

EXPLORING THE REGULATION OF SECONDARY
METABOLITES IN *STREPTOMYCES*

EXPLORING THE REGULATION OF SECONDARY
METABOLITES IN *STREPTOMYCES*

Christopher Donald Firby

A Thesis
Submitted to the School of Graduate Studies
In Partial Fulfillment of the Requirements of the Degree of
Master of Science

McMaster University
July, 2017

Master of Science (2017)
(Biology)

McMaster University
Hamilton, Ontario

TITLE: Exploring the regulation of secondary metabolites in *Streptomyces*

AUTHOR: Chris Firby, B.Sc. (McMaster University)

SUPERVISOR: Dr. Marie A. Elliot

NUMBER OF PAGES:

ABSTRACT:

The genus *Streptomyces* harbors a wealth of untapped pharmaceutical potential within uncharacterized biosynthetic clusters whose transcription may be silent or their product undetected. Regulation of these clusters is, in part, managed by global transcription factors, nucleoid-associated proteins, and pathway-specific regulatory proteins. Consequently, a complicated interplay of regulatory elements governs secondary metabolite production and a lack of understanding of this has undermined efforts to screen for novel drug candidates. In an effort to better understand this regulation, we have focussed on one highly expressed NAP, and one silent and cryptic biosynthetic cluster. In parallel, we have leveraged our existing knowledge to improve our screening of *Streptomyces* for novel bioactive compounds.

Here, we investigate the nucleoid-associated protein *Streptomyces* integration host factor (sIHF), which non-specifically binds DNA and impacts development and antibiotic production. Through targeted mutations, we probe the interfaces of sIHF thought responsible for its interactions with DNA and assess the relevance of these interfaces to *in vivo* activity. Further, we examine the transcriptionally silent *SCO6429-SCO6438* biosynthetic cluster and attempt to activate transcription from its promoter by manipulating its potential regulatory proteins. Binding sites for the putative global regulator cAMP-receptor protein (Crp) in close proximity to this cluster and a potential pathway specific regulator are examined. Finally, we overexpress Crp in a library of *Streptomyces* and assay for novel bioactivity against known clinical pathogens.

Collectively, this work establishes a foundation for further exploring the regulatory networks within *Streptomyces* and how they may be influenced to discover novel pharmaceuticals.

ACKNOWLEDGEMENTS:

This work would not have been possible without Dr. Marie Elliot, whose experience, guidance, and support has dramatically improved my scientific literacy, refined my approach to project management, and instilled in me a sense of optimism and motive to succeed. I have the highest admiration for her intelligence, compassion, and dedication to science and have no doubts that her research will continue to impress and inspire generations of *Streptomyces* researchers.

I would like to thank my committee members, Dr. Turlough Finan and Dr. Alba Guarné, whose advice and encouragement have been instrumental in completing this work. It was during my undergraduate work at McMaster that Dr. Elliot and Dr. Guarné instilled in me a passion for science and the drive to join the Masters program. I came to know Dr. Finan early in my graduate work, and his charming approach to discussing science has allowed me to expand my knowledge of microbial systems. I thank all of you for your continued support.

My colleagues in the Elliot Lab are some of the most wonderful people I have ever worked with. Matt Moody has been a source of humour who brought levity to the lab during tough times. He has a refined sense of the scientific method and brings probing questions to every conversation, and I have valued every insight. Danielle Sexton and Dr. Renee St-Onge have been role models for developing good lab habits and establishing effective protocols. Their generous support and advice over the years has been greatly appreciated. Dr. Emma Sherwood and Dr. Julia Swiercz provided the backbone for this research and I thank them for their guidance and inspiring talks. Rachel Young and Andrew Duong are extremely kind-hearted and assisted throughout my stay in the Elliot Lab. Andrew, I'll never forget the inspiring words you gave me. Stephanie Jones, thank you for your support and constant supply of intriguing science discussions. I look forward to future papers on *S. venezuelae*'s spreading/smothering phenotype. Savannah and David have brought a wealth of fresh perspectives to the lab and helped reinvigorate my passion for science.

I thank my family for their support and housing over the last few years, and my mother and father in particular for encouraging me to finish while struggling with health issues. To my good friend, Nabil Sakran, I thank you from the bottom of my heart for all your help and support.

A special thank you goes out to the Ontario and Canadian medical systems that have treated both my father and me over these past few years. In particular, I would like to thank Juravinski Hospital for everything they've done and continue to do for my father and the many others across this province affected by cancer.

TABLE OF CONTENTS

Abstract	iv
Acknowledgements	v
Table of Contents	vi
List of Figures	xii
List of Tables	xiii
List of Abbreviations	xiv
Chapter 1: Introduction	
1.1 <i>Streptomyces</i> as a bacterium of unique development and historic antimicrobial importance	1
1.1.1 Model species for study within <i>Streptomyces</i>	2
1.1.2 Potential of <i>Streptomyces</i> for future bioactive molecule discoveries	3
1.2 Transcription regulation in <i>Streptomyces</i> is complex	4
1.2.1 Transcriptional control by nucleoid-associated proteins	5
1.2.2 Properties of nucleoid-associated proteins	6
1.2.2.1 <i>Histone-like nucleoid structuring protein</i>	6
1.2.2.2 <i>Integration host factors</i>	7
1.3 Aims of this work	8
Chapter 2: Materials and Methods	10
2.1 Bacterial culturing, transformation and selection	10
2.1.1 Preparation of cultures & storage of strain stocks	10

2.1.2 Preparing chemically competent & electrocompetent <i>E. coli</i> cells	10
2.1.3 Transformation of <i>E. coli</i> cells for molecular cloning	10
2.1.4 Harvesting and storage of <i>Streptomyces</i> spore stocks	11
2.1.5 Culture density measurements	12
2.2 Molecular cloning and DNA preparation	12
2.2.1 Extraction of plasmids from <i>E. coli</i>	12
2.2.2 Extraction of genomic DNA from <i>Streptomyces</i>	12
2.2.3 Oligonucleotides & sequencing	13
2.2.4 Amplification of target DNA & DNA visualization	13
2.2.5 Gene stitching by overlap extension PCR	14
2.2.6 Restriction enzyme digestion & DNA ligation	14
2.3 Genetic engineering of <i>Streptomyces</i>	15
2.3.1 Conjugation of integrative plasmid vectors	15
2.3.2 CRISPR-mediated deletions	16
2.4 Microscopy & cell imaging	17
2.4.1 Spore phenotype and nucleoid imaging	17
2.4.2 Spore length and nucleoid area calculations	18
2.5 Protein extraction & detection	18
2.5.1 Extraction of proteins from <i>Streptomyces</i>	18
2.5.2 Protein quantification	19

	2.5.3 Protein visualization and detection by Western blotting	19
2.6 <i>In silico</i> analyses & statistical testing		20
	2.6.1 RNA-seq <i>in silico</i> processing	20
2.7 Biochemical Assays		21
	2.7.1 Reporter assay of promoter activity by <i>gusA</i>	21
2.8 WAC strain screening		22
Chapter 3: Characterization of sIHF activity in <i>Streptomyces</i>		31
3.A ACKNOWLEDGEMENTS		31
	3.1 INTRODUCTION	31
	3.1.1 sIHF as a regulator of development and secondary metabolite production	31
	3.1.2 The structure of sIHF and its proposed DNA-binding affinity	31
	3.1.3 Developing a framework to elucidate the function(s) of sIHF	31
3.2 RESULTS		31
	3.2.1 Transcriptome analyses reveal potential genes regulated by sIHF	31
	3.2.2 sIHF DNA-binding interface mutations	34
	3.2.3 sIHF function <i>in vivo</i> is sensitive to mutations at the N-terminal and DNA binding interfaces	35
3.3 DISCUSSION		37
	3.3.1 Differentially expressed genes from sIHF knockout	37

	3.3.2 <i>In vivo</i> effects of targeted sIHF mutations	38
3.4 FUTURE INVESTIGATIONS		40
	3.4.1 Flag-tagging sIHF for ChIP-seq and Co-immunoprecipitation	40
	3.4.2 Proteomic studies into sIHF and manipulation of its activity	42
Chapter 4: Regulation of the silent & cryptic NRPS cluster <i>SCO6429-SCO6438</i> in <i>S. coelicolor</i>		54
4.A ACKNOWLEDGEMENTS		54
4.1 INTRODUCTION		54
	4.1.1 Activation of specific silent/cryptic antibiotic clusters in <i>Streptomyces</i>	54
	4.1.2 Regulation of the silent <i>SCO6429-6439</i> biosynthetic cluster in <i>S. coelicolor</i>	55
4.2 RESULTS		56
	4.2.1 Utilizing a reporter system to assay promoter activity	56
	4.2.2 Detection of a second Crp binding site upstream of the <i>SCO6429-SCO6438</i> cluster	57
	4.2.3 <i>P_{SCO6429}</i> reporters were revealed undetectable activity during growth on solid media	57

	4.2.4 Upstream and intragenic binding sites for Crp influence <i>P_{SCO6429}</i> activity	58
	4.2.5 <i>SCO6439</i> is a CSD-containing regulator with a potential role in controlling <i>SCO6429-38</i> expression.	60
4.3 DISCUSSION		61
	4.3.1 <i>SCO6429-SCO6438</i> operon is regulated by Crp	61
	4.3.2 <i>SCO6439</i> as a putative activator of <i>SCO6429-SCO6438</i> expression	62
4.4 FUTURE DIRECTIONS		63
Chapter 5: Induction of antibiotic production in a library of actinomycetes		71
5.A ACKNOWLEDGEMENTS		71
5.1 INTRODUCTION		71
	5.1.1 Antibiotic screenings are limited by metabolome coverage	71
	5.1.2 Development of a “genetic toolkit” for enhancing antibiotic yield from existing <i>Streptomyces</i> libraries	72
5.2 RESULTS		74
	5.2.1 Select <i>Streptomyces</i> strains show activity against known pathogens	74

	5.2.2 Media conditions have a profound effect on native and Crp-mediated antibiotic production	75
5.3 DISCUSSION		76
5.4 FUTURE DIRECTIONS		78
Chapter 6: Summary		85
	6.1 Summary of work to date	85
APPENDIX		87
	A1	87
	A2	96
REFERENCES		97

LIST OF FIGURES

Figure 1.1: *Streptomyces* development cycle during solid culture.

Figure 2.1: RNA-seq data processing workflow

Figure 3.1: *sIHF* sequence and crystal structure

Figure 3.2: Cluster and PCA plots of RNA-seq data from RNA extracted at three timepoints, from two independent biological replicates

Figure 3.3: Distribution of encoded gene classifications that were upregulated/downregulated by *sIHF* knockout in *S. venezuelae*.

Figure 3.4: EMSAs using mutant *sIHF* proteins

Figure 3.5: Western blotting reveals most *sIHF* mutant variants

Figure 3.6: Spore length measures of the Δ *sIHF* strain complemented with wild type and mutant *sIHF* variants.

Figure 3.7: Nucleoid area measures of *sIHF* mutant complemented strains

Figure 3.8: Sporulation and antibiotic production defects associated with the different *sIHF* mutant variants

Figure 4.1: *SCO6429* operon encoding a predicted dipeptide

Figure 4.2: RNA-seq reads covering *SCO6429-6439*.

Figure 4.3: Crp binding sites C1 and C2 consensus comparison

Figure 4.4: Summary of the different reporter fusions created to monitor the effect of two Crp binding sites on *SCO6429* expression.

Figure 4.5: Reporter assays assessing the transcriptional activity of different promoter fragments (from Fig. 4.3) of *SCO6429*.

Figure 4.6: CRISPR-directed mutational strategy for *SCO6439*.

Figure 5.1: Integrating plasmid vector used in overexpressing the antibiotic regulator Crp.

Figure 5.2: WAC strain antibiotic activity assay

LIST OF TABLES

Table 2.1: Species used

Table 2.2: Working antibiotic concentrations

Table 2.3: Media types

Table 2.4: Plasmids used in this work

Table 2.5: Oligonucleotides

Table 2.6: PCR conditions

Table 3.1: sIHF mutant complementation phenotypes.

Table 5.1: WAC strains hits for bacterial growth inhibition.

Table 5.2: Media types used to enhance antibiotic production in WAC strains, as determined by bioassay.

LIST OF ABBREVIATIONS

ABBREVIATION	MEANING
A	alanine
A.U.	arbitrary units
A_{405NM}	absorbance at 405nm wavelength
C1	Crp site 1
C2	Crp site 2
CAMP	cyclic-adenine monophosphate
CDA	calcium-dependent antibiotic
CHIP-SEQ	Chromatin immunoprecipitation-sequencing
CRP	cAMP-receptor protein
CSD	cold shock domain
DAPI	4',6-diamidino-2-phenylindole
DAVID	Database for Annotation, Visualization, and Integrated Discovery.
DH₂O	distilled water
DINA	Difco nutrient agar
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acids
DNTP	deoxyribonucleoside triphosphates
EDTA	Ethylenediaminetetracetic
H-NS	histone-like nucleoid structuring protein
IC-MASS SPECTROMETRY	ion chromatography mass spectrometry
IGV	integrated genomics viewer
IHF	integration host factor
KDA	kildalton
LB	Luria-Bertani broth
MB	megabase
MG	milligrams
MIN	minutes
ML	milliLitres
MM	milliMolar
MYM	maltose-yeast extract-malt extract
N	asparagine
NAP	nucleoid-associated protein
NG	nanograms

NRPS	Non-ribosomal peptide synthetase
O/E	overexpression
°C	degrees celsius
OD	optical density
PAM	protospacer adjacent motif
PCR	polymerase chain reaction
PNPG	4-Nitrophenyl β -D-glucopyranoside
Q	glutamine
R	arginine
RNAP	ribonucleic acid polymerase
RNASE	ribonuclease
RNA-SEQ	RNA sequencing
RT-QPCR	real time quantitative polymerase chain reaction
S	serine
SAXS	Small-angle X-ray scattering
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	seconds
SFM	Soy flour mannitol
SIHF	<i>Streptomyces</i> integration host factor
SOB	super optimal broth
T1	timepoint 1
T2	timepoint 2
T3	timepoint 3
TBE	Tris Boric Acid EDTA buffer
TF	transcription factors
UTR	untranslated region
WAC	Wright actinomycete collection
X-GAL	5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside
YEME-TSB	Yeast extract-malt extract - Tryptone soy broth

1 INTRODUCTION:

1.1 *Streptomyces* as a bacterium of unique development and historic antimicrobial importance

Streptomyces is a genus of ubiquitous Gram-positive bacteria that inhabit soil and marine environments and are renowned for their production of secondary metabolites (Barka et al., 2016). These bacteria belong to the order *Actinomycetales*, which includes *Mycobacterium*, *Corynebacteria*, and *Propionibacteria*: all species of significant inquiry for their role in pathogenesis (Beylot et al., 2014; Rahman et al., 2014; Sangal & Hoskisson, 2016).

