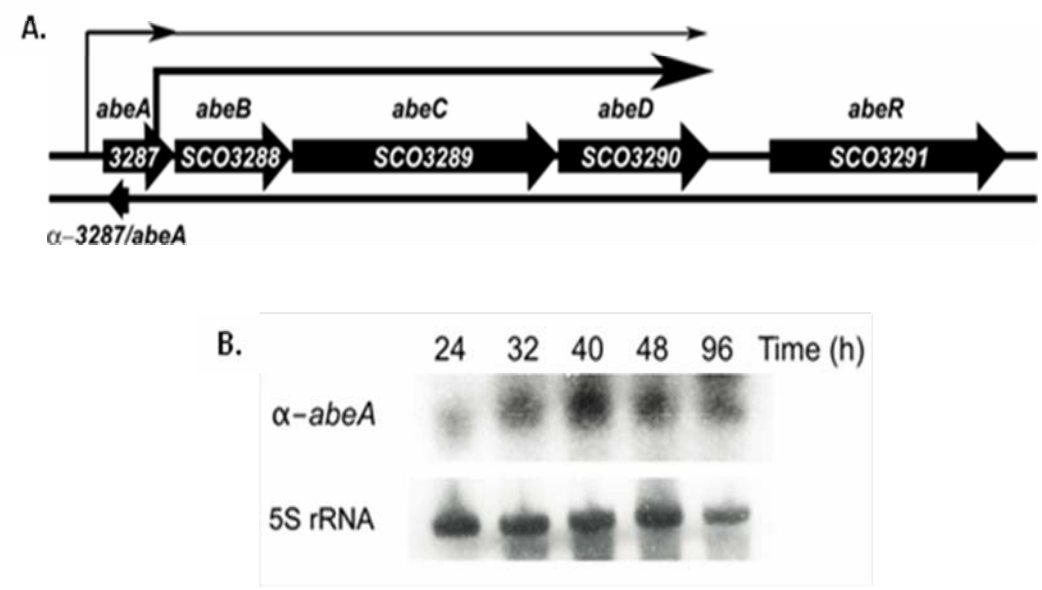


Figure 4.1. Transcription profiles of *α-3287* (or *α-abeA*) in *S. coelicolor* wild-type strain M600. (A) Genetic organization of *SCO3287-SCO3291* (the *abe* gene cluster) and the α-abeA antisense RNA-encoding gene. Transcription start sites are indicated by vertical lines, with the line widths (and associated arrows) approximating relative transcript abundance. (B) Northern blot analysis of α-*abeA* expression during growth on rich (R2YE) agar plates. RNA samples were isolated at the indicated time points. 5S rRNA served as a control for RNA levels and integrity.

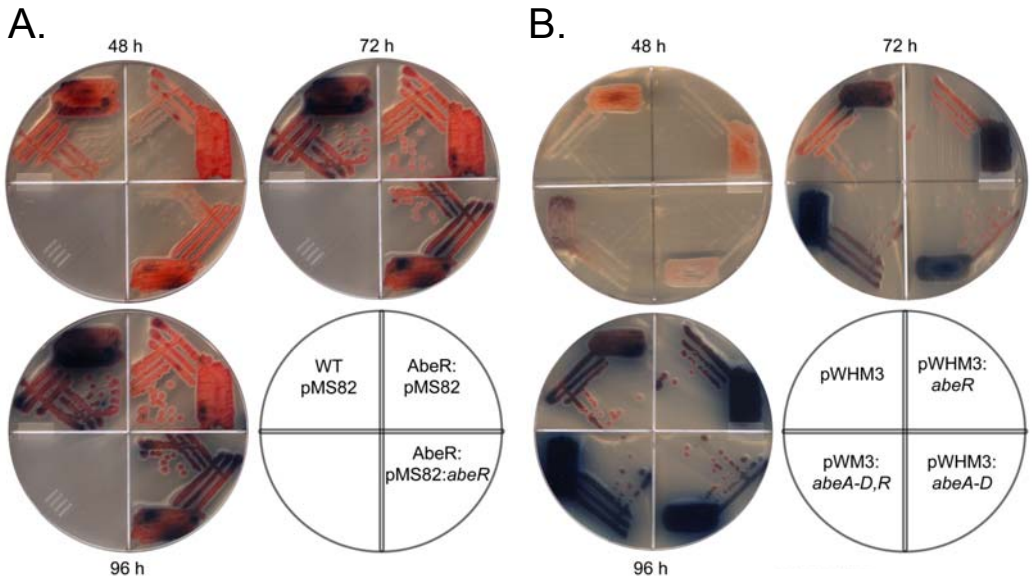


Figure 4.2. Phenotypic effects of *abeR* deletion and *abe* gene overexpression during growth on rich R2YE agar medium. (A) Comparison of wild-type *S. coelicolor* (M600) and *abeR* deletion strains carrying the empty integrating (single-copy) plasmid vector pMS82, with the *abeR* mutant strain carrying a wild-type copy of *abeR* in pMS82. (B) Wild-type strain containing the high-copy number plasmid pWHM3 alone or carrying *abeR* (*SCO3291*), *abeABCD* (*SCO3287-3290*), or both *abeR* and *abeABCD.* After 2 to 4 days of growth, pictures were taken from the underside of the petri plates. To ensure reproducibility, experiments were repeated at least four times using a minimum of two independent spore stocks as the inoculum.

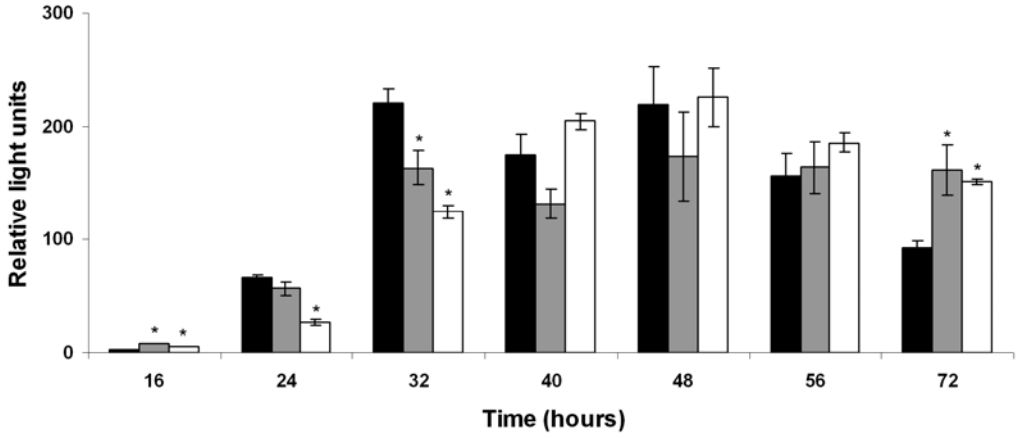


Figure 4.3. Expression of *actII-ORF4* as measured using the lux reporter system during a 72-h time course. Luminescence levels are shown as relative light units on the y-axis. Black bar: wild-type *S. coelicolor* containing the high-copy number plasmid pWHM3 (control plasmid); grey bar: wild-type strain containing pWHM3+*abeR*; white bar: wild-type strain containing pWHM3+*abeABCD*. All strains carried a chromosomally integrated pMU1 reporter plasmid containing the *actII*-ORF4 promoter cloned upstream of the *lux* genes. Strains were grown on R2YE agar plugs in 96 well plates. Luminescence levels reflect the means (+/- standard errors) from three independent experiments following subtraction of background levels (luminescence of each strain carrying pMU1 alone). Asterisks indicate those samples in which the mean luminescence levels were significantly different (at a 95% confidence level) between *abeR/abeABCD* overexpression plasmid and control plasmid (pWHM3)-containing strains.

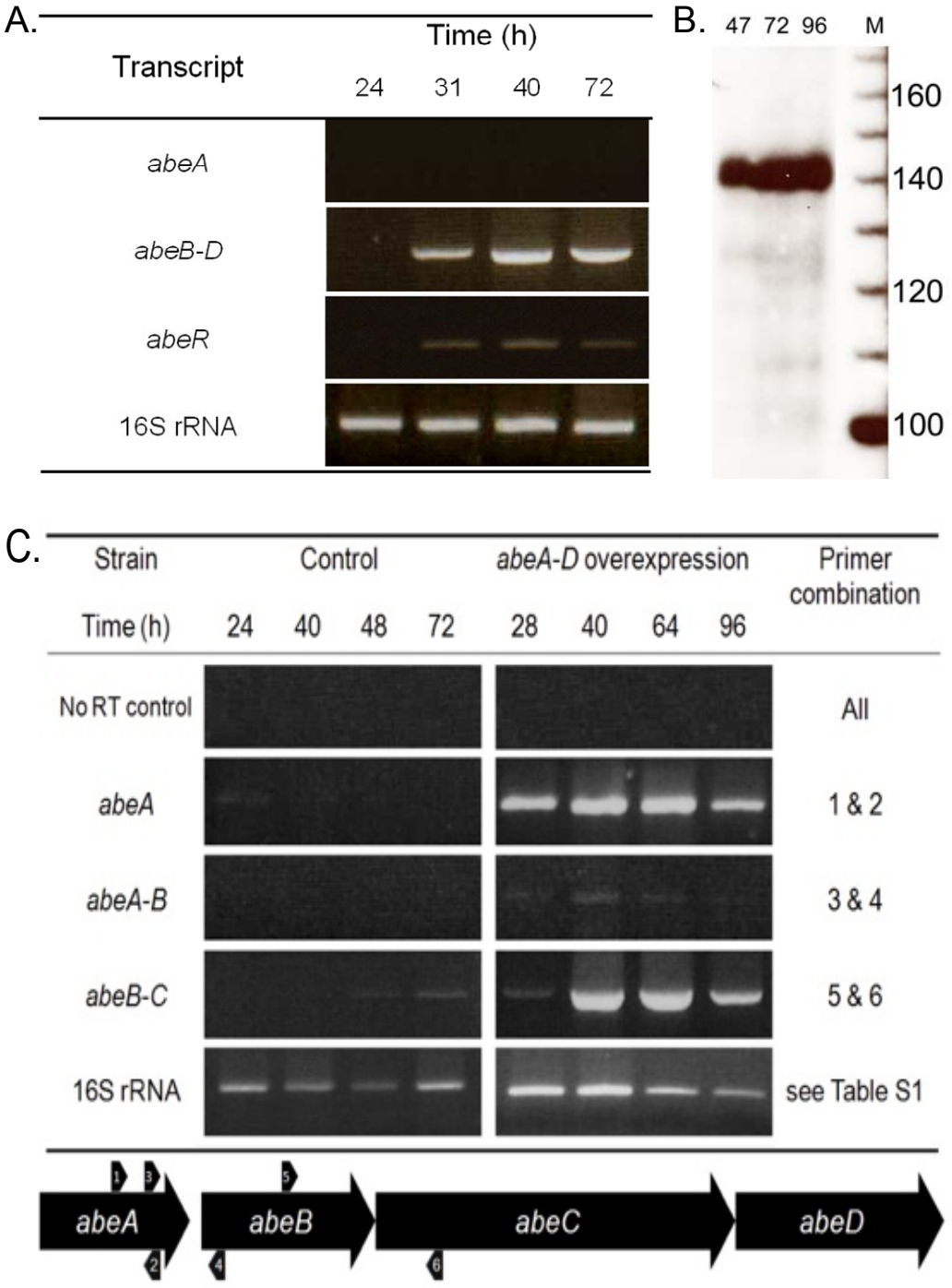
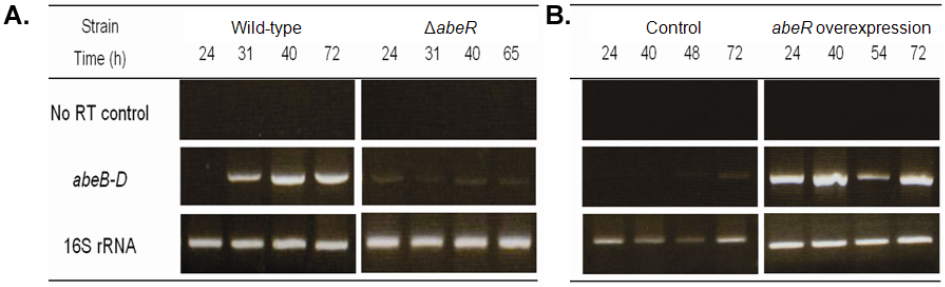
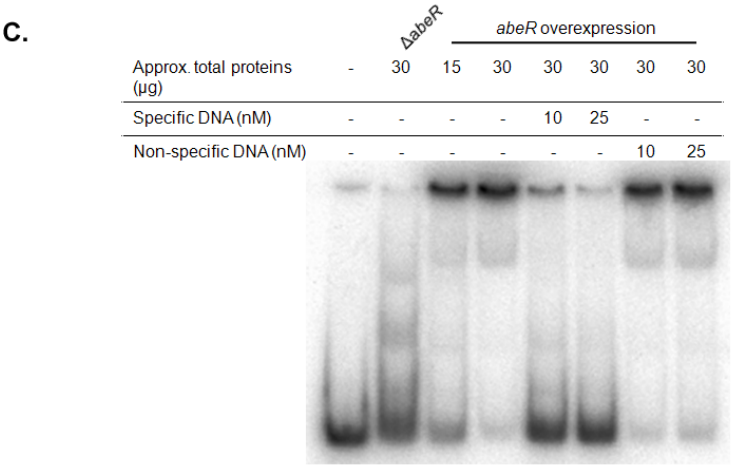
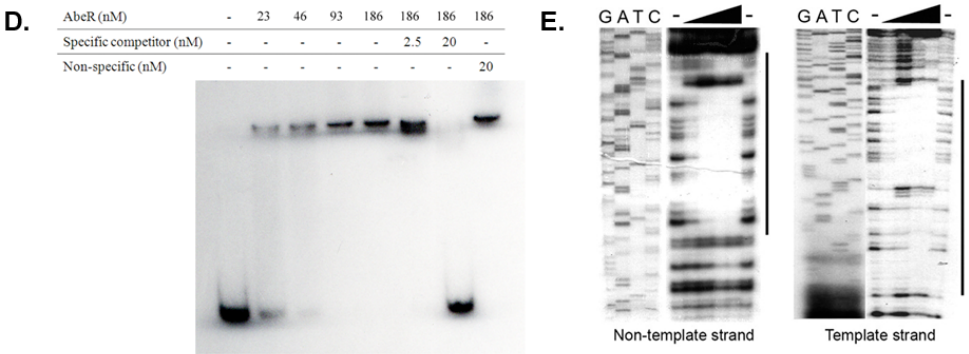