The lifecycle of *Streptomyces* is remarkably similar to that of fungi. *Streptomyces* spores germinate in nutrient rich conditions and develop as branching mycelia through hyphal-tip extension and branching (Figure 1). When the mycelia become sufficiently dense and/or nutrients become limited, *Streptomyces* will raise aerial hyphae and sporulate (Mcgregor, 1954). Sporulation yields a dormant cell state covered by an exterior amyloid sheath (Williams et al., 1972). Spores are sheared from the spore chain by mechanical wind and water forces, dispersing them to colonize distant locales (Ruddick & Williams, 1972).

Upon initiating aerial development, streptomycetes upregulate the production of secondary metabolites (Bibb, 2005). Notably, secondary metabolism is spatially segregated from the aerial hyphae, being largely confined to the vegetative cells (Bibb, 2005). The secondary metabolites can serve as ion scavengers to maintain homeostasis, secondary-messenger/quorum sensing signals, or as potent inhibitors of competing organisms (e.g. Hassan et al., 2016; Hendlin et al., 2017). It is thought that the coupling of aerial development with the production of these molecules serves as a bet-hedging strategy that allows streptomycetes to confront encroaching organisms while simultaneously deploying spores to find a less competitive niche (Chater, 2006; Demain, 1974). The diversity of compounds produced by *Streptomyces* was originally recognized in the 1940s, and has since been found to include many medically-relevant drugs such as

anti-cancer agents, anti-parasitic compounds, and especially antibiotics (Lam, 2006; Watve et al., 2001). Approximately 67% of the antibiotics in clinical use are natural products, or derivatives of compounds produced by *Streptomyces* (Watve et al., 2001). Unfortunately, the rate of antibiotic discovery has significantly declined since the 1980s and the frequent use of these drugs has led to a growing number of antibiotic resistant pathogens (French, 2010; Ventola, 2015). There is, therefore, an immediate demand for novel antibiotics, prompting further investigation of the *Streptomyces* metabolome.

1.1.1 Model species for study within *Streptomyces*

The model species for studying the streptomycetes is *Streptomyces coelicolor*, which develops canonically on solid substrates and produces two characteristic pigments: the mycelium-bound red-coloured undecylprodigiosin, and the secreted blue-coloured actinorhodin (Hobbs et al., 1990; Wezel et al., 2000). In liquid media cultures, it grows in mycelial clusters and does not sporulate: developmental studies of this organism require growth on solid media (Manteca et al., 2008). *S. coelicolor* is slow-growing (4-5 days to robust colony formation), making genetic manipulation time-intensive. This has prompted increased study of the emerging model species *Streptomyces venezuelae*. *S. venezuelae* is faster growing, sporulates in liquid media, and produces a brown pigment (melanin) that marks the onset of secondary metabolite production (Glazebrook, Doull, Stuttard, & Vining, 1990). These characteristics have made *S. venezuelae* an ideal candidate for evaluating the impact of genetic alterations in *Streptomyces*. While most characterized genetic elements in *Streptomyces* have been studied primarily in *S. coelicolor*; there is considerable value in conducting genetic research concurrently in both species.

S. coelicolor and *S. venezuelae*, like all streptomycetes, have large, linear genomes of ~9 MB (Bentley et al., 2002; Hopwood, 2006). The chromosomes are organized such that most conserved genes are found in the middle “core” of the chromosome, while the flanking “arms” contain more variation between species (Bentley et al., 2002; Hopwood, 2006). The “arms” also tend to harbour most of the biosynthetic clusters that direct secondary metabolite production (Bentley et al., 2002; Chen et al., 2002). Such

secondary metabolites are primarily produced from polycistronic operons directing the production of a dedicated set of synthesis/tailoring enzymes, secretion channels and/or resistance determinants (Bibb, 2005; Lawrence & Roth, 1996). Additionally, pathway-specific regulators are often encoded within these clusters and exert specific control over cluster activity (Huang et al., 2001). These regulators may themselves be subject to extensive regulation by other elements. Each *Streptomyces* species has a unique complement of secondary metabolic gene clusters, suggesting that this genus has the capacity to produce an enormous and diverse range of bioactive natural products (Hopwood, 2006).

1.1.2 Potential of *Streptomyces* for future bioactive molecule discoveries

Streptomyces are thought to possess a large, untapped reservoir of bioactive compounds, as the quantity of secondary metabolic clusters inferred from genome sequencing far exceeds the quantity of secondary metabolites identified (Challis, 2008). In the well-studied *S. coelicolor* alone, there are 22 biosynthetic clusters that are transcriptionally silent and whose products are unknown or undetected (“cryptic”) under conventional laboratory growth conditions. Indeed, it has been demonstrated that some *Streptomyces*-synthesized compounds are only produced under particular environmental stimuli (Anzai et al., 2009; Holden et al., 1998; Scherlach & Hertweck, 2009). This suggests that efforts to discover new clinically-relevant antibiotics may suffer from a culture-medium bias. Typically, high-throughput screening for novel antibiotics is achieved by growing a library of microorganisms, especially *Streptomyces* species, on a particular medium and then using solvating chemicals to extract metabolites (Fiedler, 1993; Miyadoh, 1993). These extracts are then screened for activity, purified, re-tested and identified (Kumar et al., 2014). It is inferred that the bias imposed by any particular growth medium prevents bioactivity screening from accessing a large portion of the streptomycete metabolome (Katz et al., 2016; Ramadhar et al., 2013). As a consequence, current efforts to screen microorganisms for antibiotic activity are plagued by a high rate

of rediscovery - hits corresponding to compounds that are already known (M. Katz et al., 2016; Ramadhar et al., 2013).

Cryptic metabolites may comprise a wealth of novel bioactive compounds, but require stimulation by specific environmental conditions (Scherlach & Hertweck, 2009). It is critical that we develop a better understanding of the regulatory mechanisms governing secondary metabolism in *Streptomyces*, such that we can activate the cryptic metabolic clusters. The complex development cycle of *Streptomyces* and their multitude of secondary metabolic clusters, suggest that these bacteria employ intricate regulatory networks to control gene expression. While some of these elements have been identified in other organisms, there are many unknown regulatory networks associated with development and metabolism in *Streptomyces*.

1.2 Transcription regulation in *Streptomyces* is complex

In bacteria, sigma factors associate with RNA polymerase (RNAP), and direct RNAP to distinct promoter sequences. The housekeeping sigma factor σ -70 in *E. coli* recognizes a TATAAT sequence (-10 region), and a TCTTGACAT sequence (-35 region) positioned ~17 nucleotides upstream (Lisser & Margalit, 1993). *E. coli* encodes seven sigma factors, while *S. coelicolor* encodes an astonishing 65 sigma factors (Bentley et al., 2002; Helmann & Chamberlin, 1988; Mazurakova et al., 2006). Few *Streptomyces* sigma factors have been studied, and as such, it is prohibitively difficult to determine promoter locations based on empirically-derived consensus binding sequences. Instead, the regulation of promoter activity is derived from methods to identify transcription start sites, transcription factor binding sites, and changes in mRNA transcript abundance under particular conditions.

Global transcription factors control the expression of large numbers of genes in bacteria. In *Escherichia coli*, over 50% of the genome is influenced by seven regulator proteins, Crp, FNR, IHF, Fis, ArcA, NarL, Lrp (Martínez-Antonio & Collado-Vides, 2003). Some of these proteins have homologues in the *Streptomyces*, including the cAMP-receptor protein (Crp). Crp binds to its activating cofactor cAMP, increasing its

DNA binding affinity for a consensus sequence (a 22-bp sequence in *E. coli*) upstream of a promoter (Liu-Johnson et al., 1986). Crp interacts with the α -C-terminal domain (CTD) of RNA polymerase (RNAP), recruiting it to the promoter and increasing transcriptional activity (Niu, et al., 1996).

In *Streptomyces*, Crp is a regulator of development that binds sites throughout the chromosome and affects the expression of many genes including those directing antibiotic production (Gao et al., 2012). In many cases this effect is thought to be direct, as chromatin immunoprecipitation (ChIP) experiments have revealed Crp to bind regions within 8 of the 22 silent biosynthetic clusters in *S. coelicolor* (Gao et al., 2012). Deletion of *crp* results in delayed growth, hypersporulation, and reduced production of the blue-pigmented antibiotic actinorhodin and the calcium-dependent antibiotic (CDA) (Gao et al., 2012).

1.2.1 Transcriptional control by nucleoid-associated proteins

Transcription factors can have positive or negative effects on gene expression, and while activators are almost always found to bind upstream of promoter regions, repressors have more diverse binding sites (Browning & Busby, 2004). Many repressors act by binding to the promoter itself, occluding binding by RNAP, while others compete with transcriptional activators for binding (Browning & Busby, 2004). Others still act to alter the spatial position and torsion of the DNA, and consequently their effect on RNAP is indirect (Becker et al., 2007). Indeed, the physical topology and organization of the DNA within the cell has a significant impact on gene expression (Zimmerman & Murphy, 1996). Proteins affecting DNA topology in bacteria have been generally termed ‘nucleoid-associated proteins’ (NAPs) and tend to be abundant within the cell (Wang et al., 2012). Analogous to the histones present in eukaryotes that can shut down gene transcription by condensing the DNA, NAPs promote DNA condensation by bending, bridging, and/or wrapping DNA (Wang & Maier, 2015; Wang et al., 2012). Some NAPs coordinate chromosome replication and segregation during cell division, while others

influence global or local gene expression by inducing a specific state of supercoiling (Bernhardt & De Boer, 2005).

Many transcription factors (TFs) bend DNA in facilitating the recruitment of the transcriptional machinery, and in doing so affect the architecture of the chromosome (Horikoshi et al., 1992). Similarly, the structural effects of NAPs on DNA are associated with changes in transcription (Dillon & Dorman, 2010). Proteins belonging to these groups exhibit a range of DNA binding affinities, consensus sequence recognition motifs, and gene expression effects. Thus, there is significant overlap in the activities of TFs and NAPs and a lack of any definitive, distinguishing features. Understanding the role of NAPs in the organization of the *Streptomyces* chromosome, and in the control of gene expression, will ultimately provide novel insights into the regulatory networks influencing transcription initiation. This understanding will allow us to exploit or modulate NAP activity, such that we may activate the transcription of lowly expressed operons in *Streptomyces*, including silent secondary metabolic clusters.

1.2.2 Properties of nucleoid-associated proteins

1.2.2.1 *Histone-like nucleoid structuring protein*

One of the best-studied NAPs is the histone-like nucleoid structuring protein, or H-NS, from *E. coli*. H-NS frequently binds at multiple sites upstream of promoters, and can either bridge duplex DNA strands, and/or oligomerize along the DNA (Esposito et al., 2002). This aggregated structure can block RNA polymerase from accessing the promoter and introduce superstructure domains into the DNA (Esposito et al., 2002; Liu et al., 2010; Oshima et al., 2006). H-NS preferentially binds A-T rich sequences (Oshima et al., 2006). Horizontally-acquired genes are often enriched for A-T residues, and are demonstrably repressed by H-NS activity. It is theorized that H-NS acts as a repressive sentinel to mitigate the effects of acquired genes that may be detrimental to the cell (Oshima et al., 2006).

While NAPs are integral to the chromosome organization and regulation in all bacteria, different species have distinct NAP complements. Though *E. coli* uses H-NS as a global silencer, the genus *Pseudomonas* employs the similarly structured MvaT/MvaU, and *Bacillus* uses Rok (Li et al., 2009; Smits & Grossman, 2010). There is no H-NS found in the actinobacteria; however, *Mycobacterium* possess an H-NS-like protein termed Lsr2, which has an orthologue in *Streptomyces*. Lsr2 does not share any sequence homology with H-NS, but it is functionally interchangeable (Gordon et al., 2008). Lsr2 acts as a dimer and binds to a nucleation site rich in A-T sequences, where it can then oligomerize along the DNA, compacting the chromosome and occluding transcriptional start sites (Chen et al., 2008).

1.2.2.2 Integration host factors

Another known NAP, the integration host factor (IHF) is a NAP originally identified in *E. coli* that binds DNA as a heterodimer, bending the local DNA around it (Lin et al., 2012). Its name is derived from initial observations that it was necessary for the integration of phage DNA into the chromosome (Craig & Nash, 1984). However, subsequent investigation revealed that IHF also impacts the expression of many genes including ATP-binding cassette (ABC) transporters, and is necessary for CRISPR-mediated integration of ssDNA to form new protospacer elements (Lin et al., 2012). It is also involved in the organisation of DNA inside the cell as it can form higher order DNA structures and influence gene expression, making it a histone-like protein (Lin et al., 2012). IHF is found extensively throughout the *Enterobacteriaceae* and related organisms.

There is no obvious IHF homologue in *Streptomyces*, although a functional homologue of IHF has been identified in *Mycobacterium*, where it is termed mIHF (Pedulla & Hatfull, 1998). The orthologue of mIHF in *Streptomyces* is a 14 kDa protein that is highly conserved within the genus and is one of the most abundant proteins in the cell (Swiercz et al., 2013). mIHF improves phage integration into the chromosome but lacks a specific DNA-recognition site, unlike IHF (Pedulla & Hatfull, 1998). *mIHF* is an

essential gene in *M. smegmatis* and its deletion affects the growth of *M. tuberculosis* (Pedulla & Hatfull, 1998; Sasseti et al., 2003). It is capable of bending DNA throughout the chromosome, compacting it (Mishra et al., 2013). It is thought to be a general transcriptional regulator with broad implications for cell development (Mishra et al., 2013).

1.3 Aims of this work

Here, we investigate the nucleoid-associated protein *Streptomyces* integration host factor (sIHF), which non-specifically binds DNA and impacts development and antibiotic production. Through targeted mutations, we probe the interfaces of sIHF thought responsible for its interactions with DNA and assess the relevance of these interfaces to *in vivo* activity. Further, we examine the transcriptionally silent *SCO6429-SCO6438* biosynthetic cluster and attempt to activate transcription from its promoter by manipulating its potential regulatory proteins. Binding sites for Crp found in close proximity to this cluster led to our analysis of Crp effects on cluster expression, alongside the role of a potential pathway specific regulator. Finally, we overexpress Crp in a library of *Streptomyces* and assay these strains for novel bioactivity against known clinical pathogens.

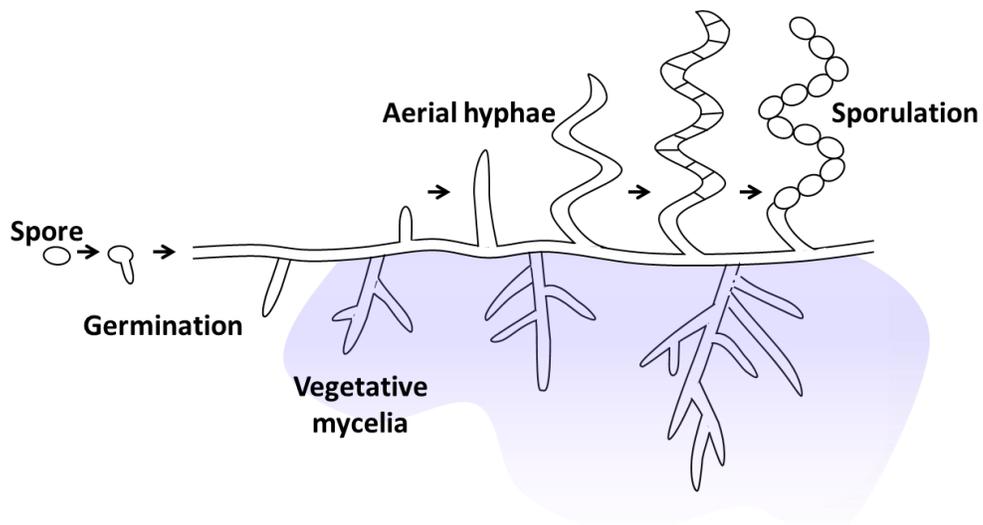


Figure 1.1: *Streptomyces* development cycle during solid culture. Spores germinate in nutrient-rich environments and grow by hyphal-tip extension and branching, forming a penetrative vegetative mycelia. As vegetative mycelial density increases, *Streptomyces* raise aerial hyphae and secrete secondary metabolites (blue). Aerial hyphae metamorphose into spore chains, and individual spores are then dispersed by environmental and mechanical forces.

Chapter 2: Materials and Methods

2.1 Bacterial culturing, transformation and selection

2.1.1 Preparation of cultures & storage of strain stocks

E. coli cultures (strains summarized in Table 2.1) were grown for 16-22 h in Luria-Bertani (LB) broth liquid medium (Miller, 1972) or Super Optimal Broth (SOB) medium (Hanahan, 1983), supplemented with antibiotics when appropriate (as indicated in Table 2.2). To create stocks, 700 μ l of a dense overnight culture mixed with 700 μ l of sterile 20% glycerol in screwcap tubes and stored at -80°C. *E. coli* selection on solid medium was carried out on solid LB or Difco-nutrient agar (DNA), with appropriate antibiotics. *Streptomyces* cultures (strains summarized in Table 2.1) were grown in or on a range of media (see Table 2.3).

2.1.2 Preparing chemically competent & electrocompetent *E. coli* cells

Overnight cultures of *E. coli* cells were subcultured in 50 ml of LB liquid medium to an optical density at 600 nm (OD_{600}) of 0.6-0.8 and pelleted in a benchtop centrifuge at 5000 $\times g$ for 5 min at 4°C. To prepare chemically competent cells for heat shock, cell pellets were washed three times by resuspension in 10 ml ice-cold 100 mM $CaCl_2$, before collecting them again by centrifugation and decanting the supernatant. Following the third wash, cell pellets were resuspended in 2 ml of 100 mM $CaCl_2$ and 45 μ l aliquots were dispensed into 1.5 ml microcentrifuge tubes. To prepare electrocompetent cells, *E. coli* subcultures were collected once they reached an optical density (OD_{600nm}) of 0.6-0.8. Cells were then washed (pelleted, and then resuspended) three times in 10 ml of 4°C 10% sterile glycerol. After the third wash, the resulting cell pellet was resuspended in 1 ml of 10% sterile glycerol and aliquoted into 50 μ l volumes in 1.5 ml microcentrifuge tubes.

2.1.3 Transformation of *E. coli* cells for molecular cloning

Plasmids (summarized in Table 2.4) were heat-shocked into chemically competent *E. coli* DH5 α cells (smaller constructs) (Invitrogen) or electroporated into electrocompetent *E. coli* DH5 α cells (larger constructs). Introduction of plasmid DNA into chemically competent *E. coli* cells entailed mixing 3 μ l of extracted plasmid stock with 45 μ l of chemically competent cells in 1.5 ml microcentrifuge tubes and incubating on ice for 15 min. The cells were then heat shocked by placing the tubes in a 37°C water bath for 30 s, and then placed immediately on ice for 2 min. For electroporations, 3 μ l of extracted plasmid stock were mixed with 50 μ l of electrocompetent cells, and incubated on ice for 15 min. Cells were transferred to an electroporation cuvette and pulsed once in an electroporator (BioRad) set to EEC2/bacteria/time (ms). Both heat-shocked and electroporated cells were then grown in 1 ml of LB medium for 1 h at 37°C, shaking at 200 rpm. Cells were then pelleted by centrifugation in a benchtop centrifuge at 5000 \times g for 5 min at 4°C, prior to decanting of the supernatant, and resuspension of the cells in the residual medium. Resuspended cells were plated on LB or DNA solid agar medium supplemented with appropriate antibiotics, and were grown for 16-22 h until single colonies developed.

2.1.4 Harvesting and storage of *Streptomyces* spore stocks

Soy flour mannitol (SFM) plates overlaid with sterile cellophane disks were inoculated with a cell suspension that resulted from the crushing and resuspending of a single *Streptomyces* colony in 200 μ l of sterile water. These plates were grown for 4-5 days at 30°C until cell growth was confluent and the colonies were sporulating. A sterile flat edge was used to scrape the cell biomass from the cellophane disk into 10 ml of sterile water in a glass vial. These solutions were then sonicated in a water bath sonicator for 2 min until the solutions turned a dark colour and only small pieces of debris remained. Spores were then filtered through a small bed of sterile cotton batting in a syringe, into 15 ml Falcon tubes and centrifuged at 5000 \times g for 5 min. The water was decanted and the pelleted spore stocks resuspended in sterile 20% glycerol.

2.1.5 Culture density measurements

The density of *E. coli* and *S. venezuelae* liquid cultures was determined using a spectrophotometer measuring the OD₆₀₀ relative to a liquid medium blank. *S. coelicolor* culture growth was monitored using cell weight: 1 ml culture medium was centrifuged at 5000 ×g at 4°C for 5 min, the supernatant extracted, and the resulting cell pellet was weighed after drying.

2.2 Molecular cloning and DNA preparation

2.2.1 Extraction of plasmids from *E. coli*

The Miniprep PureLink extraction kit (Invitrogen) was used to isolate plasmids from *E. coli*. One microlitre of an overnight culture of *E. coli* was centrifuged at 10,000 ×g for 1 min to collect the cells. The resulting cell pellet was resuspended and lysed, and the plasmid DNA extracted as per the manufacturer's instructions.

2.2.2 Extraction of genomic DNA from *Streptomyces*

Streptomyces genomic DNA was extracted either using a phenol-chloroform-based extraction procedure, or using the Genomic DNA isolation kit (Norgen), following the manufacturer's instructions. In each case, *Streptomyces* spores (~3000) were used to inoculate a 10 ml liquid culture composed of 5 ml tryptone-soya broth (TSB) and 5 ml yeast extract-malt extract (YEME) media. This culture was grown at 30°C for 4 days, after which 2 ml were pelleted by centrifugation at 5000 ×g for 5 min at 4°C. Phenol-chloroform extractions then involved cells being resuspended in R3 lysis buffer + RNase (Invitrogen) followed by the addition of a 49:49:1 phenol:chloroform:isoamyl alcohol solution in a 1:1 volumetric ratio. The cell-phenol solution was vortex mixed before being centrifuged at 8000 ×g in a Sorvall benchtop centrifuge for 5 min. The upper aqueous phase was then extracted. Genomic DNA was precipitated by adding two volumes of 95% ethanol and incubating at -20°C for 2 h. DNA was pelleted by centrifugation at 8000 ×g and 4°C for 5 min, after which the ethanol was decanted, and

the resulting DNA pellet air-dried before being resuspended in ~200 µL of nuclease-free water.

When checking the genomic integrity of *Streptomyces* strains that had been subjected to genetic engineering, in lieu of extracting genomic DNA, template genomic DNA was instead prepared rapidly by transferring a 1 mm² area of confluent cell growth on agar plates into 30 µl of nuclease-free water. This cell suspension was then incubated at 95°C for 10 min. Three microliters of the resulting DNA suspension was then used as template DNA for amplification by PCR, as described below.

2.2.3 Oligonucleotides & sequencing

Oligonucleotides used as primers for cloning are listed in Table 2.5 and were designed with the assistance of the programs SnapGENE and NCBI Primerblast. All oligonucleotides were purchased from MOBIX DNA facilities (McMaster University) operating in conjunction with Integrated DNA Technologies. MOBIX DNA facilities also provided all sequencing of prepared plasmid constructs and PCR samples.

2.2.4 Amplification of target DNA & DNA visualization

Taq DNA polymerase (Invitrogen) was used for determining optimal primer annealing conditions, validating the success of cloning reactions and plasmid conjugation into *Streptomyces*, and assessing clone orientation when necessary. Amplification reactions were conducted as per the manufacturer's instructions, while reaction conditions varied depending on the primers used, and the desired product length. Typical reaction and amplification programs are outlined in Table 2.6. High-Fidelity Phusion polymerase (New England Biolabs) was used as recommended by the manufacturer, to amplify any DNA fragments to be used for cloning purposes (New England Biolabs, Phusion polymerase manual). Typical amplification programs are summarized in Table 2.6. PCRs were carried out in an Eppendorf or Biorad thermocycler. When needed for subsequent reactions (*e.g.* cloning), the amplified DNA was purified using a PCR purification kit

(Invitrogen) (when a single reaction product was obtained), or a DNA gel extraction kit (Qiagen) (when more than one reaction product was obtained). Large batches of DNA were purified using phenol/chloroform extraction. DNA was separated on a 1% agarose Tris-Borate-EDTA (TBE) gel containing ethidium bromide at 100 V in 1× TBE buffer. Separated DNA was visualized using UV light (395 nm wavelength).

2.2.5 Gene stitching by overlap extension PCR

To combine two DNA sequences without using restriction enzyme-mediated attachment, as in the case for CRISPR right and left flank stitching, overlap extension PCR was used. This process involved individually amplifying the two DNA segments that were to be stitched together. One of the two products was amplified using primers having complementary extensions on one end, such that the 3' end of that sequence overlapped with the 5' end of the other sequence by 17-20 nucleotides. These two DNA products were then purified, and added into a typical Phusion PCR reaction mixture in a 1:1 molar ratio without any additional primers, before being subjected to 10 cycles of DNA extension, to generate a full-length product. Flanking primers were then added to the mixture and a conventional Phusion PCR protocol was performed to amplify the 'overlapped' DNA. The resulting DNA product was then gel extracted if necessary.

2.2.6 Restriction enzyme digestion & DNA ligation

All DNA digests were performed in the Cutsmart Restriction Enzyme buffer system (New England Biolabs), unless the selected enzyme had an alternate recommended buffer for optimal activity. DNA was diluted 1/50 in Cutsmart buffer to a final volume of 100 µl, before adding 5-10 units of the appropriate restriction enzyme(s), and incubating in a 37°C water bath for 30 min. An additional 5-10 units of enzyme were added to the reaction, followed by an additional incubation of up to 2 h. Digested plasmid DNA that was to be used for cloning purposes was dephosphorylated using 1 µl of shrimp alkaline phosphatase (Invitrogen), and the reaction was incubated at 37°C for 1 h. Blunt-end cloning of DNA fragments required phosphorylation of the insert DNA by polynucleotide

kinase (Invitrogen) for 1 h at 37°C. Ligation of DNA inserts into digested plasmids was carried out using the T4 DNA ligase kit (Invitrogen), where a 3:1 insert:vector molar ratio was mixed together in a 10 µl reaction (~100 ng of DNA), and this reaction was incubated overnight at 16°C before introduction via heat-shock into commercially competent *E. coli* DH5α cells (Invitrogen). Alternatively, ligations were carried out using the Rapid DNA Ligation kit (Roche), where again, a 3:1 insert:vector molar ratio was combined in a 20 µl reaction and incubated at ambient temperature for 10 min, before being introduced into *E. coli* DH5α cells, as described above.

2.3 Genetic engineering of *Streptomyces*

2.3.1 Conjugation of integrative plasmid vectors

Plasmid vectors used in this study are outlined in Table 2.4. All conjugative vectors used here harbor a ΦC31 or ΦBT1 integrase site, the corresponding phage integrase gene, an *E. coli* origin of replication, an origin of transfer, and selectable antibiotic resistance marker(s).

Introducing plasmids into *Streptomyces* strains was achieved by conjugation from a methylation-deficient *E. coli* strain (*E. coli* ET12567) carrying a conjugation competent ‘helper plasmid’ pUZ8002. Plasmids were electroporated into *E. coli* ET12567/pUZ8002 and successful transformants were selected using antibiotics specific for the plasmid resistance marker. A single colony was then selected and inoculated into LB medium supplemented with the plasmid-specific antibiotic, along with kanamycin (50 µg/ml) and chloramphenicol (25 µg/ml) (to maintain selection for the pUZ8002 plasmid, and the methylation mutation), and was grown in liquid medium for 16-22 h. The cells were then subcultured into 50 ml of LB liquid medium supplemented with the same antibiotics as the overnight culture, and grown to an OD₆₀₀ of 0.6-0.8. Cells were collected by centrifugation at 8000 ×g for 5 min and were washed by sequentially resuspending and pelleting the cells three times with fresh LB medium. Cells were finally resuspended in 3 ml of fresh LB medium. Ten microlitres of a concentrated *Streptomyces* spore stock (approximately 50 million spores/µl) were added to 500 µl of 2×YT (yeast extract-

tryptone) liquid medium and incubated at 50°C for 10 minutes. This spore suspension was cooled on ice for up to 2 h and then combined with 500 µl of the *E. coli* ET12567/pUZ8002 plasmid-carrying strain and plated on solid SFM medium. The *E. coli-Streptomyces* mixture was then grown at 30°C for 16-22 hours before being overlaid with 1 ml sterile water containing appropriate antibiotics for both selection of the conjugative plasmid and selection against the *E. coli* plasmid donor (antibiotic concentrations used are described in Table 2.2). Strains were incubated for an additional 5 days, or until exconjugant colony growth was robust. These colonies were then re-streaked on the appropriate selective medium and a single colony was used to generate spore stocks (Kieser et al., 2000).

PCR was used to validate the successful chromosomal integration of the conjugative vector. Briefly, genomic DNA was isolated and was used as template DNA, in conjunction with insert and vector specific primers that provided a diagnostic product. The resulting PCR fragment was then purified and sequenced.