Figure 4.4. Transcription profiles for *abe* genes during growth on rich R2YE agar medium. (A) Semi-quantitative RT-PCR using RNA samples harvested from wild-type *S. coelicolor* (M600), with the number of amplification cycles having been optimized for each transcript (*abeA*: 27 cycles; *abeBCD*: 25 cycles;

abeR: 28 cycles; and 16S rDNA: 15 cycles). 16S rRNA served as a positive control for the RT-PCR reactions, and also as a control for RNA levels and RNA integrity. (B) S1 nuclease mapping of the 5' end of the *abeBCD* transcript. RNA samples were harvested from the *abeA-D* overexpression strain (strain E327, Table 2.1) after 47, 72, and 96 hours of growth, and were run adjacent to a 10 bp ladder (M). (C) Semi-quantitative RT-PCR using RNA samples harvested from the wild-type strains carrying the high-copy number plasmid pWHM3 alone (left panels) and pWHM3+*abeABCD* (right panels), with the optimized amplification cycles for each transcript (*abeA*, *abeAB, abeBC*: 27 cycles; and 16S rRNA gene: 15 cycles). The relative locations of primers used for amplification are indicated by numbered arrowheads in the schematic diagram. A negative control (No RT control) was included to ensure that there was no DNA contamination in any reaction.


A.
Strain
Time (h)
Wild-type
24 31 40 72
ΔabeR
24 31 40 65
No RT control
abeB-D
16S rRNA
B.
Control
24 40 48 72
abeR overexpression
24 40 54 72


C.
ΔabeR
abeR overexpression
Approx. total proteins (μg) - 30 15 30 30 30 30 30
Specific DNA (nM) - - - - 10 25 - -
Non-specific DNA (nM) - - - - - - 10 25


D.
AbeR (nM) - 23 46 93 186 186 186 186
Specific competitor (nM) - - - - - 2.5 20 -
Non-specific (nM) - - - - - - - 20
E.
G A T C - -
G A T C - -
Non-template strand
Template strand

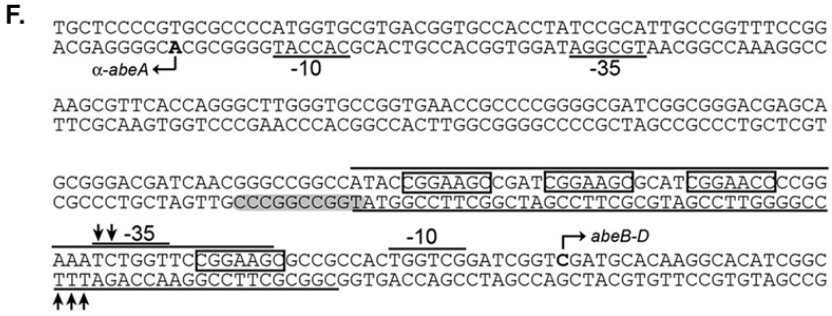

F.
TGCTCCCCGTGCGCCCCATGGTGCGTGACGGTGCCACCTATCCGCATTGCCGGTTTCCGG
ACGAGGGGCACGCGGGGTACCACGCACTGCCACGGTGGATAGGCGTAACGGCCAAAGGCC
α-abeA
-10
-35
AAGCGTTCACCAGGGCTTGGGTGCCGGTGAACCGCCCCGGGGCGATCGGCGGGACGAGCA
TTCGCAAGTGGTCCCGAACCCACGGCCACTTGGCGGGGCCCCGCTAGCCGCCCTGCTCGT
GCGGGACGATCAACGGGCCGGCCATACCGGAAGCCGATCGGAAGCGCATCGGAACCCCGG
CGCCCTGCTAGTTGCCCGGCCGGTATGGCCTTCGGCTAGCCTTCGCGTAGCCTTGGGGCC
-35
-10
abeB-D
AAATCTGGTTCCGGAAGCGCCGCCACTGGTCGGATCGGTCGATGCACAAGGCACATCGGC
TTTAGACCAAGGCCTTCGCGGCGGTGACCAGCCTAGCCAGCTACGTGTTCCGTGTAGCCG

Figure 4.5. Regulation of *abeBCD* trancription by AbeR. Transcription profiles of *abeBCD* in (A) *S. coelicolor* wild-type and *abeR* deletion strains, and (B) wild-type strains carrying the high-copy number plasmid pWHM3 (control) alone and pWHM3+*abeR* (*abeR* overexpression), as determined using semi-quantitative RT-PCR. RNA samples used in the RT-PCR were harvested during growth on R2YE agar plates (rich medium); the number of amplification cycles was optimized for each experiment: 25 cycles for (A) and 27 cycles for (B). For the "No RT control" PCR, RNA (not subjected to reverse transcriptase) served as template to ensure no DNA contamination. The 16S rRNA gene was amplified (15 cycles) as a control for RNA levels and RNA integrity. (C) Electrophoretic mobility shift assay (EMSA) of [γ-^{32}P]ATP radiolabeled probe (214 bp) containing the *abeBCD* upstream region together with crude lysates from *S. coelicolor* strains. AbeR binding specificity was tested using specific (non-radiolabeled probe) and non-specific competitor DNA (*chpD* coding sequence). (D) Similar experiment shown in (c) using His_6-AbeR purified from *E. coli*. (E) DNase I footprinting assay of His_6-AbeR binding to the *abeBCD* promoter region. Protected regions are indicated by the vertical lines to the right of each footprint. Increasing concentrations of purified AbeR (0-1.67 μM, with increasing concentrations of protein indicated by the black triangle) were incubated together with 6 nM of singly end-labeled probe. Sequencing ladders are shown to the left of each footprint. (F) Coding sequence of *abeA* encompassing the approximate -10 and -35 promoter sequences of α-*abeA* and *abeBCD*. Mapped transcriptional start sites for both transcripts are designated by bold letters. DNA sequences protected by AbeR are indicated by horizontal lines, and the protected four direct heptameric repeats are boxed. Sites of DNase I hypersensitivity are indicated by vertical arrows, while a potential AdpA binding site is shaded in grey.

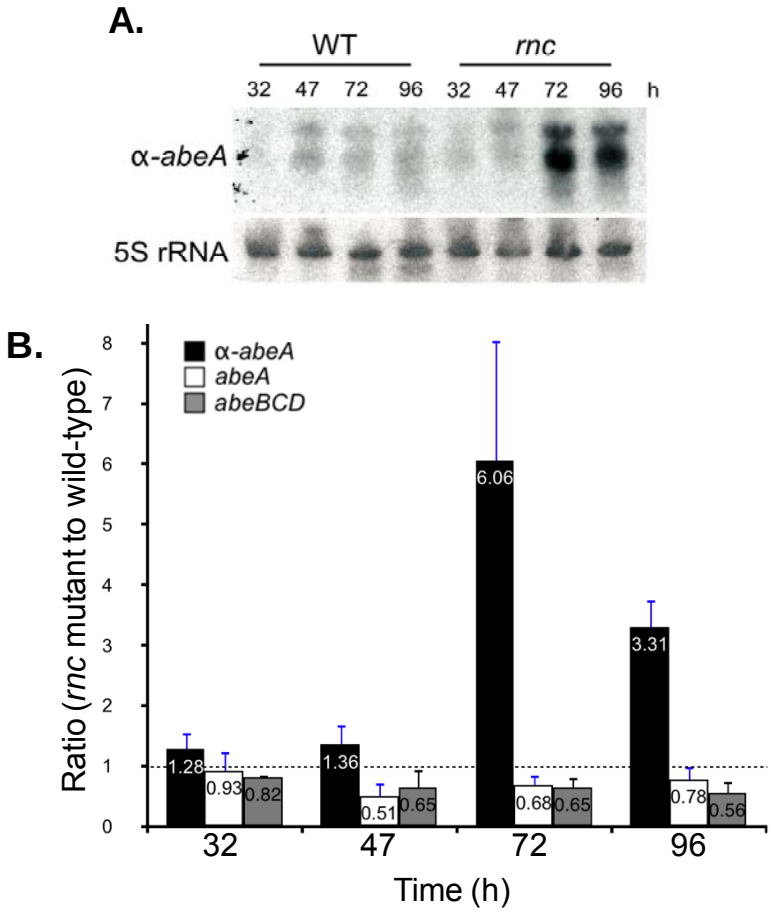


Figure 4.6. Expression of *α-abeA, abeA,* and *abeBCD* in *S. coelicolor* wild-type (strain M600) and mutant strain lacking RNase III (*rnc/absB*) during growth on R2YE (rich) agar plates. (A) Northern blot analysis of α-*abeA* expression in wild-type and *rnc* mutant strains carrying the *abeABCD* overexpression plasmid pMC116 (Table 2.2). The 5S rRNA transcript served as an RNA loading control. (B) The ratio of α-abeA (black bars), *abeA* (white bars) and *abeBCD* (grey bars) transcript levels between the *rnc* mutant and wild-type strains carrying pMC116. Transcript levels of α-abeA were normalized to 5S rRNA. For *abeA* and *abeBCD,* the levels were normalized to 16S rRNA gene and *tsr* levels to account for differences in input RNA samples and levels of plasmid copy number, respectively. The intensity of bands observed in northern blotting and semi-quantitative PCR was quantified using ImageJ 1.42q (Abramoff *et al*., 2004).