2.3.2 CRISPR-mediated deletions

Deletion of *SCO6439* was performed using pCRISPomyces-2 (Table 2.4), a conjugative vector harboring the Cas9 protein-encoding gene, an *E. coli* origin of replication, a temperature sensitive *Streptomyces* origin of replication, an origin of transfer, a guide RNA (gRNA), and an editing template, using a protocol adapted from (Cobb et al., 2015). The gRNA protospacer was selected by finding an NGG nucleotide motif (protospacer adjacent motif, or PAM) on the non-template strand that, in combination with the immediate upstream 15 nucleotides, did not have complete complementarity elsewhere in the genome, as determined using BLASTn (NCBI, (Altshul, et al, 1990)). The gRNA protospacer was produced by annealing two nucleotides comprising complementary sequences specific for the 20 nucleotides upstream of the PAM sequence (gRNA-6439 For, gRNA-6439 Rev), with single stranded BbsI specific overhangs on either end. Introduction of the protospacer sequence into pCrispomyces-2 was carried out by Golden Gate assembly and confirmed by sequencing

using the downstream pMC600R primer (Table 2.5). Separately, the editing template for CRISPR was designed by PCR amplifying two 1 kb products flanking the protospacer/PAM region (1kLF/1kRF). The intervening sequence ultimately represents the deletion that will be created, following the stitching together of the two flanking products (as per Section 2.2.4), and the ultimate double cross-over recombination that occurred between them and the chromosomal DNA. Following stitching together of the two flanking sequences, the resulting 2 kb product was cloned into the unique *Xba*I site in pCrispomyces-2. The resulting construct was then conjugated into *S. coelicolor*, and grown at 30°C on SFM agar, overlaid with apramycin (to select for the plasmid) and nalidixic acid (to select against the donor *E. coli*). Exconjugants were re-streaked on SFM medium (without nalidixic acid) and grown at 37°C for 4 days until single colonies emerged; the pCRISPomyces *Streptomyces* origin of replication is temperature sensitive, and thus growing at 37°C is sufficient to cure the thermosensitive vector. Confirmation of plasmid curing was achieved by patching colonies onto SFM, and SFM supplemented with apramycin, and growing the cultures at 30°C for 4 days. Those colonies that were now sensitive to apramycin (and had thus lost the pCRISPomyces plasmid) were screened via PCR. Potential mutants were then further tested using PCR amplification across the deleted region, followed by sequencing to confirm the deletion of interest.

2.4 Microscopy & cell imaging

2.4.1 Spore phenotype and nucleoid imaging

Coverslips were immersed in ethanol, flame-sterilized, and inserted at a 45° angle in SFM plates. *S. coelicolor* Δ *sihf* strains were inoculated underneath the coverslip slants (2 μ l, ~5000 spores), and were grown for 4 days at 30°C. The coverslips, with adherent cells, were then removed from the agar, fixed by flooding the coverslip with 100 μ l of methanol and allowing to dry. Two microlitres of DAPI stain diluted 1 in 5 in provided dilution solution (Invitrogen) applied immediately to the fixed cells for 2 mins, before the cells and coverslip were washed briefly with 1 mL of sterile water. Coverslips were mounted on 10 μ l of 10% glycerol on a microscope slide, and the edges were sealed using clear

nail polish. DAPI-stained cells were imaged using a Zeiss Axio microscope with a 100× oil-immersion objective.

2.4.2 Spore length and nucleoid area calculations

Spore lengths for $\Delta sihf$ mutant and complemented strains were determined by measuring the maximum length of individual spores using the measure tool in FIJI-ImageJ (Schindelin et al., 2012; Schneider, 2012), and a 0.06 $\mu\text{m}/\text{pixel}$ conversion ratio. Nucleoid areas were determined using the measure>particle analysis tool in FIJI-ImageJ with the Otsu threshold option. More than 270 spore lengths/nucleoids were measured for all strains. A chi-square test was performed to evaluate differences between nucleoid areas or spore lengths for mutant and complemented strains.

2.5 Protein extraction & detection

2.5.1 Extraction of proteins from *Streptomyces*

Fifty millilitre cultures of 1:1 YEME-TSB-grown *Streptomyces* strains ($\Delta sihf$ containing the integrating plasmid pIJ82 alone, or bearing wildtype and mutant variants of *siHF*; Table 2.4) were pelleted by centrifugation prior to the cells being resuspended in lysis buffer (10 mM Tris-HCl, 10 mM KCl, 50 mM NaCl, 2 mM MgCl₂, 1 mM DTT, EDTA-free protease inhibitor (Roche), 0.4 mg/ μl lysozyme). Cell suspensions were then incubated for 1 h at 37°C in a water bath and inverted every 5 min. Following this incubation, the suspensions were subjected to sonication using a tip sonicator at duty 40%, Power level 5, in 10 s bursts, interspersed by 10 s on ice, for 2 mins, or until the suspension achieved a silky texture. Suspensions were centrifuged at 4°C for 30 mins at 12,000 $\times g$ to separate soluble proteins from cell debris and insoluble compounds. The supernatant was then collected and stored at -20 °C.

2.5.2 Protein quantification

To determine the concentration of proteins in each protein extract, soluble proteins were quantified using a Bradford assay. Bradford dye reagent (BioRad) was diluted 1 in

5 in ultrapure water prior to being filtered through Whatman #1 filter paper to remove any settled aggregates. Ten microlitres of each protein sample was added to wells in a 96 well plate in technical triplicate, after which 200 μ l of the diluted Bradford reagent was added. The samples were incubated at ambient temperature for 5 min, prior to measuring the absorbance at 595 nm. A standard curve was generated using BSA concentrations ranging from 0.06 mg/ml to 1 mg/ml, and after graphing the absorbances versus protein concentration, the resultant slope was used to determine the quantity of protein in each of the *Streptomyces* extracts. To ensure that absorbances were within the linear dynamic range of the assay (specified by the BSA standards), samples were diluted with water when necessary.

2.5.3 Protein visualization and detection by Western blotting

To detect sIHF in the Δ *sihf* complemented strains, protein samples were diluted to the lowest concentration, diluted in 4 \times Buffer A loading buffer (40% glycerol, 200 mM Tris-HCl pH 6.8, 8% SDS, 0.04% bromophenol blue, 5% 2-mercaptoethanol) and boiled at 95°C for 10 min. Samples were briefly centrifuged, prior to loading 15 μ L (1 mg) of protein extract onto a 15% Tris-Tricine SDS-PAGE gel. Samples were electrophoresed for 1.5 h at 200 V. To ensure that equivalent amounts of proteins were run for each sample, one gel was stained with Coomassie Brilliant Blue for 1 h while shaking, followed by washing three times for 15 mins in destain (40% methanol, 10% glacial acetic acid) while shaking. The second gel was subject to western blotting. This involved first treating a PVDF membrane with methanol, and then equilibrating it in transfer buffer (25 mM Tris pH 8.0, 190 mM glycine, 20% methanol, pH 8.3). The equilibrated membrane was then overlaid with the SDS-PAGE gel, and the membrane/gel were then placed between two filter membranes soaked in transfer buffer. A cylinder was pressed and rolled over the stack, in order to squeeze out excess transfer buffer and remove bubbles between the gel and PVDF membrane. Proteins were transferred from the SDS-PAGE to the PVDF membrane using a BioRad Semi-Dry transfer apparatus at 15 V for 1 hour. sIHF detection was initiated by blocking the membrane in blocking solution [TBS-

T (Tris buffered saline: 20 mM Tris pH 8.0, 150 mM NaCl)-0.01% Tween) with 5% skim milk powder] for 1 h. The membrane was then washed three times in 10 ml of TBS-T for 15 min each. Next, incubation with the primary (sIHF-specific polyclonal) antibody (Swiercz, 2013) involved a 1 h incubation of the blocked membrane in blocking solution, into which the primary antibody was added in a 1/500 concentration. Following this incubation, the membrane was then washed three times in TBS-T, again for 15 min per wash. The final membrane incubation involved blocking solution containing a 1/3000 dilution of secondary antibody [goat-anti-rabbit antibody, HRP-labelled (Invitrogen)]. After 1 h, the membrane was subjected to a final three washes in TBS-T, as described above. sIHF detection was achieved using horseradish peroxidase detection solutions, which were mixed as per the manufacturer's instructions (General Electric) and added directly to the membrane for 30 s in dark conditions. The membrane was then wrapped in cling film, and was immediately overlaid with an X-ray film for 30 s to 5 min, depending on signal strength following development of short-timed films. Film was developed in a film developer.

2.6 *In silico* analyses & statistical testing

2.6.1 RNA-seq *in silico* processing

RNA-seq data was generated by Dr. Emma Sherwood. She isolated total mRNA, in duplicate, from *S. venezuelae* *Δsihf* and wildtype strains at three time points (exponential vegetative growth; transition phase fragmentation; stationary phase and onset of sporulation). Data processing was carried out using the workflow indicated in Figure 2.1: Trimmomatic was used to trim reads (Bolger, 2014), FastQC was used to evaluate the quality of the trimmed reads before Bowtie2 (Langmead, 2013) mapped reads to the genome. Output .bam files were modified to .sam files for counting in HTseq (Anders, 2015). Read counts at each gene were analyzed using the DeSeq2 statistical package in R (Love, 2014). Appendix A1 contains the final BASH/R scripts used to process this data workflow. RNA reads were processed using servers provided by Dr. Brian Golding (McMaster University) and statistical analyses were performed in R using

the DeSeq2 statistical packages (version 1.11.9). Read trimming was performed by Trimmomatic version 0.25, mapping performed by bowtie2 version 2.2.2, and counting performed by HTseq version 0.6.1.

2.7 Biochemical Assays

2.7.1 Reporter assay of promoter activity by *gusA*

Streptomyces M1146 strains harboring pGUS vectors (*gusA* under P_{ErmE*}) were grown in biological triplicate, in 14 ml of YEME-TSB liquid medium, together with ~10 sterile glass beads to assist with cell dispersal. One millilitre of culture was harvested at the 24 hour 48 hour and 72 hour timepoints, the cells were pelleted, and washed in 1 ml of sterile water before being pelleted again and stored at -80°C. Frozen pellets were then thawed on ice, prior to the cells being resuspended in lysis buffer (50 mM NaH₂PO₄, 0.03% 2-mercaptoethanol, 0.1% Triton X-100, 4 mg/ml lysozyme) and incubated at 37°C for 1 h. To ensure appropriate mixing, the lysis suspensions were inverted every 10 min. Cell debris was then pelleted by centrifugation and the lysate supernatant was disposed. One hundred microliters of supernatant were added to a transparent 96 well plate in triplicate for each replicate. To this supernatant was added 66.6 µl of Z-buffer (150 mM NaH₂PO₄, 20 mM KCl, 5 mM MgCl₂, 0.03% 2-mercaptoethanol, pH 7) and 33.3 µl of 4 mg/ml 4-Nitrophenyl β-D-glucopyranoside (PNPG) were added. Absorbance measurements were collected at a wavelength of 420 nm in a microplate reader (Synergy) every 3 min for 4 h.

The average absorbance change was measured within the linear region of the assay for each replicate, and was normalized by protein content as measured by a Bradford assay (described above). The calculation for the arbitrary units is as follows:

$$AU = \frac{\text{slope}\left(\frac{\Delta A_{420nm}}{\text{time}(mins)}\right)}{\text{protein concentration}\left(\frac{mg}{ml}\right)}$$

Calculation of fold changes was made by dividing the arbitrary units of promoter segments with Crp-O/E by its corresponding promoter segment with the empty plasmid.

2.8 Wright Actinomycete Collection (WAC) Screening

Strains from the WAC collection (200) as well as those transformed with either empty plasmid or Crp-O/E were screened against MRSA and/or *P. aeruginosa*. A small aliquot of spores for each strain was grown in a spot on a Bennett's medium agar plate for four days at 30°C before an overlay of Bennett's agar inoculated with Methicillin-resistant *Staphylococcus aureus* (MRSA) or *Pseudomonas aeruginosa* was added. It was incubated for another 24 hours at 30°C and zones of inhibition measured with a ruler.

Table 2.1: Species used

<i>Streptomyces</i> species	Strain	Genotype/Usage	Reference
<i>S. venezuelae</i>	ATCC 10712	Wildtype	Gift from M. Bibb and M. Buttner
<i>S. coelicolor</i>	M145	SCP1 ⁻ SCP2 ⁻	-(Keiser et al., 2000)
<i>S. coelicolor</i>	M1146	$\Delta act, \Delta red, \Delta cpk, \Delta cda$	(Gomez-Escribano & Bibb, 2011)
<i>E. coli</i> species			
<i>E. coli</i>	DH5 α	Molecular cloning	Bethesda Research Laboratories
<i>E. coli</i>	ET 12567 pUZ8002	Source of <i>tra</i> genes and conjugation into <i>Streptomyces</i>	(MacNeil et al., 1992)

Table 2.2: Working antibiotic concentrations

Antibiotic	Working Culture Concentration	Conjugation Overlay Concentration (1 ml sterile dH₂O overlays)
Apramycin	50 µg/ml	1.25 mg/ml
Hygromycin	50 µg/ml	1.25 mg/ml
Ampicillin	100 µg/ml	2.5 mg/ml
Nalidixic Acid	25 µg/ml	0.625 mg/ml
Chloramphenicol	25 µg/ml	0.625 mg/ml
Kanamycin	50 µg/ml	1.25 mg/ml

Table 2.3: Media types

Media type	Nutrient conditions	Usage
<i>Soy flour mannitol (SFM)</i>	Good	Analyzing <i>Streptomyces</i> development/making spore stocks
<i>R5 medium</i>	Rich	Source of robust nutrient conditions
<i>Streptomyces antibiotic medium (SAM)</i>	Rich	Promoting antibiotic production
<i>Bennett's medium</i>	Rich	Growth of many species
<i>Minimal medium (MM)</i>	Poor	For evaluating supplements/carbon sources

Table 2.4: Plasmids used in this work

Project	Plasmid Name	Usage	Origin
6439 k/o CRISPR (knockout of predicted cluster regulator)	pCrispomycetes-2	CRISPR template vector designed for use in <i>Streptomyces</i>	(Cobb et al., 2015)
	pCrispomycetes-2 + EditSCO6439_gRNASCO6439	CRISPR vector for introducing frameshift into SCO6439	This work
pGUS 6429 GUS assays on 6429 promoter activity	pIJ10257: ErmE*	empty integrative vector control (ϕ BT1)	(Hong, Hutchings, Hill, & Buttner, 2005)
	pIJ10257 + ErmE*-Crp	integrative vector + Crp overexpression under ErmE*	internal
	pGUS	empty Gus reporter vector control (ϕ C31)	(Myronovskiy, Welle, Fedorenko, & Luzhetskyy, 2011)
	pGUS:ErmE*	Gus reporter vector with <i>gusA</i> under ErmE* promoter	Internal
	pGUS: Psc06429 1B-2B	Gus reporter vector with <i>gusA</i> under Psc06429 1B-2B region	This work
	pGUS: Psc06429 1B-2C	Gus reporter vector with <i>gusA</i> under Psc06429 1B-2C region	This work
	pGUS: Psc06429 1A-2B	Gus reporter vector with <i>gusA</i> under Psc06429 1A-2B region	This work
	pGUS: Psc06429 1A-2C	Gus reporter vector with <i>gusA</i> under Psc06429 1A-2C region	This work
SVEN sIHF FLAG-tagging constructs	pCrispomycetes-2	Crispr template vector designed for use in <i>Streptomyces</i>	(Cobb et al., 2015)
	pCrispomycetes-2 + EditSVEN1080-5XGly-3XFLAG_gRNASVEN1080	Crispr vector for introducing 3 \times Flag tag onto SVEN1080	This work
	pCrispomycetes-2 + EditSVEN1080-[SerGly ₃] ₄ -3XFLAG_gRNASVEN1080	Crispr vector for introducing 3 \times Flag tag onto SVEN1080	This work
SCO1480 mutant complement vectors	pSET152	Empty integrative vector	(Bierman et al., 1992)
	pSET152 + SCO1480	Integrative vector + SCO1480	This work
	pSET152 + SCO1480: NQ93SS	Integrative vector + SCO1480 mutant	This work
	pSET152 + SCO1480: RR85AS	Integrative vector + SCO1480 mutant	This work
	pSET152 + SCO1480: INSGly66	Integrative vector + SCO1480 mutant	This work
	pSET152 + SCO1480: Groove	Integrative vector + SCO1480 mutant	This work
	pSET152 + SCO1480: Δ 1-13	Integrative vector + SCO1480 mutant	This work
	pSET152 + SCO1480: Δ 1-36	Integrative vector + SCO1480 mutant	This work
WAC strain transformation	pIJ12551	Integrative empty negative control plasmid (ϕ C31)	(Sherwood, Hesketh, & Bibb, 2013)
	pIJ12551 + P _{ErmE*} -Crp	Integrative empty negative control plasmid with Crp-O/E(ϕ C31)	This work

Table 2.5: Oligonucleotides: Red sequences are hybrid sequences, RE sites underlined

Project	Name	Oligonucleotide Sequence (5'-3')	Usage
6439 k/o CRISPR (knockout of predicted cluster regulator)	NdeI-6439_For	CGACGT <u>CATATGGTTCCGTTCCGGGGATG</u>	SCO6439 clone screening
	6439-HinDIII_Rev	CGACG <u>AAGCTTTTCAGTCTCGACCCAGCCGTG</u>	SCO6439 clone screening
	gRNA-6439 For	ACGCTCCGGAACGCACCGACTCCT	CRISPR-mediated deletion of SCO6439
	gRNA-6439 Rev	AAACAGGAGTCGGTGCCTCCGGA	CRISPR-mediated deletion of SCO6439
	1kLF_6439_For	CATAGTCTAGAGACGATGTCCTGATGATACG	CRISPR-mediated deletion of SCO6439
	1kLF_6439_Rev	CGGGAATGAGCAGGTCGGTC	CRISPR-mediated deletion of SCO6439
	1kRF_6439_For	CGACCTGCTCAITCCCGCTGGTCGTGAATC	CRISPR-mediated deletion of SCO6439
1kRF_6439_Rev	GCACTTCTAGACTGCACATGCACATCGGCTC	CRISPR-mediated deletion of SCO6439	
pGUS 6429 GUS assays on 6429 promoter activity	pGUS_For	AAGCTTGCTCAATCAATCACC	pGus clone screening [multi-binding]
	pGUS_Rev	CGTCCAGCTTCTTGATCTCG	pGus clone screening
	Crp1-P6429_1A_For	CATAGTCTAGAGTATCGGCGGTCTCTGGATC	Psc6429 cloning into pGUS
	P6429_1B_For	TCTAGAAGTCCGATTGCGTCCGACCCGTC	Psc6429 cloning into pGUS
	P6429_2B_Rev	CAGATGGTACCGTCTGATGCCGCTTTCGCAT	Psc6429 cloning into pGUS
P6429-Crp2_2C_Rev	CAGATGGTACCCTGATTCCGTCGTGGACGCTC	Psc6429 cloning into pGUS	
SVEN sIHF FLAG-tagging constructs	SVEN_1080_1kL_For	CATAGTCTAGACCTTCTCGAGCGCTTCGGCTC	CRISPR-mediated in situ FLAG-tagging of SVEN1080
	SVEN_1080_1kL_Rev	CTCGAGGGCGCCGGTCTGCCGAATC	CRISPR-mediated in situ FLAG-tagging of SVEN1080
	SVEN_1080_1kR_For	TCGGCAGCACCGGCGCCCTCGAGCCCGGCGACAGGTCGCCGCTG	CRISPR-mediated in situ FLAG-tagging of SVEN1080
	SVEN_1080_1kR_Rev	CATAGTCTAGATCGAGCAGCTCGTCGATCGGC	CRISPR-mediated in situ FLAG-tagging of SVEN1080
	5xGly linker For	TCGACGGCGCGCGCGCGGCC	Linker for FLAG-tagging of SVEN1080
	5xGly linker Rev	TCGAGGCCGCGCGCGCGGCC	Linker for FLAG-tagging of SVEN1080
	[SerGly4]3 linker For	TCGACTCCGGCGCGCGGCTCCGGCGCGCGGCTCCGGCGCGCGCGGCC	Linker for FLAG-tagging of SVEN1080
	[SerGly4]3 linker Rev	TCGAGGCCGCGCGCGCGGAGCCGCCCGCGGAGCCGCCCGCGCGCGGAG	Linker for FLAG-tagging of SVEN1080
	FLAG-Forward (xhoI)	TCGAGGACTACAAGGACCACGACGGCGACTACAAGGACCACGACATCGACTACAAGGACGATGACGACAAGTGAC	Linker for FLAG-tagging of SVEN1080
	FLAG-Reverse (xhoI)	TCGAGTCACTGTCTGTCATCGTCTTGTAGTCTGATGTCGTTGGTCTTGTAGTCCCGCTCGTGGTCTTGTAGTCC	Linker for FLAG-tagging of SVEN1080
	1080 gRNA For	ACGCGCCGGGCGGCAGCGGGAC	CRISPR-mediated in situ FLAG-tagging of SVEN1080
	1080 gRNA Rev	AAACGTCCCGTCCCGCCCGGCG	CRISPR-mediated in situ FLAG-tagging of SVEN1080
Overlap_For	TCGGCAGCACCGGCGCCCTCGA	Clone check of pCrispomyces-2: SVEN1080 constructs	
SCO1480 sIHF check primers	NdeI_SCO1480_For	GGGTGCCATATGGCTCTCCGCCCTTACCC	Cloning of SCO1480
	SCO1480_Rev_BamHI	GGGTGCCGATCCTCAGCTGCCGGTGTGCC	Cloning of SCO1480
	Psc1480_For	CGTTGCCGACTGCTGCCGACCG	Check of SCO1480 mutant-complement conjugations
	Tsco1480_Rev	GATGCAGATCTATATGGGCGACGACCTGTCTC	Check of SCO1480 mutant-complement conjugations
	Δ1-36_For	CCTGAACAGCGCGCAG	Check of SCO1480 mutant-complement conjugations
General Primers	M13F	GTAAACGACGGCCAGT	Cloning checks of compatible vectors
	M13R	CAGGAAACAGCTATGAC	Cloning checks of compatible vectors
	T7_For	TAATACGACTCACTATAGGG	Cloning checks of compatible vectors
	T7_Rev	GCTAGTTATTGCTCAGCGG	Cloning checks of compatible vectors
	pMC600R	TTTGTGATGCTCGTCAGGGG	Cloning checks of compatible vectors

Table 2.6: PCR conditions

Taq polymerase			
Reagent	Final concentration		
Buffer	1×		
dNTP mix (10 mM)	0.2 mM		
DMSO (50%)	5%		
Primer Forward	0.3 μ M		
Primer Reverse	0.3 μ M		
Polymerase	0.125 U/ μ l		
Nuclease-Free water	To desired volume		
Protocol			
Step	Temperature ($^{\circ}$C)	Time	Cycles
Denaturation	95	3 min	1
Denaturation	95	30 s	×34
Annealing	T _m -5	30 s	
Extension	72	30 s/1 kb	
Last Extension	72	5 min	1
Phusion polymerase			
Reagent	Final concentration		
Buffer	1×		
dNTP mix (10 mM)	0.2 mM		
DMSO (50%)	5%		
Primer Forward	0.5 μ M		
Primer Reverse	0.5 μ M		
Polymerase	0.02 U/ μ L		
Nuclease-Free water	To desired volume		
Protocol			
Step	Temperature ($^{\circ}$C)	Time	Cycles
Denaturation	98	3 min	1
Denaturation	98	10 s	×30
Annealing	T _m -5	30 s	
Extension	72	15s/kb	
Last Extension	72	5 min	1

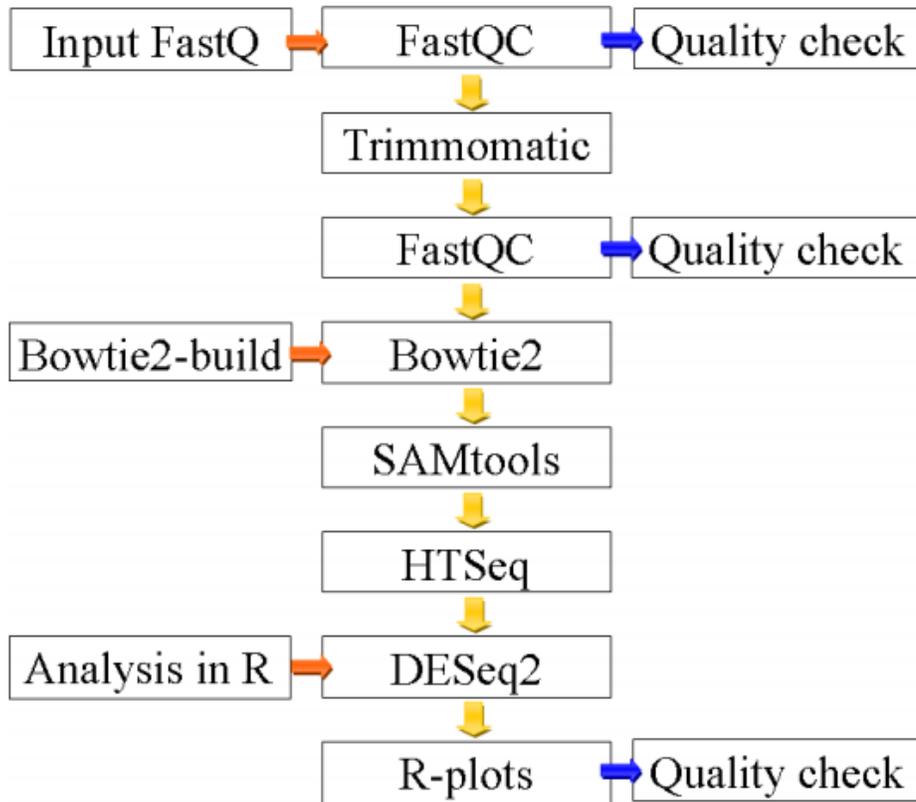


Figure 2.1: RNA-seq data processing workflow. FastQC evaluates the quality of raw reads, Trimmomatic trims reads to lengths of sufficient/high quality. Bowtie2 maps trimmed reads to *S. venezuelae* genome and SAMtools converts mappings to .sam/.bam files for counting. HTSeq counts the reads at each gene position in the genome. DESeq2 applies statistical packages to read profiles and evaluates changes between treatments. R was used to generate plots from DESeq2 processing to evaluate quality of treatment read relationships.

Chapter 3: Characterization of sIHF activity in *Streptomyces*

3.A ACKNOWLEDGEMENTS

Dr. Emma Sherwood, conducted the RNA extractions, provided RNA-sequencing read data, and performed EMSAs of sIHF proteins in complex with DNA.

3.1 INTRODUCTION

3.1.1 sIHF as a regulator of development and secondary metabolite production

sIHF localizes to the nucleoid in *S. coelicolor* and is produced constitutively throughout the *S. coelicolor* lifecycle (Swiercz et al., 2013). Deleting *sIHF* in *S. coelicolor* results in delayed development, aberrant antibiotic production, and elongated spores with diffuse nucleoids (Swiercz et al., 2013). In *S. venezuelae*, *sIHF* deletion results in aberrant development, enhanced melanin production, and diffuse nucleoids, though no discernible impact on spore length was observed (E. Sherwood; unpublished data). Of particular interest is the specific role of sIHF in organizing the bacterial nucleoid and how this activity impacts antibiotic production.

3.1.2 The structure of sIHF and its proposed DNA-binding affinity

X-ray crystallography of sIHF in complex with DNA oligonucleotides has revealed sIHF to comprise five consecutive alpha helices that are arranged similarly to domains within topoisomerase VI and the ribosomal protein S13 (Swiercz et al., 2013). This arrangement includes a canonical helix-2-turns-helix (H2TH) motif implicated in peripheral DNA binding across many protein families, and a large N-terminal helix ($\alpha 1$) that likely stabilizes the hydrophobic protein core and may utilize positively charged residues to interact with DNA (Nanji, 2014; Swiercz et al., 2013).

The first 13 residues at the N-terminus appear to adopt a flexible structure that could not be resolved in the proposed sIHF crystal model (Swiercz et al., 2013). These

residues are followed by a 21-residue alpha helix 1 ($\alpha 1$). $\alpha 1$ contains seven positively charged residues that face the core of the protein and form a positively charged cleft (Figure 3.1). Structural analysis by small-angle x-ray scattering (SAXS) has suggested that this may be a DNA binding pocket, where $\alpha 1$ wraps around the DNA, such that the positive moieties lining the cleft interact with the negative backbone of the DNA helix (Nanji et al., under revision). However, the crystal-packing structure reveals that sIHF-DNA interactions are possible at two additional interfaces: interface 1 (an inter helix loop that forms a lid over the hydrophobic core) and interface 2 (the H2TH motif) (Figure 3.1).

In vitro work has shown sIHF binds DNA non-specifically and it can inhibit TopA-mediated relaxation of plasmid DNA, where TopA encodes the only type I topoisomerase in *Streptomyces* (Swiercz et al., 2013). More recently, targeted mutations have been conducted to tease apart the relative contributions of the different binding interfaces, to the different functions ascribed to sIHF (Nanji, 2014).

3.1.3 Developing a framework to elucidate the function(s) of sIHF

Currently, we lack a specific list of genetic elements influenced by sIHF activity, and do not know whether sIHF interacts with any other proteins in the cell. Here, we work towards deciphering the effect of sIHF on global gene expression, determining the impact of sIHF mutations on *in vivo* activity, and preparing sIHF for future experiments to identify any protein binding partners and determine its chromosomal binding sites.