Error bars indicate the standard errors from the average ratio obtained from four repeated experiments (two experimental replicates of two independent RNA time courses).

CHAPTER 5

SUMMARY AND FUTURE DIRECTION

The identification of small RNAs (sRNAs) in the streptomycetes was firstly reported by Panek *et al*. (2008). They used bioinformatic searches to predict 32 sRNA candidate sequences, and a subset of these sequences was verified using RT-PCR and microarray analyses. Additional *S. coelicolor* sRNAs were identified in our study (Swiercz *et al*., 2008) using both bioinformatic- and cloning-based approaches. In our work, we detected nine sRNAs using northern blotting. The transcript ends for each were mapped using a combination of RACE experiments and northern blotting using "primer walking" with overlapping probes. Two of these identified sRNAs were also identified in Panek *et al.* (2008).

After these two reports of sRNA identification, additional sRNAs were identified in both *S. griseus* (Tezuka *et al*., 2009) and *S. coelicolor* (D'Alia *et al*., 2010, Vockenhuber *et al*., 2011). The latter two studies reported and verified a number of antisense RNAs in addition to *trans*-encoded sRNAs. The increasing number of antisense RNAs being discovered in the streptomycetes and other bacteria, collectively, implies an obvious question as to physiological role of these RNA molecules. Here, we have begun our exploration into the functioning of two antisense RNAs, as outlined in chapters 3 and 4.

5.1. Regulation of the *SCO4677-4675* genetic region containing the *scr4677* antisense RNA gene

The work described in chapter 3 examined a potential role for scr4677 in regulating expression of its flanking genes, *SCO4676* and *SCO4677*. We observed that nutritional status affected the transcription profiles of *scr4677* as scr4677 transcripts were more abundant when *S. coelicolor* was cultured on R2YE (rich medium) than on minimal medium. Prior to investigating function of scr4677 within the genetic region, we profiled the expression of its flanking genes. While expression of *SCO4677* was not affected by carbon source/media composition, the expression of *SCO4676* was impacted. In contrast to *scr4677,* transcript levels of *SCO4676* were less abundant when cultures were grown on R2YE, relative to the levels when they grown on minimal medium. We hypothesized that this differential expression was mediated by scr4677. Intriguingly, *SCO4676* transcription stemmed from a promoter located immediately upstream, as well as from read-through transcription from *SCO4677* (despite the fact that *SCO4676* and *SCO4677* are separated by ~250 bp). The abundance of the *SCO4677-4676* read-through transcript was also reduced during growth on R2YE.

To begin to understand the role of *scr4677* in the cell, we chose to increase the expression of *scr4677 in trans* through the cloning of its gene into a high copy plasmid, and followed its expression and the resulting transcript levels of *SCO4677, SCO4676,* and *SCO4677-4676* mRNAs. We observed different

effects of *scr4677* overexpression on each mRNA transcript (section 3.4), and a full understanding of this will require additional investigation into the mechanism of action. As it was possible that scr4677 promoted the destabilization of the *SCO4676*-specific mRNA, we have investigated whether this process involved RNase III (a double-stranded RNA-specific nuclease) mediated-degradation of base-paired scr4677-*SCO4676* transcripts. This enzyme did not seem to be involved, suggesting that other RNases may contribute in this case (discussed below).

As we were overexpressing scr4677 *in trans,* it seems unlikely that transcription interference (due to colliding RNA polymerase enzymes or altered local supercoiling) could explain our observations. Alternatively, the reduced levels of *SCO4676* transcripts in the *scr4677* overexpression strain could be an indirect result of the altered accessibility of ribosomal binding on *SCO4676* mRNA (discussed in section 3.3). To further investigate this possibility, future experiments can be directed to use *in vitro* toeprinting assays as detection of SCO4676 has not been successful using western blotting. The toeprinting assays will involve the isolation of ribosomal 30S subunits from *S. coelicolor* and *in vitro* transcription of *SCO4676* and *scr4677*. After ribosomal loading on the RBS (upon addition of $tRNA^{fMet}$), the translation initiation reaction is followed by reverse transcription, and this would extend cDNA until being terminated at +15/+16 from the start codon (+1). If scr4677 binding interferes with the docking of ribosomal 30S subunits on the RBS of *SCO4676*, the cDNA synthesis would be

expected to extend beyond the +15/+16 position. Therefore, this approach will serve as an alternative methodology to examine the mechanism of action of scr4677.

5.2. Regulation of the *SCO3287-3290 (abeABCD)* genetic region containing the *α-3287 (α-abeA)* antisense RNA gene

We undertook a similar approach to examine whether the α-abeA antisense RNA transcript had an impact on the expression of *abeA*. Differential expression of genes in this region was observed, and this was mainly due to an internal promoter driving the expression of *abeBCD* (section 4.2.5). While α-abeA did not appear to be involved in regulating the expression of *abeA*-containing transcripts at transcriptional/post-transcriptional levels, we cannot exclude the possibility that α-*abeA* may affect the translatability of *abe* mRNA at this point. We found that the product of the immediately downstream gene *abeR* activated the transcription of *abeBCD* directly, but not *abeABCD* or α-*abeA*. Further biochemical tests revealed that AbeR bound target sites within the coding sequence of *abeA*.

Transcriptional profiling showed that expression of *abeR* and *abeBCD* occurred at time when secondary metabolism had been initiated. Following genetic modification through deletion and overexpression of these genes, phenotypic analyses were conducted. We observed that overexpressing *abeABCD* and the SARP (*Streptomyces* antibiotic regulatory protein)-like protein-encoding

gene *abeR* contributed to secondary metabolism, particularly enhancing actinorhodin production. Interestingly, deleting *abeR* resulted in reduced actinorhodin production, while deleting the entire *abeABCD* gene cluster did not affect its production. We examined whether enhanced actinorhodin production following overexpression of *abeABCD* or *abeR* was the result of changes in *actII-ORF4* expression, where *actII-ORF4* encodes a transcriptional activator of the actinorhodin biosynthetic gene cluster (Arias *et al*., 1999). We observed that overexpressing either *abeR* or the entire gene cluster *abeABCD* did not increase the overall expression of *actII-ORF4*, but it extended the expression of *actII-ORF4* at later stages of growth.

As briefly mentioned in section 4.2.2, the products of *abeBCD* may be involved in signal transduction or protein-protein interaction. Future work can be aimed at verifying interactions among these gene products. The experimental approaches include two-hybrid techniques or the isolation of AbeB/C/D interacting partners using immunoprecipitation or pull down assays. Initial experiments can be focused on AbeD since it is likely a cytoplasmic protein. AbeD-specific antibodies or tag fusion variants of AbeD could be generated for use in subsequent immunoprecipitation or pull down assays, respectively. These assays coupled with mass spectrometry can be assigned to identify interacting partners of AbeD. Verification of interaction between AbeB, AbeC and AbeD, as well as identification of AbeD-interacting partners will ultimately provide new

insight of AbeD function, in addition to establish a more solid connection between the gene products of *abe* genes.

5.3. RNase III-independent degradation of the potential base-paired scr4677-*SCO4676* or α-abeA-*abeA* transcripts

In the event that antisense RNAs and their cognate sense mRNAs form a duplex structure that can be targeted by RNase III (Chang *et al*., 2005), both transcripts should be more abundant in a strain lacking RNase III (Δ*rnc/absB*). In sections 3.2.7 and 4.2.7, we showed decreased abundance of *SCO4676, abeA*, and *abeBCD* transcripts in an *rnc* mutant. This effect is reminiscent of the reduced expression of genes in the antibiotic biosynthetic clusters of actinorhodin, undecylprodigiosin, calcium-dependent antibiotic, and a cryptic polyketide observed in an *absB* mutant (Huang *et al.*, 2005). Although our findings were not unusual, our data suggest that RNase III does not appear to mediate the destabilization of potential duplex of scr4677-*SCO4676 or* α-abeA-*abeA*.

Given that it is widely accepted that RNase III plays an important role in processing 30S rRNA to 23S and 16S rRNA, it is possible that a strain lacking RNase III may be impacted in its functional ribosome, which can lead to the premature decay of ribosome-unbound mRNAs (reviewed in Deana and Belasco, 2005; Kaberdin and Bläsi, 2006). Thus, it is not surprising to observe reduced expression of many genes (or reduced levels of mRNA transcripts) in an *rnc* mutant. Another possibility is that transcription factors of these genes are negatively or positively affected by RNase III.

While it is possible that RNase III may contribute to the degradation of other pairs of antisense-mRNA transcripts in *S. coelicolor*, it is also worth considering other alternatives. Interaction between scr4677 and *SCO4676*, despite the extent of complementarity between the two transcripts, may expose a region that could be recognized by RNase E. It is well established that, in *E. coli,* RNase E degrades many sRNA-targeted mRNAs (sections 1.6 and 1.7). Therefore, it may be useful to probe further into the essential nature (or not) of RNase E in *S. coelicolor.* If it turns out that RNase E is an essential protein, it would then be reasonable to create individual RNase E depletion strain, *e.g.* by disrupting RNase E-encoding gene in the presence of second copy of this gene under control of inducible-promoter. This would allow us to examine whether these proteins are involved in sRNA-mediated RNA degradation in *S. coelicolor.*

Notably, *B. subtilis* does not have an RNase E homologue. Instead, it possesses RNase J enzymes, which are suggested to be the functional equivalent of RNase E in terms of their preference for cleaving substrates with monophosphorylated 5′ end, despite the fact that they share little sequence similarity with RNase E (Even *et al.*, 2005). Additionally, RNase J1 has 5′-to-3′ exoribonuclease activity (Mathy *et al.*, 2007), which could play a role in degrading the 5′ ends of mRNAs, including sRNA-targeted mRNAs. *Streptomyces coelicolor* has a single RNase J orthologue (encoded by *SCO5745*), with 39% identity and 61% similarity to the *Bacillus* protein. Since it is not known yet whether SCO5745 serves as an endoribonuclease and/or

exoribonuclease in *S. coelicolor*, further characterization of SCO5745, to investigate its putative nuclease activity, would be another point for future direction.

5.4. Towards understanding the physiological role of antisense RNAs in *S. coelicolor*

Through genetic and transcriptional analyses in our work, we propose that scr4677 may regulate *SCO4676* expression; however, the physiological roles of scr4677 and SCO4676 are not yet understood. Modulation of scr4677 transcript levels, followed by a comprehensive phenotypic analysis, did not show any phenotypic changes relevant to scr4677. This is not unexpected since sRNAs often function to "fine-tune" the expression of target genes (reviewed in Gottesman and Storz, 2011). Other *S. coelicolor* studies have reported that increasing the expression of several sRNAs had phenotypic effects on growth rate, antibiotic production, and agar utilization (section 3.3); however, the mechanistic details of these sRNA-mediated effects are also not yet understood. Using a genome-wide approach (RNA-Seq), a coworker in the Elliot laboratory has identified a collection of antisense RNAs (non-coding RNAs in general) in several *Streptomyces* species and started to analyze many candidate mRNA targets. Collectively, further characterization of known antisense RNAs and additional sRNAs will be necessary to improve our understanding of their relevance to *Streptomyces* biology.

APPENDIX

Table 2.3. Oligonucleotides used in this study

Name	Sequence (5' to 3′)	Function
KO3287CL-1	CGGCCGGCCGCCGCCGATGTTGGTGGGGTGACGCCCGTGATTCCGG GGATCCGTCGACC	Creation of *abeABCD* (*SCO3287-3290*) knockout
KO3287CL-2	ATCCGCCGGACGACTGCGCCGGCGGTGGCAGGAGGCCTATGTAGG CTGGAGCTGCTTC	Creation of *abeABCD* (*SCO3287-3290*) knockout
3287-1	CTTGCCCTGCTCCACCTTGC	PCR and Southern blot analysis to confirm *abeABCD* (*SCO3287-3290*) knockout
3290-2	CCACTCGCTTCACACTCTCC	PCR and Southern blot analysis to confirm *abeABCD* (*SCO3287-3290*) knockout
3288-F	CCGACGACACGGCCACCGAGGACACCACC	PCR analysis to confirm *abeABCD* null mutation; RT-PCR for *abeBCD* (*SCO3288-3290*)
MIDCL-R	CGAGGGCGATCAGCTGGTTGCGGCGCAGG	PCR analysis to confirm *abeABCD* null mutation; RT-PCR for *abeBCD* (*SCO3288-3290*)
3287-2	CCTTGTGCATCGACCGATCC	RT-PCR for *abeA* (*SCO3287*); PCR amplification of probe for EMSA and DNaseI footprinting
3287-3	ATGGTGCGTGACGGTGCCAC	RT-PCR for *abeA*; PCR amplification of probe for EMSA
KO3291-1	GGTCGGGAGAGTGTGAAGCGAGTGGAGACCACTTGTATGATTCCG GGGATCCGTCGACC	Creation of *abeR* (*SCO3291*) knockout
KO3291-2	AAGCGGGGTCGCTGATGCCGTAAGGGCATGTCGACGTCATGTAGG CTGGAGCTGCTTC	Creation of *abeR* (*SCO3291*) knockout
3291-1	CGCACGAACCGACCGAAACC	PCR analysis to confirm *abeR* knockout
3291-2	GATGCCGTAAGGGCATGTCG	PCR analysis to confirm *abeR* knockout
3291-R	GAACAGGTCGCCGCGCACC	PCR analysis to confirm *abeR* knockout
3291-1IN	GTATCCGTGTGAGCGTCCTG	RT-PCR for *abeR*
3291-2IN	TTTGCACTTCGTTGCGGTGG	RT-PCR for *abeR*
3287-5	ATCGGAAGCGCATCGGAACC	RT-PCR for *abeABCD* read-through transcription
3287-6	TGGCGGGTGCTCCGTCGTC	RT-PCR for rough mapping of *abeBCD* promoter
3287-7	GGATACCGGCGGCAATGAGG	RT-PCR for rough mapping of *abeBCD* promoter

Table 2.3. Oligonucleotides used in this study (continued)

Name	Sequence (5' to 3′)	Function
M13F	CGCCAGGGTTTTCCCAGTCACG	Probe generation for S1 nuclease mapping of *abeBCD* promoter; Sequencing of clones
3288-R	TGAGTGCGAGCCCCGGTGC	S1 nuclease mapping probe; RT-PCR for *abeABCD* read-through transcription
M13R	GCGGATAACAATTTCACACAGG	Sequencing of clones
OEA3287-1	GCGAGTCTTCGAACGGACCG	PCR amplification of *α-abeA* (*α-3287*) for overexpression
OEA3287-2	GATCGTCCCGCTGCTCGTCCC	PCR amplification of *α-abeA* (*α-3287*) for overexpression
OE3291-1	GTAGGCCTCCTGCCACCG	PCR amplification of *abeR* for overexpression
OE3291-2	CCGCCGACCTGGTACTCG	PCR amplification of *abeR* for overexpression
OE3291-3	GACTCCATATGGCGCGTATCCG	PCR amplification of *abeR* for protein overexpression in *Escherichia coli*
OE3291-4	CGGTAGCCTAGGACGTCACCG	PCR amplification of *abeR* for protein overexpression in *Escherichia coli*
CHPD-F	GCGCTGCTGTCGTCGCGGG	PCR amplification of non-specific competitor DNA for EMSA
CHPD-R	GGTTCCCGTGCACGTTCCC	PCR amplification of non-specific competitor DNA for EMSA
3287MSAF	GCTTGGGTGCCGGTGAACC	PCR amplification of probe for DNaseI footprinting
KO2792-1	GAAGGGCGCCGCGACCACCGAGGGGGGCTTAGCAGTATGATTCCG GGGATCCGTCGACC	Creation of *adpA* (*SCO2792*) knockout
KO2792-2	GGCCCGTCCGGCGTTCGAGGGAGCCGTCTGCTCACCTCATGTAGGC TGGAGCTGCTTC	Creation of *adpA* (*SCO2792*) knockout
2792-1O	AATGTCCCAAGTGCCGAACC	PCR analysis to confirm *adpA* knockout
2792-2IN	GCGAGATCGCCTCCAGTCC	PCR analysis to confirm *adpA* knockout
THIO-1	GACTGAGTTGGACACCATCG	RT-PCR for thiostrepton resistance gene (*tsr*) transcript
THIO-2	CCTTGAACAACTGGTTGACG	RT-PCR for thiostrepton resistance gene (*tsr*) transcript
RRNA-1	AGAGTTTGATCCTGGCTCAG	RT-PCR for 16S rRNA transcript
RRNA-2	CGAACCTCGCAGATGCCTG	RT-PCR for 16S rRNA transcript
5S	CCCTGCAGTACCATCGGCGCT	Probe for northern blot for 5S rRNA
3287 comp	TCGTCTTCGTCGTCATCGTC	Probe for northern blot for *α-abeA*
4677-1	CCCCGCTGGGACTGGCATCC	Probe for northern blot for *scr4677*
4675-2IN	CGCTCTGCGGTCTCGAAGC	PCR analysis to confirm *SCO4676* knockout, RT-PCR for *SCO4676-4675* read-through transcripts, PCR amplification of *SCO4676* for complementation
4677-3E	AACCCACGTTCTGAGGTTCC	PCR amplification of *scr4677* for overexpression, PCR analysis to confirm *SCO4676* knockout, PCR for amplifying *SCO4677-4676* read-through cDNA transcript, PCR amplification of *SCO4676* for complementation

Table 2.3. Oligonucleotides used in this study (continued)

Name	Sequence (5' to 3′)	Function
4676-6IN	GAGCGACAGCTCCTTGTGC	RT of *SCO4676* and *SCO4677-4676* read-through transcripts, PCR for amplifying *SCO4676* cDNA transcripts
4676-1IN	GGGTACCGAAAACACTAGCG	PCR for amplifying *SCO4676* cDNA transcripts
KO4676-1	CCCCCGCGCGCCCAGTCAGCCCCCGCCGACAATCCGACCATTCCGG GGATCCGTCGACC	Creation of *SCO4676* knockout
KO4676-2	CGTGGGACGAGCTGACCATGGAAGTCCTTGGTTCGGTTATGTAGGC TGGAGCTGCTTC	Creation of *SCO4676* knockout
4676 HOT	ACCAGGTGGTTGCCGATCAC	S1 nuclease mapping probe; PCR amplification of *scr4677* for overexpression ; PCR analysis to confirm *SCO4677* knockout ; PCR for amplifying *SCO4676* cDNA transcripts
46S177	ACATCCTCCGCATCGAGGTC	Probe generation for S1 nuclease mapping of *SCO4676* promoter
KO4677-1	TTCTCACGGTACGCACACATCGCCACGCTGAGTGACGTGATTCCGG GGATCCGTCGACC	Creation of *SCO4677* knockout
KO4677-2	TTCTTGTCTTGGTAGGGCGCCGACGGCACCGGAACCTCATGTAGGC TGGAGCTGCTTC	Creation of *SCO4677* knockout
4677-1IN	GTGAATCAGGAATCCACCACC	PCR analysis to confirm *SCO4677* knockout, RT-PCR for *SCO4677*
4677-2IN	TACAGGGTGAGCCGGAAATC	RT-PCR for *SCO4677*
MSA4677-5	ACCAGGTGGTTGCCGATCACCACGAAGCGGGTGGTGTGGCGGGTG TTCTCGTGGACGATGCCGCCGCCGTGGCGACGCGTACG	EMSA probe encompassing upstream of *scr4677*
MSA4677-6	CGTACGCGTCGCCACGGCGGCGGCATCGTCCACGAGAACACCCGC CACACCACCCGCTTCGTGGTGATCGGCAACCACCTGGT	EMSA probe encompassing upstream of *scr4677*
RPOBF	TCGACCACTTCGGCAACCGC	RT-PCR for *rpoB*
RPOBR	GCGCTCCATACGGGCGAGAC	RT-PCR for *rpoB*
4676-4IN	GCGCTAGTGTTTTCGGTACC	PCR for amplifying *SCO4677-4676* read-through cDNA transcript
4675-4IN	GAGTGAGGGTGGGCAATTCG	RT-PCR for *SCO4675*
4675-1IN	CGTGTGGACACACTGCTGC	RT-PCR for *SCO4675*
4676-3E	CGGCCTCCCACCCGTAACC	RT-PCR for *SCO4676-4675* read-through transcripts

Table 6.1. Sensitivity of deletion strains (Δ*SC04676* and Δ*SC04677*) compared to wild-type parent strain to various chemical stressors

Stressors	Zone of inhibition (mm)* Day 1			Zone of inhibition (mm)* Day 2		
	WT	Δ4676	Δ4677	WT	Δ4676	Δ4677
1% (v/v) SDS	16	14	16	12	11	10
50 mM EDTA	12	12	12	10	11	11
Lysozyme (30 mg/mL)	NI	NI	NI	7	7	7
Vancomycin (16 mg/mL)	14	14	14	14	14	14
0.05 M Diamide	23	23	20	13	13	13
0.1% (v/v) H_2O_2	22	20	20	Growth was no longer inhibited		
0.25% (v/v) H_2O_2	27	29	27	25	25	23
0.1 M Dipyridyl	NC	NC	NC	19	18	19

*) Zone of inhibition was measured based on the diameter of inihibition area observed after 1 or 2 days from inoculation, the values are the averages computed from three experiments.

NI = no zone of inhibition observed, NC = zone of inhibition was not clear.

Table 6.2. Sensitivity of *scr4677* overexpression strain compared to the plasmid control-carrying strain to various chemical stressors

Stressors	Zone of inhibition (mm)*	
	control	OEscr4677
1% (v/v) SDS	12	12
50 mM EDTA	10	10
Lysozyme (30 mg/mL)	7	7
Vancomycin (16 mg/mL)	14	14
0.05 M Diamide	18	17
0.25% (v/v) H_2O_2	24	23
0.1 M Dipyridyl	20	20

*) Zone of inhibition was measured based on the diameter of inihibition area observed after 2 days of inoculation, the values are the averages computed from three experiments. Both *scr4677* overexpression and control strains were inoculated on nutrient agar with thiostrepton. These strains grew slower than no plasmid-carrying strains (no thiostrepton supplemented to the growth medium); thus, zone of inhibition was not clear on day 1.

REFERENCES

Abramoff M.D., Magalhaes P.J., and Ram S.J. (2004) Image Processing with ImageJ. *Biophotonics Int* **11**: 36-42.

Aceti D.J., and Champness W.C. (1998) Transcriptional regulation of *Streptomyces coelicolor* pathway-specific antibiotic regulators by the *absA* and *absB* loci. *J Bacteriol* **180:** 3100-3106.

Adamidis T., and Champness W. (1992) Genetic analysis of *absB*, a *Streptomyces coelicolor* locus involved in global antibiotic regulation. *J Bacteriol* **174:** 4622-4628.

Altschul S.F., Madden T.L., Schaffer A.A., Zhang J., Zhang Z., Miller W., and Lipman D.J. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res* **25:** 3389-3402.

Anderson T.B., Brian P., and Champness W.C. (2001) Genetic and transcriptional analysis of *absA*, an antibiotic gene cluster-linked two-component system that regulates multiple antibiotics in *Streptomyces coelicolor*. *Mol Microbiol* **39:** 553-566.