3.2 RESULTS

3.2.1 Transcriptome analyses reveal potential genes regulated by sIHF

The pathways through which sIHF affects wild type *Streptomyces* development and metabolism are unknown. It is proposed that its non-specific DNA binding and chromosome condensation activity may globally alter transcription, but there is as yet nothing known about which genes are influenced by sIHF, nor if its developmental effects are direct or indirect. To determine the impact of sIHF at a transcriptional level,

differences in mRNA transcript abundance between wild type *S. venezuelae* and a $\Delta sIHF$ mutant were analyzed following an RNA-sequencing (RNA-seq) experiment (Dr. Emma Sherwood, unpublished). To track the impact of sIHF through development during growth in liquid medium, RNA was harvested in duplicate at three growth-stages: early exponential phase (T1), mid-exponential phase (T2), and early stationary phase (T3). An analysis of the RNA-seq read data was conducted using multiple parameter sets, as the method by which RNA-seq data is processed can influence the output (Appendix A2). Initial processing of the data used the read mapping program eXpress and normalized reads before subjecting them to statistical tests (DESeq2 statistical package), producing a list of 360 differentially expressed genes ($p < 0.01$, $q < 0.01$). We anticipated this list to include genes that could explain the observed spore defects in *S. venezuelae* eg. genes that directly participate in spore separation (e.g. *ftsZ*) or those that regulate sporulation (e.g. *bldD*). Yet such genes did not appear to be differentially expressed in the $\Delta sIHF$ strain.

We employed an alternative processing method that minimizes false positives and circumvents any bias introduced by read normalization, as outlined in Figure 2.1. Sequenced reads of ~75 bp were processed through a trimming algorithm that clips low quality nucleotides (PHRED > 25) from the sequences to ensure a high degree of accuracy when mapping to the chromosome. Next an indexed database was created for the sequenced chromosome using Bowtie2. The reads were mapped to this database using HTSeq instead of eXpress, as HTSeq ignores multi-mapped reads to minimize false positives and optimize the data for differential expression analysis (Anders et al., 2015). The number of reads at each designated gene location was counted by HTSeq. A statistical suite, DESeq2 was then used to compile counts from replicates, and assess differential expression between the wild type and $\Delta sIHF$ strains.

A principle component analysis (PCA) plot and cluster analysis revealed that only the third timepoint (T3) had sufficient similarities between replicates for reliable comparisons to be made (Figure 3.2); any transcriptional changes observed at either T1 or T2 were less likely to be reproducible. Therefore, our analyses focussed on changes in gene expression occurring at the early stationary (T3) growth phase.

At this final timepoint, our analyses ultimately yielded a list of genes whose transcripts were influenced by sIHF, as determined by differences in expression ($p < 0.01$, $q < 0.1$) in the wild type strain compared with the *sIHF* mutant strain. Notably, the differentially expressed genes did not include any known regulatory or developmental genes, suggesting that the sIHF effect on development and antibiotic production may be through a novel mechanism.

We found there were 304 genes upregulated and 254 genes downregulated by at least 2-fold in the $\Delta sIHF$ strain, when compared with the wild type. The upregulated genes could be broadly grouped into four categories: glucose metabolism, amino acid metabolism, metal-ion transport, and fatty acid synthesis/metabolism (Figure 3.3). Of the downregulated genes, most were of unknown function, with 120 being annotated as encoding “hypothetical proteins” (Figure 3.3). While there was no obvious relationship between these upregulated and downregulated genes and the $\Delta sIHF$ mutant phenotype, these results implicated sIHF as a regulator of primary metabolic pathways in late stage development.

3.2.2 sIHF DNA-binding interface mutations

Given the many potential DNA binding interfaces associated with sIHF, it would be useful to know which was the major mediator of DNA association when considering both sIHF function, and our attempts to generate a functional sIHF fusion. To probe the DNA binding domains of sIHF, a series of targeted mutations were made to sIHF at the three sites suggested by the structural studies: interface 1, interface 2, and cleft (Nanji, 2014). The NQ93SS and RR85AS point mutations each removed two charged residues at interface 1 that had the potential to interact with the DNA backbone, though R86 is oriented such that its charged surface lies along the cleft interface (Figure 3.1, and (Nanji, 2014). Mutation of interface 2 involved inserting an additional glycine residue (alternatively referred to as Gly, +Gly or G66+) to disrupt the ordered association of the amide peptide backbone of this region with the negatively charged phosphates on the DNA (Nanji, 2014). To explore the contribution made by the N-terminal positively-

charged cleft, the positively charged residues were replaced with alanines in $\alpha 1$ to eliminate any association between this region and the DNA backbone (Nanji, 2014). Finally, a series of N-terminally truncated proteins were created, lacking either the first 13 ($\Delta 1-13$) amino acids (which were absent from the protein structure due to their presumed flexibility), or the first 36 amino acids ($\Delta 1-36$), which removed the entire $\alpha 1$ helix and associated cleft (Nanji, 2014). Electrophoretic mobility shift assay (EMSA) experiments revealed that both N-terminally truncated versions of sIHF, as well as the cleft mutant, lost their ability to bind DNA, while the NQ93SS, RR85AS, and INSG66 were able to bind, but showed reduced affinity for DNA (Figure 3.4).

Here, we evaluated the impact of these mutations on sIHF *in vivo* function, by introducing each *sIHF* mutant variant under the control of its native promoter, into a $\Delta sIHF$ strain of *S. coelicolor*. We further attempted to generate a functional FLAG-tagged sIHF variant using linkers to separate the FLAG epitope tag from the core protein, to use in ChIP-seq experiments to identify sIHF-associated sequences.

3.2.3 sIHF function *in vivo* is sensitive to mutations at the N-terminal and DNA binding interfaces

Each of the six mutant *sIHF* genes described above and summarized in Table 3.1 and Figure 3.1, was cloned into the integrative, conjugative vector pSET152 and introduced to *S. coelicolor* $\Delta sIHF$ by conjugation. Successful plasmid integration and the integrity of the *sIHF* alleles were verified by PCR amplification of each allele using a combination of plasmid-specific and gene-specific primers, followed by sequencing of the resulting PCR products (Mobix Laboratories). Additionally, antibodies raised against the native sIHF protein were used to detect the mutant proteins in cell lysates from each complemented strain (Figure 3.5). We could detect all the mutant sIHF proteins except for the $\Delta 1-36$ mutant, suggesting that this protein was not stable *in vivo*.

Given the $\Delta sIHF$ defects in sporulation (heterogeneous spore sizes), nucleoid compaction (more diffuse nucleoids in spores relative to wild type), and altered antibiotic

production, we were interested in testing the ability of each mutant protein to complement these specific characteristics.

Spore lengths for each mutant-complemented strain were measured from microscopy images in FIJI/ImageJ (>270 spores per strain), and were compared with both a plasmid-alone strain (negative control), and a strain complemented with wild type *sIHF* (Figure 3.6). The mean spore length increased from 1.1 μm in the wild type-complemented strain, to 1.2 μm in the $\Delta sIHF$ mutant (plasmid-alone) strain. Consistent with previous work, we also observed an increase in the variance of spore length in the $\Delta sIHF$ mutant from 0.05 to 0.11, consistent with an increased upper range for $\Delta sIHF$ of 1.41 μm compared to wild type at 1.21 μm (5 percentile bounds were at 0.98 μm and 0.97 μm respectively). When considering the different mutant *sIHF* variants, the RR85AS (lid/interface 1) mutant appeared to function as wild type, complementing the $\Delta sIHF$ spore length, with both wild type median spore length and minimal variation in spore length. Interestingly, however, NQ93SS, also at interface 1, did not complement spore lengths. The +Gly mutant at interface 2 exhibited an intermediate median spore length, but the variance remained similar to that of the $\Delta sIHF$ strain. Interestingly, all mutations to the N-terminus (cleft mutant, $\Delta 1-13$ and $\Delta 1-36$) led to variably sized spores most similar to those of the $\Delta sIHF$ strain, suggesting an inability to complement the mutant phenotype, at least for the cleft mutant and $\Delta 1-13$ strain (the $\Delta 1-36$ protein was not stable, and thus would have been entirely unable to complement).

Nucleoid area measurements from DAPI-stained spores were calculated using Fiji-ImageJ and revealed enhanced chromosome diffusion in the $\Delta sIHF$ (plasmid-alone) control strain relative to that complemented with wild type *sIHF* (Figure 3.7). Consistent with our observations for spore length, the +Gly and RR85AS mutants exhibited wild type-like nucleoid areas (compact), although notably, the RR85AS had lower variance than the wild type-complemented strain. The NQ93SS showed greater variance in nucleoid compaction, but a median similar to that of wild type. Neither the N-terminal deletions nor the cleft mutant complemented the $\Delta sIHF$ nucleoid-area phenotype.

Finally, actinorhodin production for each of these strains was monitored by blue pigment secretion after 5 days of growth on SFM agar (Figure 3.8). The wild type-complemented strain produced copious amounts of blue pigment, which was also observed for the +Gly, RR85AS, and NQ93SS mutants, but not the cleft or N-terminal deletions (Table 3.1).

3.3 DISCUSSION

3.3.1 Differentially expressed genes from *sIHF* knockout

The RNA-seq data, processed by two separate analyses did not yield any obvious candidates that could be responsible for the spore differentiation or antibiotic defects associated with the *sIHF* mutation in *S. coelicolor*. Only a few uncharacterized regulators from the GntR, TetR, and ArsR families were identified as being differentially regulated in these analyses, and their corresponding transcripts were generally lowly abundant. Instead, most of the differentially expressed genes identified were predicted to contribute to some form of primary metabolism. As primary metabolites serve as the precursors for secondary metabolite formation, it is possible that the observed changes in antibiotic production could be attributable to *sIHF* altering primary metabolism.

Interestingly, we could only be confident in our observations at T3 given that the cluster plots show low correspondence between replicates at T1 and T2. As T3 corresponds to entry into sporulation, it is conceivable that *sIHF*-mediated chromosome compaction during sporulation prevents further metabolic activity, and helps to promote metabolic downregulation and the ensuing spore dormancy. These changes at T3 could be due to direct regulatory control by *sIHF*, or indirect control through intermediate partners or the action of other regulators.

To differentiate between direct and indirect regulation by *sIHF*, we had previously attempted to introduce a FLAG epitope tag onto *sIHF* to facilitate co-immunoprecipitation with any protein binding partners. This would have allowed us to perform ChIP-seq experiments to determine if there were any specific loci/sequences

bound by sIHF. However, neither N-terminal nor C-terminal FLAG fusions to sIHF were able to complement the null mutant phenotype. This has made it a challenge to probe the role of sIHF in regulating gene expression in *Streptomyces*, although it does suggest that both the N- and C- termini of sIHF may be critical for protein function and cannot be occluded by the immediate fusion of a FLAG tag. This would be consistent with the known structure of sIHF, as the first 13 amino acids are necessary for binding to DNA (Nanji, et al., under revision).

3.3.2 *In vivo* effects of targeted sIHF mutations

Collectively, the phenotypic characteristics observed for each of the different mutant proteins indicate that sIHF has functions that are structurally separable. The strains expressing the N-terminally truncated sIHF variants and the cleft mutant did not complement the spore compartmentalization defect, the diffuse nucleoids, or the lack of actinorhodin production by the $\Delta sIHF$ strain. Evidently, such perturbations at the N-terminus completely ablate sIHF activity.

The NQ93SS interface 1 mutant complemented actinorhodin production but showed null mutant spore lengths and a varied distribution of nucleoid areas, though the median was similar to wild type. The nearby RR85AS interface 1 mutant completely complemented the spore size and diffuse nucleoid phenotypes of the $\Delta sIHF$ strain. This suggests that the NQ93 residues at interface 1 contributed to spore size, and possibly nucleoid compaction. The capacity of NQ93SS to complement actinorhodin production presents two possibilities: 1) sIHF regulates spore and nucleoid size through interface 1, while actinorhodin production is controlled by a distinct region of the protein, or, 2) the NQ93SS mutant alters the malleability of the sIHF structure such that its ability to associate with DNA at other interfaces (most likely the N-terminal region) are influenced to different degrees: it meets the activity threshold to restore actinorhodin production but not the threshold to restore spore size to those seen in wild type.

At interface 2, the +Gly mutant complementation restored nucleoid compaction but did not fully restore spore length. As described previously, this interface was targeted

due to a predicted interaction between the peptide backbone and an adjacent DNA strand in the crystal. This proposed interaction with DNA was somewhat corroborated by EMSAs showing reduced DNA binding affinity by the +Gly mutant protein. However, the fact that this variant could restore wild type DNA compaction and actinorhodin production suggested that this weakened affinity was insufficient to inhibit *in vivo* functionality. Such an interaction would be consistent with the peripheral role for H2TH motifs in DNA/RNA binding by other proteins (Gilboa et al., 2002).

Overall, these data indicate that the DNA binding cleft suggested by small-angle x-ray scattering SAXS, within the N-terminal region of sIHF seems to be the most important DNA binding determinant for this protein, as replacing the positive charges within $\alpha 1$ or truncating the N-terminal and $\alpha 1$ region resulted in variants that failed to complement any of the observed $\Delta sIHF$ phenotypes *in vivo*. Interestingly, the $\Delta 1-13$ mutant eliminated sIHF activity, even though these residues did not appear in the original crystal structure. Interface 1 is not particularly sensitive to perturbation, as the mutations of RR85 did not influence sIHF activity *in vivo*, although they did compromise the DNA binding ability *in vitro*. The NQ93 residues, however, appeared to influence sIHF activity, as discussed above. Interface 2 did not appear to affect DNA compaction, but this region was implicated in the control of sporulation, based on the phenotypes associated with the +Gly addition.

An explanation for these phenotypes may be that sIHF binds DNA through the SAXS-predicted N-terminal cleft, while interfaces 1 and 2 provide stabilizing contacts with DNA. It is also possible that sIHF acts in association with as-yet-to-be-determined partner proteins, and these interact with sIHF through interfaces 1 and/or 2, or their interactions with sIHF are allosterically altered by mutations in interface 1 or 2.

sIHF appears to bind DNA in a manner that is unique among NAPs. All other known NAPs bind DNA in a multimeric manner: H-NS dimerizes across tracks of DNA, bridging the DNA and polymerizing along it, while HU proteins impart bends in DNA through dimerization. Yet sIHF does not appear to multimerize, instead associating with DNA as an abundant monomer (Swiercz et al., 2013). Interestingly, the monomeric state

of sIHF may be specific for the *Streptomyces* protein, as its orthologue in *Mycobacterium*, mIHF has been reported to exist as a dimer in solution (Mishra et al., 2013).

3.4 FUTURE INVESTIGATIONS

Here, we establish characteristics of the sIHF protein that are critical to fully understanding its role in developmental and metabolic regulation. Further research into sIHF should have three specific aims: 1) establish a means of isolating functional sIHF bound to DNA and any protein binding partners such that they can be identified, 2) determine how sIHF promotes chromosome condensation, and 3) exploit these interactions to modify secondary metabolite production.