André G., Even S., Putzer H., Burguière P., Croux C., Danchin A., Martin-Verstraete I., and Soutourina O. (2008) S-box and T-box riboswitches and antisense RNA control a sulfur metabolic operon of *Clostridium acetobutylicum*. *Nucleic Acids Res* **36:** 5955-5969.

Arias P., Fernandez-Moreno M.A., and Malpartida F. (1999) Characterization of the pathway-specific positive transcriptional regulator for actinorhodin biosynthesis in *Streptomyces coelicolor* A3(2) as a DNA-binding protein. *J Bacteriol* **181:** 6958-6968.

Babitzke P., Granger L., Olszewski J., and Kushner S.R. (1993) Analysis of mRNA decay and rRNA processing in *Escherichia coli* multiple mutants carrying a deletion in RNase III. *J Bacteriol* **175:** 229-239.

Bentley S.D., and Parkhill J. (2004) Comparative genomic structure of prokaryotes. *Annu Rev Genet* **38:** 771-792.

Bentley S.D., Chater K.F., Cerdeño-Tárraga A.M., Challis G.L., Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H., Harper D., Bateman A., Brown

S., Chandra G., Chen C. W., Collins M., Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S., Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S., Rabbinowitsch E., Rajandream M.-A., Rutherford K., Rutter S., Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K., Warren T., Wietzorrek A., Woodward J., Barrell B. G., Parkhill J., and Hopwood D.A. (2002) Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* **417:** 141-147.

Bibb M. (1996) 1995 Colworth prize lecture. the regulation of antibiotic production in *Streptomyces coelicolo*r A3(2). *Microbiol* **142:** 1335-1344.

Bibb M.J. (2005) Regulation of secondary metabolism in streptomycetes. *Curr Opin Microbiol* **8:** 208-215.

Bibb M.J., White J., Ward J.M., and Janssen G.R. (1994) The mRNA for the 23S rRNA methylase encoded by the *ermE* gene of *Saccharopolyspora erythraea* is translated in the absence of a conventional ribosome-binding site. *Mol Microbiol* **14:** 533-545.

Bradford M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72:** 248-254.

Brantl S. (2007) Regulatory mechanisms employed by cis-encoded antisense RNAs. *Curr Opin Microbiol* **10:** 102-109.

Carpousis A.J. (2007) The RNA degradosome of *Escherichia coli*: An mRNA-degrading machine assembled on RNase E. *Annu Rev Microbiol* **61:** 71-87.

Cases I., de Lorenzo V., and Ouzounis C.A. (2003) Transcription regulation and environmental adaptation in bacteria. *Trends Microbiol* **11:** 248-253.

Castro-Melchor M., Charaniya S., Karypis G., Takano E., and Hu W.S. (2010) Genome-wide inference of regulatory networks in *Streptomyces coelicolor. BMC Genomics* **11:** 578.

Celesnik H., Deana A., and Belasco J.G. (2007) Initiation of RNA decay in *Escherichia coli* by 5' pyrophosphate removal. *Mol Cell* **27:** 79-90.

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. *J Bacteriol* **179:** 5854-5861.

Challis G.L., and Hopwood D.A. (2003) Synergy and contingency as driving forces for the evolution of multiple secondary metabolite production by *Streptomyces* species. *Proc Natl Acad Sci U S A* **100:** 14555-14561.

Chang S.A., Bralley P., and Jones G.H. (2005) The *absB* gene encodes a double strand-specific endoribonuclease that cleaves the read-through transcript of the *rpsO-pnp* operon in *Streptomyces coelicolor*. *J Biol Chem* **280:** 33213-33219.

Charaniya S., Mehra S., Lian W., Jayapal K.P., Karypis G., and Hu W.S. (2007) Transcriptome dynamics-based operon prediction and verification in *Streptomyces coelicolor*. *Nucleic Acids Res* **35:** 7222-7236.

Chater K.F. (2006) *Streptomyces* inside-out: A new perspective on the bacteria that provide us with antibiotics. *Philos Trans R Soc Lond B Biol Sci* **361:** 761-768.

Chater K.F. (1998) Taking a genetic scalpel to the *Streptomyces* colony. *Microbiol* **144:** 1465-1478.

Chater K.F. (1993) Genetics of differentiation in *Streptomyces*. *Annu Rev Microbiol* **47:** 685-713.

Christen B., Abeliuk E., Collier J.M., Kalogeraki V.S., Passarelli B., Coller J.A., Fero M.J., McAdams H.H., and Shapiro L. (2011) The essential genome of a bacterium. *Mol Syst Biol* **7:** 528.

Corre C., Song L., O'Rourke S., Chater K.F., and Challis G.L. (2008) 2-alkyl-4-hydroxymethylfuran-3-carboxylic acids, antibiotic production inducers discovered by *Streptomyces coelicolor* genome mining. *Proc Natl Acad Sci U S A* **105:** 17510-17515.

Craney A., Hohenauer T., Xu Y., Navani N.K., Li Y., and Nodwell J. (2007) A synthetic *luxCDABE* gene cluster optimized for expression in high-GC bacteria. *Nucleic Acids Res* **35:** e46.

D'Alia D., Eggle D., Nieselt K., Hu W.S., Breitling R., and Takano E. (2011) Deletion of the signalling molecule synthase ScbA has pleiotropic effects on secondary metabolite biosynthesis, morphological differentiation and primary metabolism in *Streptomyces coelicolor* A3(2). *Microb Biotechnol* **4:** 239-251.

Dalton K.A., Thibessard A., Hunter J.I., and Kelemen G.H. (2007) A novel compartment, the 'subapical stem' of the aerial hyphae, is the location of a *sigN*-

dependent, developmentally distinct transcription in *Streptomyces coelicolor*. *Mol Microbiol* **64:** 719-737.

Darfeuille F., Unoson C., Vogel J., and Wagner E.G. (2007) An antisense RNA inhibits translation by competing with standby ribosomes. *Mol Cell* **26:** 381-392.

Datsenko K.A., and Wanner B.L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* **97:** 6640-6645.

Dayan F.E., Owens D.K., and Duke S.O. (2012) Rationale for a natural products approach to herbicide discovery. *Pest Manag Sci* **68:** 519-528 .

Deana A., and Belasco J.G. (2005) Lost in translation: The influence on bacterial mRNA decay. *Genes Dev* **19**: 2526-2533.

Deeble V.J., Lindley H.K., Fazeli M.R., Cove J.H., and Baumberg S. (1995) Depression of streptomycin production by *Streptomyces griseus* at elevated growth temperature: Studies using gene fusions. *Microbiol* **141:** 2511-2518.

den Hengst C.D., Tran N.T., Bibb M.J., Chandra G., Leskiw B.K., and Buttner M.J. (2010) Genes essential for morphological development and antibiotic production in *Streptomyces coelicolor* are targets of BldD during vegetative growth. *Mol Microbiol* **78:** 361-379.

Desnoyers G., Morissette A., Prévost K., and Massé E. (2009) Small RNA-induced differential degradation of the polycistronic mRNA *iscRSUA*. *EMBO J* **28:** 1551-1561.

Dornenburg J.E., Devita A.M., Palumbo M.J., and Wade J.T. (2010) Widespread antisense transcription in *Escherichia coli*. *mBio* **1:** e00024-10.

Dühring U., Axmann I.M., Hess W.R., and Wilde A. (2006) An internal antisense RNA regulates expression of the photosynthesis gene *isiA*. *Proc Natl Acad Sci U S A* **103:** 7054-7058.

Durand S., Gilet L., Bessieres P., Nicolas P., and Condon C. (2012) Three essential ribonucleases-RNase Y, J1, and III-control the abundance of a majority of *Bacillus subtilis* mRNAs. *PLoS Genet* **8:** e1002520.

Eccleston M., Willems A., Beveridge A., and Nodwell J.R. (2006) Critical residues and novel effects of overexpression of the *Streptomyces coelicolor*

developmental protein BldB: Evidence for a critical interacting partner. *J Bacteriol* **188:** 8189-8195.

Eccleston M., Ali R.A., Seyler R., Westpheling J., and Nodwell J. (2002) Structural and genetic analysis of the BldB protein of *Streptomyces coelicolor*. *J Bacteriol* **184:** 4270-4276.

El Hanafi D., and Bossi L. (2000) Activation and silencing of leu-500 promoter by transcription-induced DNA supercoiling in the *Salmonella chromosome*. *Mol Microbiol* **37:** 583-594.

Elliot M., Damji F., Passantino R., Chater K., and Leskiw B. (1998) The *bldD* gene of *Streptomyces coelicolor* A3(2): A regulatory gene involved in morphogenesis and antibiotic production. *J Bacteriol* **180:** 1549-1555.

Elliot M.A., Bibb M.J., Buttner M.J., and Leskiw B.K. (2001) BldD is a direct regulator of key developmental genes in *Streptomyces coelicolor* A3(2). *Mol Microbiol* **40:** 257-269.

Elliot M.A., Karoonuthaisiri N., Huang J., Bibb M.J., Cohen S.N., Kao C.M., and Buttner M.J. (2003) The chaplins: A family of hydrophobic cell-surface proteins involved in aerial mycelium formation in *Streptomyces coelicolor. Genes Dev* **17:** 1727-1740.

Even S., Pellegrini O., Zig L., Labas V., Vinh J., Brechemmier-Baey D., and Putzer H. (2005) Ribonucleases J1 and J2: Two novel endoribonucleases in *B.subtilis* with functional homology to *E.coli* RNase E. *Nucleic Acids Res* **33:** 2141-2152.

Fedoryshyn M., Petzke L., Welle E., Bechthold A., and Luzhetskyy A. (2008) Marker removal from actinomycetes genome using flp recombinase. *Gene* **419:** 43-47.

Flärdh K., and Buttner M.J. (2009) *Streptomyces* morphogenetics: Dissecting differentiation in a filamentous bacterium. *Nat Rev Microbiol* **7:** 36-49.

Floriano B., and Bibb M. (1996) *afsR* is a pleiotropic but conditionally required regulatory gene for antibiotic production in *Streptomyces coelicolor* A3(2). *Mol Microbiol* **21:** 385-396.

Gatewood M.L., Bralley P., Weil M.R., and Jones G.H. (2012) RNA-seq and RNA immunoprecipitation analyses of the transcriptome of *Streptomyces coelicolor* identify substrates for RNase III. *J Bacteriol* **194:** 2228-2237.

Georg J., and Hess W.R. (2011) Cis-antisense RNA, another level of gene regulation in bacteria. *Microbiol Mol Biol Rev* **75:** 286-300.

Gonzalez-Ceron G., Miranda-Olivares O.J., and Servin-Gonzalez L. (2009) Characterization of the methyl-specific restriction system of *Streptomyces coelicolor* A3(2) and of the role played by laterally acquired nucleases. *FEMS Microbiol Lett* **301:** 35-43.

Gottesman S., and Storz G. (2011) Bacterial small RNA regulators: Versatile roles and rapidly evolving variations. *Cold Spring Harb Perspect Biol* **3:** a003798.

Gravenbeek M.L., and Jones G.H. (2008) The endonuclease activity of RNase III is required for the regulation of antibiotic production by *Streptomyces coelicolor*. *Microbiol* **154:** 3547-3555.

Gregory M.A., Till R., and Smith M.C. (2003) Integration site for *Streptomyces* phage phiBT1 and development of site-specific integrating vectors. *J Bacteriol* **185:** 5320-5323.

Güell M., van Noort V., Yus E., Chen W.H., Leigh-Bell J., Michalodimitrakis K., Yamada T., Arumugam M., Doerks T., Kühner S., Rode M., Suyama M., Schmidt S., Gavin A-C., Bork P., Serrano L. (2009) Transcriptome complexity in a genome-reduced bacterium. *Science* **326:** 1268-1271.