3.4.1 Flag-tagging sIHF for ChIP-seq and Co-immunoprecipitation

The primary activity of sIHF appears to be mediated through its DNA-binding activity. It is therefore important that we distinguish between those genes that are directly transcriptionally controlled by sIHF and those that are indirectly controlled. We sought to create a functional tagged (3×FLAG) sIHF protein in *S. venezuelae*, to permit ChIP-seq for cross-referencing with the RNA-seq dataset. Given the apparent DNA-binding function of the N-terminus, we selected the C-terminus for FLAG-tag fusion, in the hopes that this would maximize the likelihood of generating a functional protein. However, previous fusions directly at the C-terminus had failed to complement the $\Delta sIHF$ phenotype *in vivo* (E. Sherwood, personal communication). In an attempt to circumvent this issue, we opted to generate two different sIHF-3×FLAG-encoding variants under control of the native *sIHF* promoter. In contrast to previous attempts, where the FLAG tag was fused directly to the sIHF protein, here, we designed it such that it was separated by different linkers: a shorter 5×Gly repeat and a longer [SerGly₄]₃ linker. The longer linker has been successfully used in fusions to other small DNA-binding proteins (Bush et al., 2016). A restriction enzyme site was introduced at the C-terminus of *sIHF* lacking the stop codon, maintaining 1 kb upstream and downstream of the *sIHF* coding region. The linker-3×FLAG segments were introduced into this restriction site, and were

confirmed by sequencing. To date, conjugation of these *sIHF3*×*FLAG* constructs into the *S. venezuelae* Δ *sIHF* strain have been unsuccessful. Previous work showed that the immediate promoter-*sIHF*-terminator region of *S. venezuelae* was not sufficient to complement the Δ *sIHF* phenotype. However, a larger construct, encompassing an additional 4 kb upstream and downstream of *sIHF*, was sufficient to complement the Δ *sIHF* phenotype (E. Sherwood, personal communication). The reason for this phenomenon is still unknown, however, it is conceivable that sIHF regulates its own expression by binding to sites within the flanking 4 kb regions and introduces specific DNA topology changes that affects its own locus. An integrative plasmid carrying the *sIHF* gene may not present a genetic environment similar enough to the chromosome to permit wild type expression of the *sIHF* gene. As such, it is possible that 1 kb upstream and downstream of *sIHF* may be insufficient for complementation, as was the case for the wild type *sIHF*. It is worth noting, however, that our RNA-seq data do not show altered transcript levels for *sIHF* in wild type and mutant strains. Alternatively, it remains possible that even in the presence of additional flanking sequence, the sIHF-3×*FLAG* protein will be non-functional.

We are continuing to screen for successful exconjugants to complement the Δ *sIHF* strain, however, our cloning strategy also permits us to introduce the *sIHF-linker-3*×*FLAG* genes *in situ* using CRISPR-Cas9. This technique reconstitutes the chromosome into exactly the same sequence as wild type (with the exception of the small linker-3×*FLAG* introduced at the C-terminus of *sIHF*), ensuring that expression from the *sIHF* promoter is in its native genetic regulatory context.

The DNA-binding sequence-specificity of sIHF is unknown. It appears to bind DNA oligomers promiscuously, although a systematic evolution of ligands by exponential enrichment (SELEX) experiment revealed a preference for AT-rich sequences with a narrow minor groove (Swiercz., 2013). Chromatin immunoprecipitation-sequencing (ChIP-seq) performed on the FLAG-tagged sIHF protein would allow us to isolate segments of DNA bound by sIHF. If these segments align to specific sites on the chromosome, it would suggest that sIHF directly influences

gene expression by repressing or activating transcription at defined sites. However, there is a greater likelihood that we detect variable sequences with broad coverage of the genome, given the apparent non-specificity of sIHF.

3.4.2 Proteomic studies into sIHF and manipulation of its activity

Work to date does not exclude the possibility that sIHF has a protein binding partner(s). Co-immunoprecipitation of the sIHF-FLAG protein using anti-FLAG antibodies would permit not only the identification of any associated DNA sequences, it would also allow us to isolate any proteins that interact with sIHF. We have begun work to test for sIHF interaction with other proteins using a bacterial two-hybrid system in which sIHF is bound at the N and C terminus to one component of a protein and a library of DNA segments from the *S. venezuelae* chromosome are bound to the other component of a protein. If sIHF interacts with any protein encoded from the DNA segments in the library, then the two fused protein components will be brought together to activate a reporter gene. To date we have not detected any positive hits using this system, however it is complicated by the known inability of sIHF to function with fusions at both its N and C terminus. Given the multiple interfaces of sIHF that appear to interact with DNA in some capacity, it is somewhat unlikely that it participates in binding interactions with other proteins, and these investigations are not our top priority.

It has been reported though that sIHF may be the target of acetylase activity, as lysine acetylation was detected at position 13 in the sIHF sequence from a proteomic analysis (Liao, Xie, Li, Cheng, & Xie, 2014). Interestingly, this acetylation lies at the base of the 13-amino acid N-terminal loop that cannot be resolved by x-ray crystallography (Swiercz et al., 2013) but is critical for the DNA binding of sIHF. This bears similarities to the N-terminal tails located on the H3 and H4 histones that comprise nucleosomes in eukaryotes, as histone tail modifications are known to act as switches that can alter the state of the nucleosomes to grant DNA access for transcription (euchromatin) or occlude DNA to prevent transcription (heterochromatin) (Grunstein, 1997). Probing this interaction may require identification of the acting acetylase and

gauging its impact on sIHF binding activity *in vitro* through EMSAs of acetylase-treated and untreated sIHF samples.

A greater understanding of sIHF will allow us to exploit the regulatory behaviour of sIHF to enhance antibiotic production. As deleting sIHF leads to altered antibiotic production, it would be interesting to determine the effects of *sIHF* deletion when combined with the deletion of other NAP proteins. The H-NS-like protein Lsr2 has been identified in *Streptomyces* as a repressor protein that binds in the vicinity of biosynthetic clusters, and whose deletion results in increased secretion of chloramphenicol in *S. venezuelae* (E. Sherwood, in preparation). It is likely that Lsr2 affects the local DNA topology, and so there is the opportunity for synergistic or antagonistic activity with sIHF.

Further, we could consider developing strategies to readily inhibit sIHF activity across a range of *Streptomyces*. As streptomycetes differ in their respective suites of biosynthetic clusters, any strategy to enhance antibiotic production would benefit from a system that is tractable in most streptomycetes. Modulating sIHF activity could involve one of several approaches: introducing highly-expressed post-translational modifiers of sIHF (if any are confirmed); over-expressing sIHF itself; or knocking down sIHF using small inhibitory RNAs to bind the *sIHF* mRNA and ablate its translation. Manipulating sIHF activity has the potential to provide access to the latent reservoir of *Streptomyces* bioactive secondary metabolites.

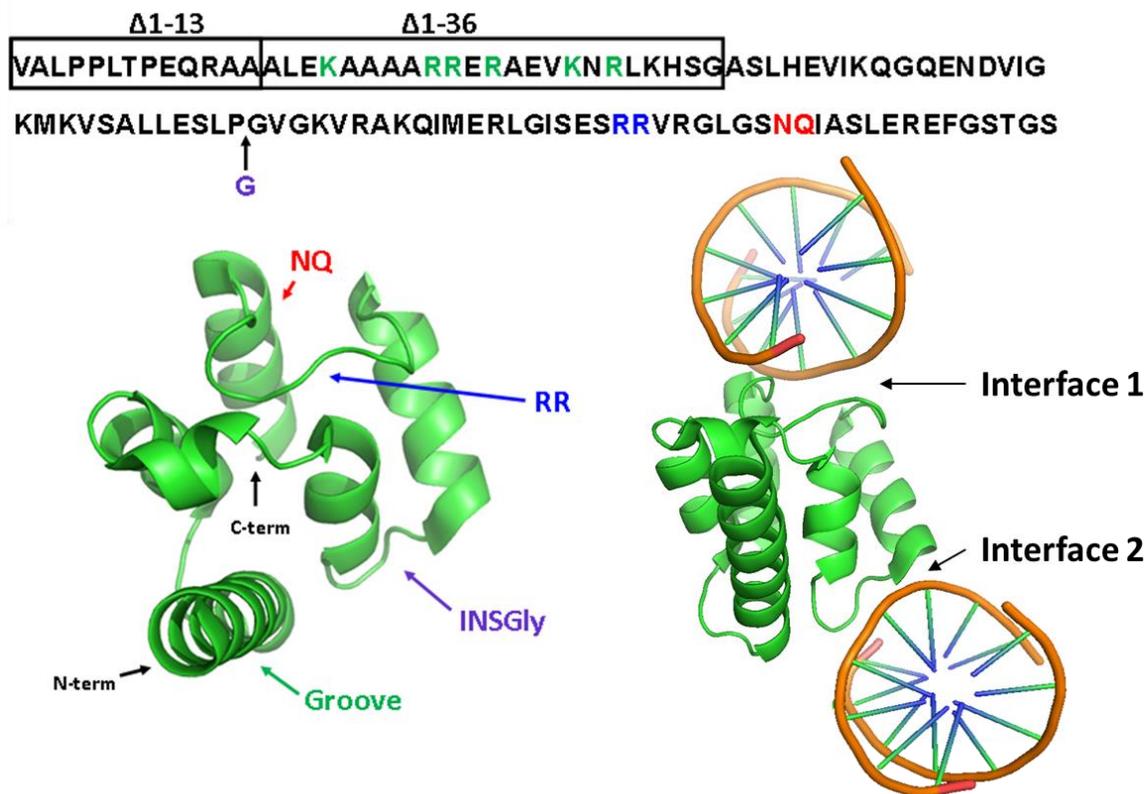


Figure 3.1: *sIHF* sequence and crystal structure. *S. coelicolor* *sIHF* crystallized in the presence of DNA oligonucleotides (Swiercz et al., 2013). Targeted mutations to probe *sIHF* activity are indicated by coloured arrows, with the associated amino acids that were changed for each mutant construct indicated in the same colours. The ‘Cleft’ mutations altered the positive charges along the N-terminal helix, while the ‘INSGly’ mutation involved introducing an additional glycine residue into the loop connecting helix 3 and helix 4.

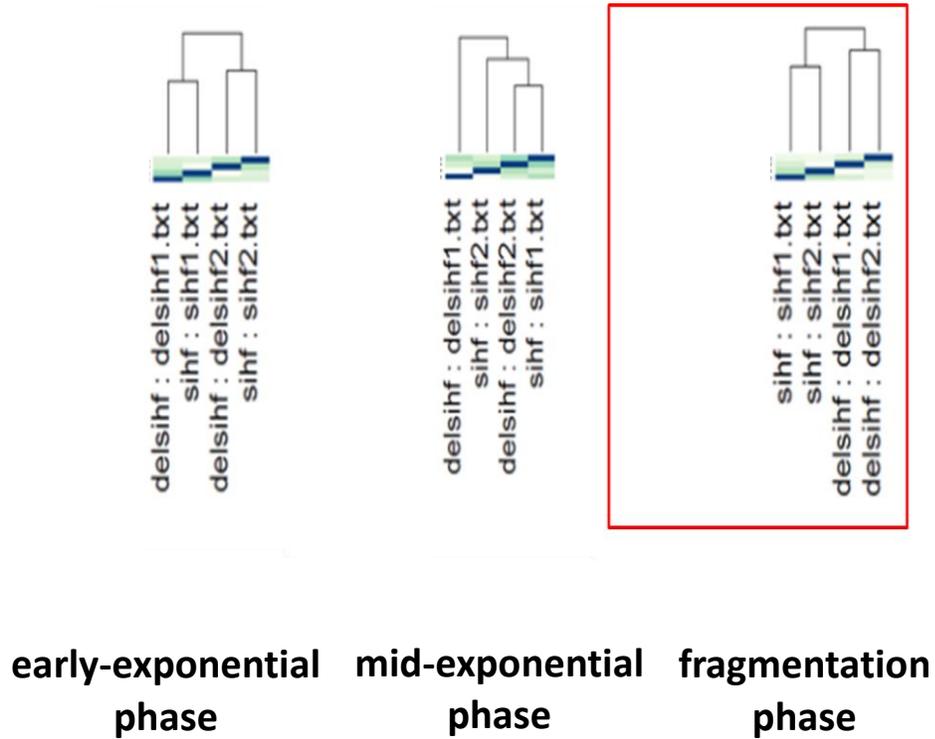


Figure 3.2: Cluster plots of RNA-seq data from RNA extracted at three timepoints, from two independent biological replicates. *S. venezuelae* processed RNA-seq timepoints sorted by cluster analysis in R, to evaluate the similarity of each replicate (1 or 2) relative to each sample [*wildtype sIHF* (*sihf*) vs Δ *sIHF* (*delsihf*)]. Replicates of the same sample were expected to sort together. The red box was used to highlight the data set (from the fragmentation phase of development) that was the most reliable, based on similarities between individual replicates.

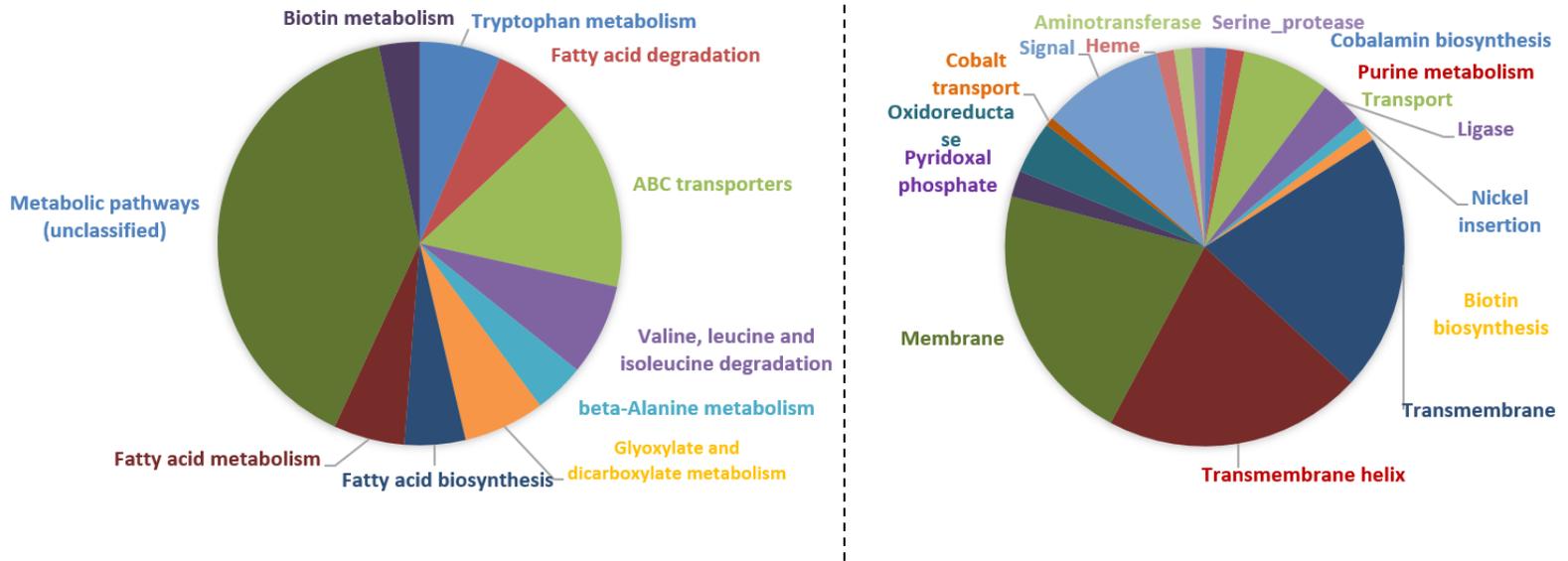


Figure 3.3: Distribution of encoded gene classifications that were upregulated/downregulated by *sIHF* knockout in *S. venezuelae*. Genes altered by $\Delta sIHF$ were grouped by KEGG pathway and Universal Protein (UP) category with the assistance of DAVID Huang, 2008). Left: genes upregulated in the *sIHF* mutant. Right: genes downregulated in the *sIHF* mutant.