Gust B., Challis G.L., Fowler K., Kieser T., and Chater K.F. (2003) PCR-targeted *Streptomyces* gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. *Proc Natl Acad Sci U S A* **100:** 1541-1546.

Haeder S., Wirth R., Herz H., and Spiteller D. (2009) Candicidin-producing *Streptomyces* support leaf-cutting ants to protect their fungus garden against the pathogenic fungus *Escovopsis*. *Proc Natl Acad Sci U S A* **106:** 4742-4746.

Hagège J.M., and Cohen S.N. (1997) A developmentally regulated *Streptomyces* endoribonuclease resembles ribonuclease E of *Escherichia coli*. *Mol Microbiol* **25:** 1077-1090.

Haiser H.J., Yousef M.R., and Elliot M.A. (2009) Cell wall hydrolases affect germination, vegetative growth, and sporulation in *Streptomyces coelicolor*. *J Bacteriol* **191:** 6501-6512.

Haiser H.J., Karginov F.V., Hannon G.J., and Elliot M.A. (2008) Developmentally regulated cleavage of tRNAs in the bacterium *Streptomyces coelicolor*. *Nucleic Acids Res* **36:** 732-741.

Han K., Kim K.S., Bak G., Park H., and Lee Y. (2010) Recognition and discrimination of target mRNAs by *sib* RNAs, a cis-encoded sRNA family. *Nucleic Acids Res* **38:** 5851-5866.

Hindra, Pak P., and Elliot M.A. (2010) Regulation of a novel gene cluster involved in secondary metabolite production in *Streptomyces coelicolor. J Bacteriol* **192:** 4973-4982.

Hjort K., Bergstrom M., Adesina M.F., Jansson J.K., Smalla K., and Sjoling S. (2010) Chitinase genes revealed and compared in bacterial isolates, DNA extracts and a metagenomic library from a phytopathogen-suppressive soil. *FEMS Microbiol Ecol* **71:** 197-207.

Hodges T.W., Slattery M., and Olson J.B. (2012) Unique actinomycetes from marine caves and coral reef sediments provide novel PKS and NRPS biosynthetic gene clusters. *Mar Biotechnol (NY)* **14:** 270-280.

Hopwood D.A., and Wright H.M. (1983) CDA is a new chromosomally-determined antibiotic from *Streptomyces coelicolor* A3(2). *J Gen Microbiol* **129:** 3575-3579.

Hsiao N.H., Soding J., Linke D., Lange C., Hertweck C., Wohlleben W., and Takano E. (2007) ScbA from *Streptomyces coelicolor* A3(2) has homology to fatty acid synthases and is able to synthesize gamma-butyrolactones. *Microbiol* **153:** 1394-1404.

Hsiao N.H., Nakayama S., Merlo M.E., de Vries M., Bunet R., Kitani S., Nihira T., and Takano E. (2009) Analysis of two additional signaling molecules in *Streptomyces coelicolor* and the development of a butyrolactone-specific reporter system. *Chem Biol* **16:** 951-960.

Huang J., Shi J., Molle V., Sohlberg B., Weaver D., Bibb M.J., Karoonuthaisiri N., Lih C-H., Kao C.M., Buttner, and Cohen S.N. (2005) Cross-regulation among disparate antibiotic biosynthetic pathways of *Streptomyces coelicolor. Mol Microbiol* **58:** 1276-1287.

Hutchings M.I., Hoskisson P.A., Chandra G., and Buttner M.J. (2004) Sensing and responding to diverse extracellular signals? analysis of the sensor kinases and response regulators of *Streptomyces coelicolor* A3(2). *Microbiol* **150:** 2795-2806.

Inagawa T., Okamoto S., Wachi M., and Ochi K. (2003) RNase ES of *Streptomyces coelicolor* A3(2) can complement the *rne* and *rng* mutations in *Escherichia coli. Biosci Biotechnol Biochem* **67:** 1767-1771.

Janssen G.R., and Bibb M.J. (1993) Derivatives of pUC18 that have BglII sites flanking a modified multiple cloning site and that retain the ability to identify recombinant clones by visual screening of *Escherichia coli* colonies. *Gene* **124:** 133-134.

Kaberdin V.R., and Bläsi U. (2006) Translation initiation and the fate of bacterial mRNAs. *FEMS Microbiol Rev* **30**: 967-979.

Kahsay R.Y., Gao G., and Liao L. (2005) An improved hidden markov model for transmembrane protein detection and topology prediction and its applications to complete genomes. *Bioinformatics* **21:** 1853-1858.

Kang S.G., Jin W., Bibb M., and Lee K.J. (1998) Actinorhodin and undecylprodigiosin production in wild-type and *relA* mutant strains of *Streptomyces coelicolor* A3(2) grown in continuous culture. *FEMS Microbiol Lett* **168:** 221-226.

Kato J.Y., Funa N., Watanabe H., Ohnishi Y., and Horinouchi S. (2007) Biosynthesis of gamma-butyrolactone autoregulators that switch on secondary metabolism and morphological development in *Streptomyces*. *Proc Natl Acad Sci U S A* **104:** 2378-2383.

Kawano M., Aravind L., and Storz G. (2007) An antisense RNA controls synthesis of an SOS-induced toxin evolved from an antitoxin. *Mol Microbiol* **64:** 738-754.

Kelemen G.H., Brian P., Flärdh K., Chamberlin L., Chater K.F., and Buttner M.J. (1998) Developmental regulation of transcription of *whiE*, a locus specifying the polyketide spore pigment in *Streptomyces coelicolor* A3(2). *J Bacteriol* **180:** 2515-2521.

Kelemen G.H., Brown G.L., Kormanec J., Potuckova L., Chater K.F., and Buttner M.J. (1996) The positions of the sigma-factor genes, *whiG* and *sigF*, in the hierarchy controlling the development of spore chains in the aerial hyphae of *Streptomyces coelicolor* A3(2). *Mol Microbiol* **21:** 593-603.

Kelley L.A., and Sternberg M.J. (2009) Protein structure prediction on the web: A case study using the phyre server. *Nat Protoc* **4:** 363-371.

Khan S.T., Komaki H., Motohashi K., Kozone I., Mukai A., Takagi M., and Shin-ya K. (2011) *Streptomyces* associated with a marine sponge *Haliclona* sp.; biosynthetic genes for secondary metabolites and products. *Environ Microbiol* **13:** 391-403.

Kieser T., Bibb M.J., Buttner M.J., Chater K.F., and Hopwood D.A. (2000) Practical *Streptomyces* genetics. The John Innes Foundation, Norwich, UK.

Kim D-W., Chater K., Lee K-J., and Hesketh A. (2005) Changes in the extracellular proteome caused by the absence of the *bldA* gene product, a developmentally significant tRNA, reveal a new target for the pleiotropic regulator AdpA in *Streptomyces coelicolor*. *J Bacteriol* **187:**2957-2966.

Kim E.S., Song J.Y., Kim D.W., Chater K.F., and Lee K.J. (2008) A possible extended family of regulators of sigma factor activity in *Streptomyces coelicolor*. *J Bacteriol* **190:** 7559-7566.

Kirby R., Wright L.F., and Hopwood D.A. (1975) Plasmid-determined antibiotic synthesis and resistance in *Streptomyces coelicolor*. *Nature* **254:** 265-267.

Kiviharju K., Leisola M., and Eerikainen T. (2004) Optimization of *Streptomyces peucetius* var. caesius N47 cultivation and epsilon-rhodomycinone production using experimental designs and response surface methods. *J Ind Microbiol Biotechnol* **31:** 475-481.

Komatsu M., Kuwahara Y., Hiroishi A., Hosono K., Beppu T., and Ueda K. (2003) Cloning of the conserved regulatory operon by its aerial mycelium-inducing activity in an *amfR* mutant of *Streptomyces griseus*. *Gene* **306:** 79-89.

Krogh A., Larsson B., von Heijne G., and Sonnhammer E.L. (2001) Predicting transmembrane protein topology with a hidden markov model: Application to complete genomes. *J Mol Biol* **305:** 567-580.

Labes G., Bibb M., and Wohlleben W. (1997) Isolation and characterization of a strong promoter element from the *Streptomyces ghanaensis* phage I19 using the gentamicin resistance gene (*aacC1*) of tn 1696 as reporter. *Microbiol* **143:** 1503-1512.

Laing E., Mersinias V., Smith C.P., and Hubbard S.J. (2006) Analysis of gene expression in operons of *Streptomyces coelicolor*. *Genome Biol* **7:** R46.

Lakey J.H., Lea E.J., Rudd B.A., Wright H.M., and Hopwood D.A. (1983) A new channel-forming antibiotic from *Streptomyces coelicolor* A3(2) which requires calcium for its activity. *J Gen Microbiol* **129:** 3565-3573.

Lawlor E.J., Baylis H.A., and Chater K.F. (1987) Pleiotropic morphological and antibiotic deficiencies result from mutations in a gene encoding a tRNA-like product in *Streptomyces coelicolor* A3(2). *Genes Dev* **1:** 1305-1310.

Le T.B., Fiedler H.P., den Hengst C.D., Ahn S.K., Maxwell A., and Buttner M.J. (2009) Coupling of the biosynthesis and export of the DNA gyrase inhibitor simocyclinone in *Streptomyces antibioticus*. *Mol Microbiol* **72:** 1462-1474.

Lee E.J., and Groisman E.A. (2010) An antisense RNA that governs the expression kinetics of a multifunctional virulence gene. *Mol Microbiol* **76:** 1020-1033.

Lee K., and Cohen S.N. (2003) A *Streptomyces coelicolor* functional orthologue of *Escherichia coli* RNase E shows shuffling of catalytic and PNPase-binding domains. *Mol Microbiol* **48:** 349-360.

Lee P.C., Umeyama T., and Horinouchi S. (2002) *afsS* is a target of AfsR, a transcriptional factor with ATPase activity that globally controls secondary metabolism in *Streptomyces coelicolor* A3(2). *Mol Microbiol* **43:** 1413-1430.

MacNeil D.J., Gewain K.M., Ruby C.L., Dezeny G., Gibbons P.H., and MacNeil T. (1992) Analysis of *Streptomyces avermitilis* genes required for avermectin biosynthesis utilizing a novel integration vector. *Gene* **111:** 61-68.

Majdalani N., Hernandez D., and Gottesman S. (2002) Regulation and mode of action of the second small RNA activator of RpoS translation, RprA. *Mol Microbiol* **46:** 813-826.

Majdalani N., Cunning C., Sledjeski D., Elliott T., and Gottesman S. (1998) DsrA RNA regulates translation of RpoS message by an anti-antisense mechanism, independent of its action as an antisilencer of transcription. *Proc Natl Acad Sci U S A* **95:** 12462-12467.

Mandin P., and Gottesman S. (2010) Integrating anaerobic/aerobic sensing and the general stress response through the ArcZ small RNA. *EMBO J* **29:** 3094-3107.

Manzoni M., and Rollini M. (2002) Biosynthesis and biotechnological production of statins by filamentous fungi and application of these cholesterol-lowering drugs. *Appl Microbiol Biotechnol* **58:** 555-564.

Marchler-Bauer A., Lu S., Anderson J.B., Chitsaz F., Derbyshire M.K., DeWeese-Scott C., Fong J.H., Geer LY., Gonzales N.R., Gwadz M., Hurwitz D.I., Jackson J.D., Ke Z., Lanczycki C.J., Lu F., Marchler G.H., Mullokandov M., Omelchenko M.V., Robertson C.L., Song J.S., Thanki N., Yamashita R.A., Zhang D., Zhang N., Zheng C., and Bryant S.H. (2011) CDD: A conserved domain database for the functional annotation of proteins. *Nucleic Acids Res* **39:** D225-9.

Martin J.F. (2004) Phosphate control of the biosynthesis of antibiotics and other secondary metabolites is mediated by the PhoR-PhoP system: An unfinished story. *J Bacteriol* **186:** 5197-5201.

Martin J.F., and Liras P. (2010) Engineering of regulatory cascades and networks controlling antibiotic biosynthesis in *Streptomyces*. *Curr Opin Microbiol* **13:** 263-273.

Martin M.U., and Wesche H. (2002) Summary and comparison of the signaling mechanisms of the Toll/interleukin-1 receptor family. *Biochim Biophys Acta* **1592:** 265-280.

Martinez-Hackert E., and Stock A.M. (1997) The DNA-binding domain of OmpR: Crystal structures of a winged helix transcription factor. *Structure* **5:** 109-124.

Massé E., Escorcia F.E., and Gottesman S. (2003) Coupled degradation of a small regulatory RNA and its mRNA targets in *Escherichia coli*. *Genes Dev* **17:** 2374-2383.

Mathy N., Benard L., Pellegrini O., Daou R., Wen T., and Condon C. (2007) 5'-to-3' exoribonuclease activity in bacteria: Role of RNase J1 in rRNA maturation and 5' stability of mRNA. *Cell* **129:** 681-692.

McKenzie N.L., and Nodwell J.R. (2007) Phosphorylated AbsA2 negatively regulates antibiotic production in *Streptomyces coelicolor* through interactions with pathway-specific regulatory gene promoters. *J Bacteriol* **189:** 5284-5292.

Mitschke J., Georg J., Scholz I., Sharma C.M., Dienst D., Bantscheff J., Voβ B., Steglich C., Wilde A., Vogel J., and Hess W.R. (2011) An experimentally anchored map of transcriptional start sites in the model *Cyanobacterium synechocystis* sp. PCC6803. *Proc Natl Acad Sci U S A* **108:** 2124-2129.

Mokni-Tlili S., Ben Abdelmalek I., Jedidi N., Belghith H., Gargouri A., Abdennaceur H., and Marzouki M.N. (2010) Exploitation of biological wastes for

the production of value-added hydrolases by *Streptomyces* sp. MSWC1 isolated from municipal solid waste compost. *Waste Manag Res* **28:** 828-837.

Morita T., Mochizuki Y., and Aiba H. (2006) Translational repression is sufficient for gene silencing by bacterial small noncoding RNAs in the absence of mRNA destruction. *Proc Natl Acad Sci U S A* **103:** 4858-4863.

Moulin L., Rahmouni A.R., and Boccard F. (2005) Topological insulators inhibit diffusion of transcription-induced positive supercoils in the chromosome of *Escherichia coli. Mol Microbiol* **55:** 601-610.

Mraheil M.A., Billion A., Mohamed W., Mukherjee K., Kuenne C., Pischimarov J., Krawitz C., Retey J., Hartsch T., Chakraborty T., and Hain T. (2011) The intracellular sRNA transcriptome of *Listeria monocytogenes* during growth in macrophages. *Nucleic Acids Res* **39:** 4235-4248.

Narva K.E., and Feitelson J.S. (1990) Nucleotide sequence and transcriptional analysis of the *redD* locus of *Streptomyces coelicolor* A3(2). *J Bacteriol* **172:** 326-333.

Nett M., Ikeda H., and Moore B.S. (2009) Genomic basis for natural product biosynthetic diversity in the actinomycetes. *Nat Prod Rep* **26:** 1362-1384.

Nguyen K.T., Tenor J., Stettler H., Nguyen L.T., Nguyen L.D., and Thompson C.J. (2003) Colonial differentiation in *Streptomyces coelicolor* depends on translation of a specific codon within the *adpA* gene. *J Bacteriol* **185:** 7291-7296.

Nicholson A. W. (2011) Ribonuclease III and the role of double-stranded RNA processing in bacterial systems. In: Ribonucleases, Nicholson, A.W., Ed. Springer Berlin, Germany pp. 269-298.

Nicolas P., Mader U., Dervyn E., Rochat T., Leduc A., Pigeonneau N., Bidnenko E., Marchadier E., Hoebeke M., Aymerich S., Becher D., Bisichhia P., Botella E., Delumeau O., Doherty G., Denham E.L., Fogg M.J., Fromion V., Goelzer A., Hansen A., Härtig E., Harwood C.R., Homuth G., Jarmer H., Jules M., Klipp E., Le Chat L., Lecointe F., Lewis P., Liebermeister W., March A., Mars R.A.T., Nannapaneni P., Noone D., Pohl S., Rinn B., Rügheimer F., Sappa P.K., Samson F., Schaffer M., Schwikowski B., Steil L., Stülke J., Wiegert T., Devine K.M., Wilkinson A.J., van Dijl J.M., Hecker M., Völker U., Bessières P., and Noirot P. (2012) Condition-dependent transcriptome reveals high-level regulatory architecture in *Bacillus subtilis. Science* **335:** 1103-1106.

Nothaft H., Rigali S., Boomsma B., Swiatek M., McDowall K.J., van Wezel G.P., and Titgemeyer F. (2010) The permease gene *nagE2* is the key to *N*-acetylglucosamine sensing and utilization in *Streptomyces coelicolor* and is subject to multi-level control. *Mol Microbiol* **75:** 1133-1144.

Ochi K. (2007) From microbial differentiation to ribosome engineering. *Biosci Biotechnol Biochem* **71:** 1373-1386.

Ohnishi Y., Kameyama S., Onaka H., and Horinouchi S. (1999) The A-factor regulatory cascade leading to streptomycin biosynthesis in *Streptomyces griseus* : Identification of a target gene of the A-factor receptor. *Mol Microbiol* **34:** 102-111.

Ohnishi Y., Yamazaki H., Kato J.Y., Tomono A., and Horinouchi S. (2005) AdpA, a central transcriptional regulator in the A-factor regulatory cascade that leads to morphological development and secondary metabolism in *Streptomyces griseus*. *Biosci Biotechnol Biochem* **69:** 431-439.

Okamoto S., Lezhava A., Hosaka T., Okamoto-Hosoya Y., and Ochi K. (2003) Enhanced expression of S-adenosylmethionine synthetase causes overproduction of actinorhodin in *Streptomyces coelicolor* A3(2). *J Bacteriol* **185:** 601-609.

O'Neill L.A., and Bowie A.G. (2007) The family of five: TIR-domain-containing adaptors in toll-like receptor signalling. *Nat Rev Immunol* **7:** 353-364.

Opdyke J.A., Kang J.G., and Storz G. (2004) GadY, a small-RNA regulator of acid response genes in *Escherichia coli*. *J Bacteriol* **186:** 6698-6705.

Opdyke J.A., Fozo E.M., Hemm M.R., and Storz G. (2011) RNase III participates in GadY-dependent cleavage of the *gadX-gadW* mRNA. *J Mol Biol* **406:** 29-43.

O'Rourke S., Wietzorrek A., Fowler K., Corre C., Challis G.L., and Chater K.F. (2009) Extracellular signalling, translational control, two repressors and an activator all contribute to the regulation of methylenomycin production in *Streptomyces coelicolor. Mol Microbiol* **71:** 763-778.

Pall G.S., and Hamilton A.J. (2008) Improved northern blot method for enhanced detection of small RNA. *Nat Protoc* **3:** 1077-1084.

Panek J., Bobek J., Mikulik K., Basler M., and Vohradsky J. (2008) Biocomputational prediction of small non-coding RNAs in *Streptomyces*. *BMC Genomics* **9:** 217.

Perocchi F., Xu Z., Clauder-Munster S., and Steinmetz L.M. (2007) Antisense artifacts in transcriptome microarray experiments are resolved by actinomycin D. *Nucleic Acids Res* **35:** e128.

Pfaffl M.W. (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* **29:** e45.

Pope M.K., Green B., and Westpheling J. (1998) The *bldB* gene encodes a small protein required for morphogenesis, antibiotic production, and catabolite control in *Streptomyces coelicolor*. *J Bacteriol* **180:** 1556-1562.

Prévost K., Desnoyers G., Jacques J.F., Lavoie F., and Massé E. (2011) Small RNA-induced mRNA degradation achieved through both translation block and activated cleavage. *Genes Dev* **25:** 385-396.

Price B., Adamidis T., Kong R., and Champness W. (1999) A *Streptomyces coelicolor* antibiotic regulatory gene, *absB*, encodes an RNase III homolog. *J Bacteriol* **181:** 6142-6151.

Redenbach M., Kieser H.M., Denapaite D., Eichner A., Cullum J., Kinashi H., and Hopwood D.A. (1996) A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb *Streptomyces coelicolor* A3(2) chromosome. *Mol Microbiol* **21:** 77-96.

Resch A., Afonyushkin T., Lombo T.B., McDowall K.J., Blasi U., and Kaberdin V.R. (2008) Translational activation by the noncoding RNA DsrA involves alternative RNase III processing in the *rpoS* 5'-leader. *RNA* **14:** 454-459.

Rigali S., Titgemeyer F., Barends S., Mulder S., Thomae A.W., Hopwood D.A., and van Wezel G.P. (2008) Feast or famine: The global regulator DasR links nutrient stress to antibiotic production by *Streptomyces*. *EMBO Rep* **9:** 670-675.

Rigali S., Nothaft H., Noens E.E., Schlicht M., Colson S., Müller M., Joris B., Koerten H.K., Hopwood D.A., Titgemeyer F., and van Wezel G.P. (2006) The sugar phosphotransferase system of *Streptomyces coelicolor* is regulated by the GntR-family regulator DasR and links N-acetylglucosamine metabolism to the control of development. *Mol Microbiol* **61:** 1237-1251.

Rudd B.A., and Hopwood D.A. (1980) A pigmented mycelial antibiotic in *Streptomyces coelicolor*: Control by a chromosomal gene cluster. *J Gen Microbiol* **119:** 333-340.

Ruiz B., Chávez A., Forero A., García-Huante Y., Romero A., Sánchez M., Rocha D., Sánchez B., Rodriguez-Sanoja R., Sánchez S., Langley E. (2010) Production of microbial secondary metabolites: Regulation by the carbon source. *Crit Rev Microbiol* **36:** 146-167.

Ryding N.J., Anderson T.B., and Champness W.C. (2002) Regulation of the *Streptomyces coelicolor* calcium-dependent antibiotic by *absA*, encoding a cluster-linked two-component system. *J Bacteriol* **184:** 794-805.

Sambrook J., and Russell D.W. (2001) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Santamarta I., Rodriguez-Garcia A., Perez-Redondo R., Martin J.F., and Liras P. (2002) CcaR is an autoregulatory protein that binds to the *ccaR* and *cefD-cmcI* promoters of the cephamycin C-clavulanic acid cluster in *Streptomyces clavuligerus*. *J Bacteriol* **184:** 3106-3113.

Santos-Beneit F., Rodriguez-Garcia A., Sola-Landa A., and Martin J.F. (2009) Cross-talk between two global regulators in *Streptomyces*: PhoP and AfsR interact in the control of *afsS, pstS* and *phoRP* transcription. *Mol Microbiol* **72:** 53-68.

Saudagar P.S., Survase S.A., and Singhal R.S. (2008) Clavulanic acid: A review. *Biotechnol Adv* **26:** 335-351.

Seipke R.F., Kaltenpoth M., and Hutchings M.I. (2011) *Streptomyces* as symbionts: An emerging and widespread theme? *FEMS Microbiol Rev* .

Selinger D.W., Cheung K.J., Mei R., Johansson E.M., Richmond C.S., Blattner F.R., Lockhart D.J., and Church G.M. (2000) RNA expression analysis using a 30 base pair resolution *Escherichia coli* genome array. *Nat Biotechnol* **18:** 1262-1268.

Sello J.K., and Buttner M.J. (2008) The gene encoding RNase III in *Streptomyces coelicolor* is transcribed during exponential phase and is required for antibiotic production and for proper sporulation. *J Bacteriol* **190:** 4079-4083.

Shahbabian K., Jamalli A., Zig L., and Putzer H. (2009) RNase Y, a novel endoribonuclease, initiates riboswitch turnover in *Bacillus subtilis*. *EMBO J* **28:** 3523-3533.

Sharma C.M., Hoffmann S., Darfeuille F., Reignier J., Findeiss S., Sittka A., Chabas S., Reiche K., Hackermüller J., Reinhardt R., Stadler P.F., and Vogel, J.

(2010) The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature* **464:** 250-255.

Sheeler N.L., MacMillan S.V., and Nodwell J.R. (2005) Biochemical activities of the *absA* two-component system of *Streptomyces coelicolor*. *J Bacteriol* **187:** 687-696.

Sheldon P.J., Busarow S.B., and Hutchinson C.R. (2002) Mapping the DNA-binding domain and target sequences of the *Streptomyces peucetius* daunorubicin biosynthesis regulatory protein, DnrI. *Mol Microbiol* **44:** 449-460.

Sledjeski D.D., Gupta A., and Gottesman S. (1996) The small RNA, DsrA, is essential for the low temperature expression of RpoS during exponential growth in *Escherichia coli*. *EMBO J* **15:** 3993-4000.

Soding J., Biegert A., and Lupas A.N. (2005) The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res* **33:** W244-8.

Sola-Landa A., Moura R.S., and Martin J.F. (2003) The two-component PhoR-PhoP system controls both primary metabolism and secondary metabolite biosynthesis in *Streptomyces lividans*. *Proc Natl Acad Sci U S A* **100:** 6133-6138.

Spear A.M., Loman N.J., Atkins H.S., and Pallen M.J. (2009) Microbial TIR domains: Not necessarily agents of subversion? *Trends Microbiol* **17:** 393-398.

Stegmann E., Frasch H.J., and Wohlleben W. (2010) Glycopeptide biosynthesis in the context of basic cellular functions. *Curr Opin Microbiol* **13:** 595-602.

Stork M., Di Lorenzo M., Welch T.J., and Crosa J.H. (2007) Transcription termination within the iron transport-biosynthesis operon of *Vibrio anguillarum* requires an antisense RNA. *J Bacteriol* **189:** 3479-3488.

Storz G., Vogel J., and Wassarman K.M. (2011) Regulation by small RNAs in bacteria: Expanding frontiers. *Mol Cell* **43:** 880-891.

Sun X., Zhulin I., and Wartell R.M. (2002) Predicted structure and phyletic distribution of the RNA-binding protein Hfq. *Nucleic Acids Res* **30:** 3662-3671.

Swiercz J., Hindra, Bobek J., Haiser H.J., Di Berardo C.,Tjaden B., and Elliot, M.A. (2008) Small non-coding RNAs in *Streptomyces coelicolor*. *Nucleic Acids Res* **36:** 7240-7251.

Tahlan K., Ahn S.K., Sing A., Bodnaruk T.D., Willems A.R., Davidson A.R., and Nodwell J.R. (2007) Initiation of actinorhodin export in *Streptomyces coelicolor*. *Mol Microbiol* **63:** 951-961.

Takano E. (2006) Gamma-butyrolactones: *Streptomyces* signalling molecules regulating antibiotic production and differentiation. *Curr Opin Microbiol* **9:** 287-294.

Takano E., Chakraburtty R., Nihira T., Yamada Y., and Bibb M.J. (2001) A complex role for the gamma-butyrolactone SCB1 in regulating antibiotic production in *Streptomyces coelicolor* A3(2). *Mol Microbiol* **41:** 1015-1028.

Takano E., Kinoshita H., Mersinias V., Bucca G., Hotchkiss G., Nihira T., Smith C.P., Bibb M., Wohlleben W., and Chater K. (2005) A bacterial hormone (the SCB1) directly controls the expression of a pathway-specific regulatory gene in the cryptic type I polyketide biosynthetic gene cluster of *Streptomyces coelicolor*. *Mol Microbiol* **56:** 465-479.

Takano E., Tao M., Long F., Bibb M.J., Wang L., Li W., Li W., Buttner M.J., Bibb M.J., Deng Z.X., and Chater K.F. (2003) A rare leucine codon in *adpA* is implicated in the morphological defect of *bldA* mutants of *Streptomyces coelicolor*. *Mol Microbiol* **50:** 475-486.

Takano H., Hashimoto K., Yamamoto Y., Beppu T., and Ueda K. (2011) Pleiotropic effect of a null mutation in the *cvn1* conservon of *Streptomyces coelicolor* A3(2). *Gene* **477:** 12-18.

Tanaka A., Takano Y., Ohnishi Y., and Horinouchi S. (2007) AfsR recruits RNA polymerase to the *afsS* promoter: A model for transcriptional activation by SARPs. *J Mol Biol* **369:** 322-333.

Taylor S., Wakem M., Dijkman G., Alsarraj M., and Nguyen M. (2010) A practical approach to RT-qPCR—Publishing data that conform to the MIQE guidelines. *Methods* **50:** S1-S5.

Tezuka T., Hara H., Ohnishi Y., and Horinouchi S. (2009) Identification and gene disruption of small noncoding RNAs in *Streptomyces griseus*. *J Bacteriol* **191:** 4896-4904.

Thomason M.K., and Storz G. (2010) Bacterial antisense RNAs: How many are there, and what are they doing? *Annu Rev Genet* **44:** 167-188.

Tian X.P., Long L.J., Wang F.Z., Xu Y., Li J., Zhang J., Zhang C-S., Zhang S., and Li W-J. (2012) *Streptomyces nanhaiensis* sp. nov., a marine streptomycete isolated from a deep-sea sediment. *Int J Syst Evol Microbiol* **62:** 864-868.

Toledo-Arana A., Dussurget O., Nikitas G., Sesto N., Guet-Revillet H., Balestrino D., Loh E., Gripenland J., Tiensuu T., Vaitkevicius K., Nahori M-A., Soubigou G., Régnault B., Coppée J-Y., Lecuit M., Johansson J., and Cossart P. (2009) The *Listeria* transcriptional landscape from saprophytism to virulence. *Nature* **459:** 950-956.

Tomono A., Tsai Y., Yamazaki H., Ohnishi Y., and Horinouchi S. (2005) Transcriptional control by A-factor of *strR*, the pathway-specific transcriptional activator for streptomycin biosynthesis in *Streptomyces griseus*. *J Bacteriol* **187:** 5595-5604.

Uguru G.C., Stephens K.E., Stead J.A., Towle J.E., Baumberg S., and McDowall K.J. (2005) Transcriptional activation of the pathway-specific regulator of the actinorhodin biosynthetic genes in *Streptomyces coelicolor*. *Mol Microbiol* **58:** 131-150.

van Wezel G.P., Konig M., Mahr K., Nothaft H., Thomae A.W., Bibb M., and Titgemeyer F. (2007) A new piece of an old jigsaw: Glucose kinase is activated posttranslationally in a glucose transport-dependent manner in *Streptomyces coelicolor* A3(2). *J Mol Microbiol Biotechnol* **12:** 67-74.

Vara J., Lewandowska-Skarbek M., Wang Y.G., Donadio S., and Hutchinson C.R. (1989) Cloning of genes governing the deoxysugar portion of the erythromycin biosynthesis pathway in *Saccharopolyspora erythraea (Streptomyces erythreus*). *J Bacteriol* **171:** 5872-5881.

Vockenhuber M.P., and Suess B. (2012) *Streptomyces coelicolor* sRNA scr5239 inhibits agarase expression by direct base pairing to the *dagA* coding region. *Microbiol* **158:** 424-435.

Vockenhuber M.P., Sharma C.M., Statt M.G., Schmidt D., Xu Z., Dietrich S., Liesegang H., Mathews D.H., and Suess B. (2011) Deep sequencing-based identification of small non-coding RNAs in *Streptomyces coelicolor*. *RNA Biol* **8:** 468-477.

Voelker F., and Altaba S. (2001) Nitrogen source governs the patterns of growth and pristinamycin production in '*Streptomyces pristinaespiralis*'. *Microbiol* **147:** 2447-2459.

Vogel J., Bartels V., Tang T.H., Churakov G., Slagter-Jager J.G., Huttenhofer A., and Wagner E.G. (2003) RNomics in *Escherichia coli* detects new sRNA species and indicates parallel transcriptional output in bacteria. *Nucleic Acids Res* **31:** 6435-6443.

Vögtli M., Chang P.C., and Cohen S.N. (1994) afsR2: A previously undetected gene encoding a 63-amino-acid protein that stimulates antibiotic production in *Streptomyces lividans*. *Mol Microbiol* **14:** 643-653.

Wadler C.S., and Vanderpool C.K. (2007) A dual function for a bacterial small RNA: SgrS performs base pairing-dependent regulation and encodes a functional polypeptide. *Proc Natl Acad Sci U S A* **104:** 20454-20459.

Wang Y., Zhu W., and Levy D.E. (2006) Nuclear and cytoplasmic mRNA quantification by SYBR green based real-time RT-PCR. *Methods* **39:** 356-362.

Waters L.S., and Storz G. (2009) Regulatory RNAs in bacteria. *Cell* **136:** 615-628.

Wietzorrek A., and Bibb M. (1997) A novel family of proteins that regulates antibiotic production in streptomycetes appears to contain an OmpR-like DNA-binding fold. *Mol Microbiol* **25:** 1181-1184.

Williamson N.R., Fineran P.C., Leeper F.J., and Salmond G.P. (2006) The biosynthesis and regulation of bacterial prodiginines. *Nat Rev Microbiol* **4:** 887-899.

Wright L.F., and Hopwood D.A. (1976) Actinorhodin is a chromosomally-determined antibiotic in *Streptomyces coelicolor* A3(2). *J Gen Microbiol* **96:** 289-297.

Wright L.F., and Hopwood D.A. (1976) Identification of the antibiotic determined by the SCP1 plasmid of *Streptomyces coelicolor* A3(2). *J Gen Microbiol* **95:** 96-106.

Xu J., Wang Y., Xie S.J., Xu J., Xiao J., and Ruan J.S. (2009) *Streptomyces xiamenensis* sp. nov., isolated from mangrove sediment. *Int J Syst Evol Microbiol* **59:** 472-476.

Xu W., Huang J., and Cohen S.N. (2008) Autoregulation of AbsB (RNase III) expression in *Streptomyces coelicolor* by endoribonucleolytic cleavage of *absB* operon transcripts. *J Bacteriol* **190:** 5526-5530.

Xu W., Huang J., Lin R., Shi J., and Cohen S.N. (2010) Regulation of morphological differentiation in *S. coelicolor* by RNase III (AbsB) cleavage of mRNA encoding the AdpA transcription factor. *Mol Microbiol* **75:** 781-791.

Yamazaki H., Tomono A., Ohnishi Y., and Horinouchi S. (2004) DNA-binding specificity of AdpA, a transcriptional activator in the A-factor regulatory cascade in *Streptomyces griseus*. *Mol Microbiol* **53:** 555-572.

Yeats C., Bentley S., and Bateman A. (2003) New knowledge from old: In silico discovery of novel protein domains in *Streptomyces coelicolor*. *BMC Microbiol* **3:** 3.

Zemanova M., Kaderabkova P., Patek M., Knoppova M., Silar R., and Nesvera J. (2008) Chromosomally encoded small antisense RNA in *Corynebacterium glutamicum. FEMS Microbiol Lett* **279:** 195-201.

Zhang A., Altuvia S., Tiwari A., Argaman L., Hengge-Aronis R., and Storz G. (1998) The OxyS regulatory RNA represses *rpoS* translation and binds the Hfq (HF-I) protein. *EMBO J* **17:** 6061-6068.