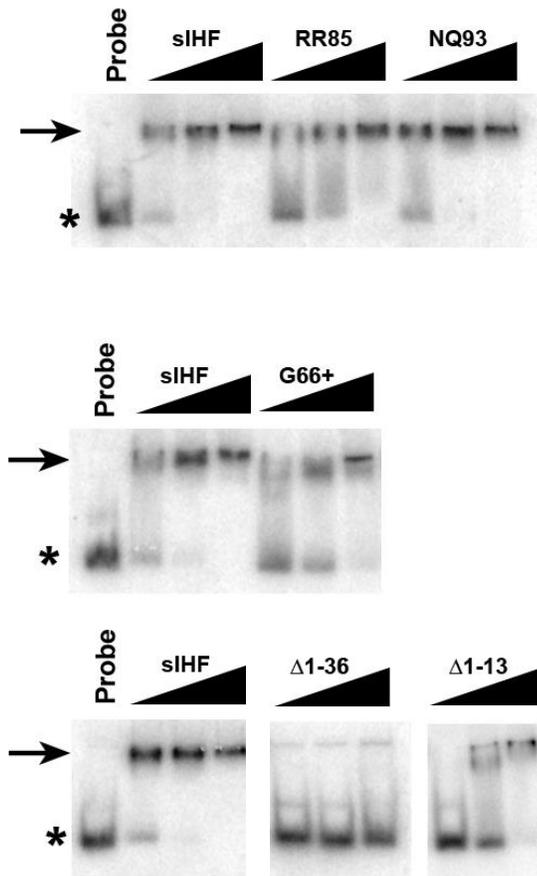


Figure 3.4: EMSAs using mutant sIHF proteins. Labelled DNA oligomers alone, or in association with increasing concentrations of sIHF protein variants, were separated on a polyacrylamide gel. * denotes the presence of the DNA oligomers alone without sIHF and the arrow indicates the shifted position of sIHF in complex with the DNA oligomers.

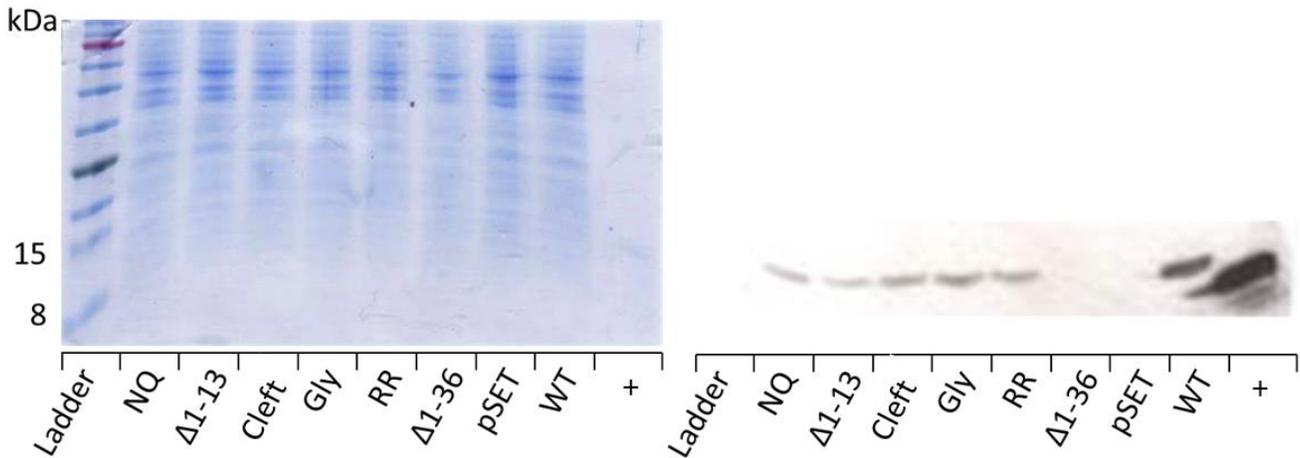


Figure 3.5: Western blotting reveals most siHF mutant variants are stable *in vivo*. *S. coelicolor* Δ siHF strains transformed with wild type and mutant siHF variants were grown in liquid culture. The cells were collected after 48 h of growth, lysed by sonication and soluble proteins were collected by centrifugation. Soluble proteins were separated on a 15% SDS-PAGE, and either stained with Coomassie brilliant blue (left), or transferred to a PVDF membrane and subjected to western blotting using an siHF-specific polyclonal antibody (right). Extracts corresponding to the different mutant variants are as indicated below each panel. NQ and RR: amino acids that were altered. Δ 1-13 and Δ 1-36: N-terminal deletion mutants. Cleft: mutant in which positively charged amino acid residues in the N-terminal cleft were mutated. Gly: extra amino acid introduced. pSET: plasmid-alone control. WT: complemented with wild type siHF. +: purified His-tagged siHF. No signal was detected for Δ 1-36 under various conditions across multiple lysis/western trials.

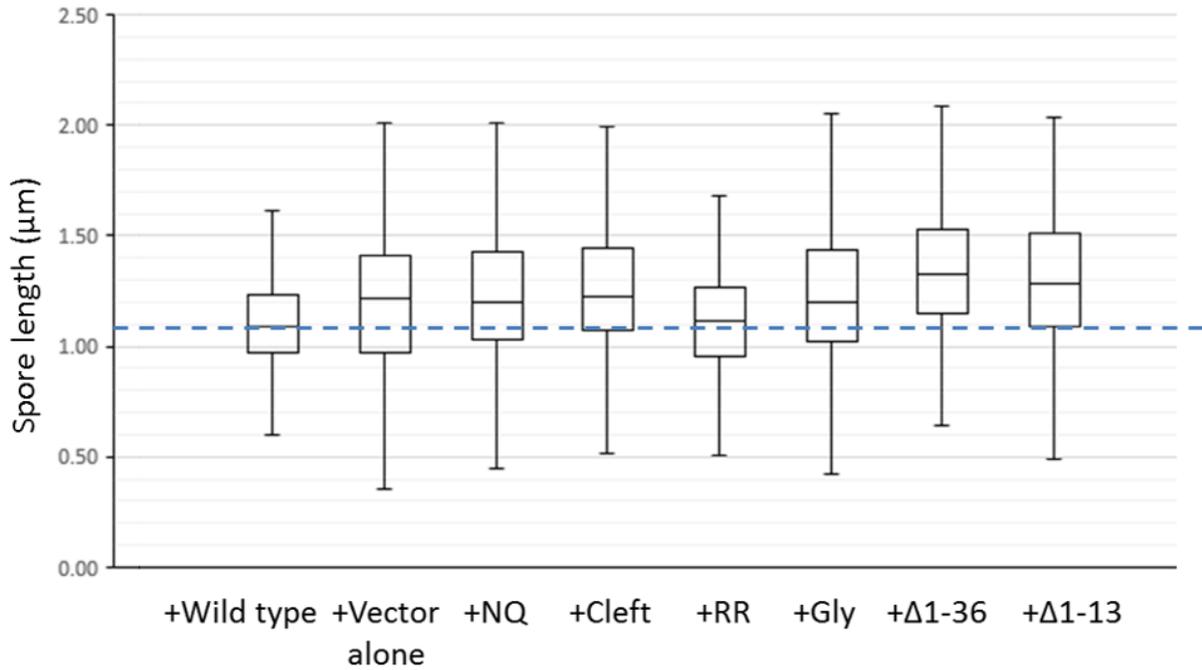


Figure 3.6: Spore length measures of the $\Delta sIHF$ strain complemented with wild type and mutant *sIHF* variants. Light microscopy images were processed in FIJI/ImageJ using the measure tool. For each strain, the lengths of >300 spores were measured. Boxes represent first to third quartile, with middle line showing median (second quartile). Whiskers account for 5th to 95th percentile. The blue dashed line represents the wild type median value.

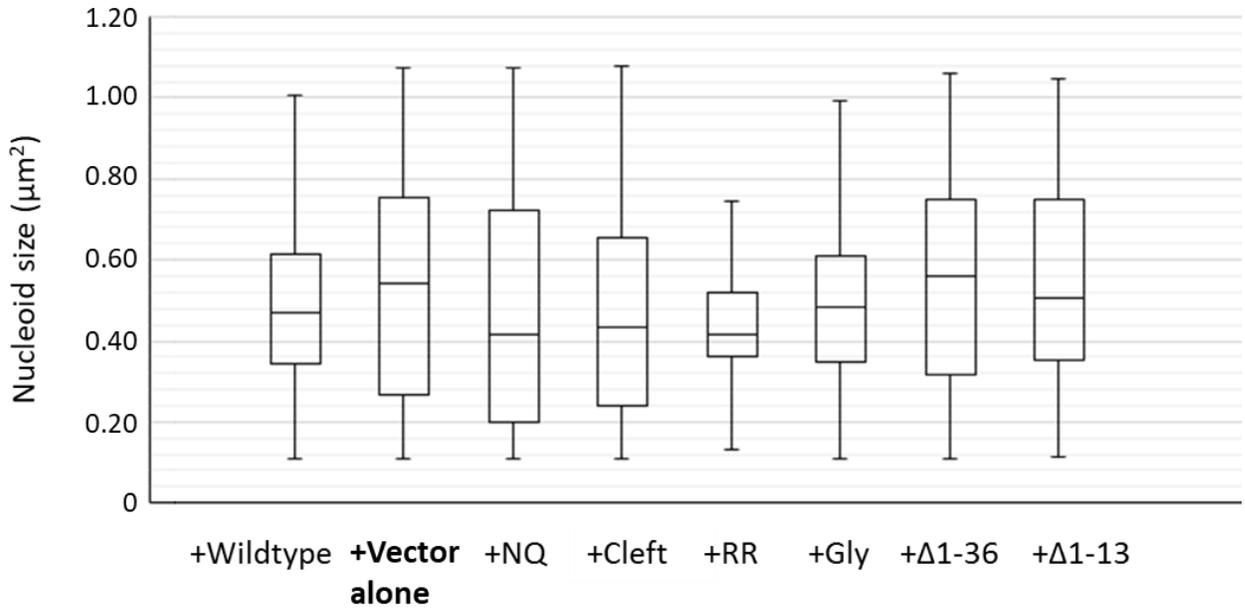


Figure 3.7: Nucleoid area measures of *sIHF* mutant complemented strains. DAPI microscopy images were processed in FIJI/ImageJ using the area threshold parsing, with >275 spore nucleoids being measured per strain. Boxes represent first to third quartile, with the middle line showing median (second quartile). Whiskers account for 5th to 95th percentile.

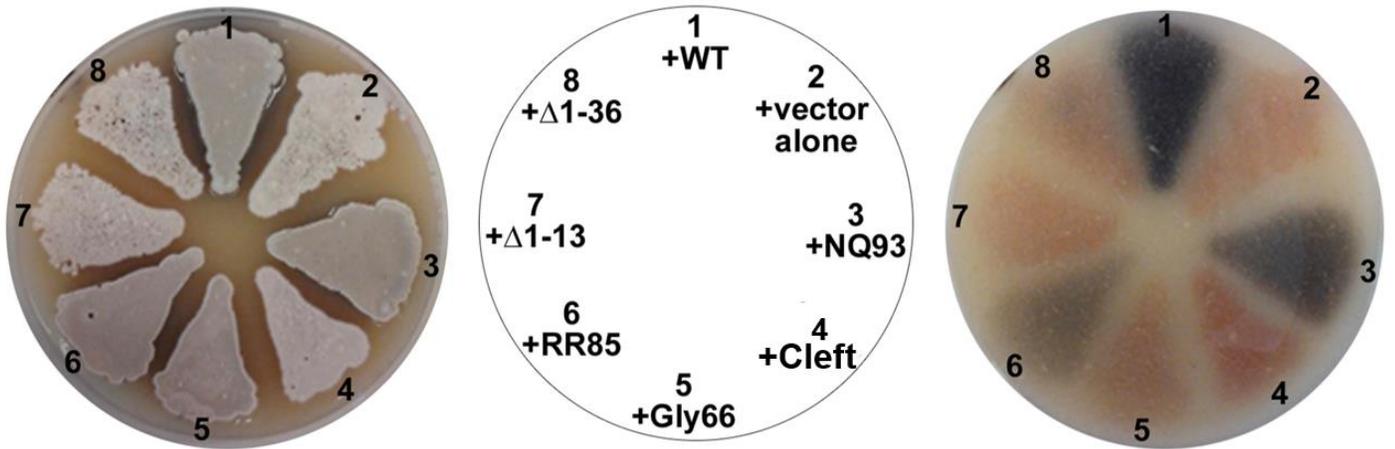


Figure 3.8: Sporulation and antibiotic production defects associated with the different *sIHF* mutant variants. Each *S. coelicolor sIHF* mutant strain was grown on SFM agar medium for 4 days at 30°C.

Table 3.1: sIHF mutant complementation phenotypes. $\Delta sIHF$ + mutant *sIHF* genes were assessed for sporulation and actinorhodin (Act) production after growth for 96 h on solid growth medium. Spore length and nucleoid size were assessed by brightfield and fluorescent microscopy, respectively.

Mutation	Sporulation	Act	Spore Length*	Nucleoid Size*
WT	+	+	n/a	n/a
$\Delta sIHF$	-	-	More varied/increased	Increased
$\Delta 1-13$	-	-	More varied/increased	Increased
$\Delta 1-36$	-	-	More varied/increased	Increased
RR85AS	+ (delayed)	+	WT	WT
NQ93SS	+ (delayed)	+	More varied/increased	Increased
Cleft	-	-	More varied/increased	Increased
+Gly66	+ (delayed)	+	WT	WT

*Relative to WT

Chapter 4: Regulation of the silent & cryptic NRPS cluster *SCO6429-SCO6438* in *S. coelicolor*

4.A ACKNOWLEDGEMENTS

I would like to thank Dr. Erin Carlson and Andy Johnson for their work in identifying the cryptic metabolite discussed here as well as their insights into NRPS cluster activity.

4.1 INTRODUCTION

4.1.1 Activation of specific silent/cryptic antibiotic clusters in *Streptomyces*

Global regulators provide us with an opportunity to influence gene expression across the chromosome; however, transcriptional control of individual genetic elements can be more complex than a simple on-off switch, with multi-level control being employed to fine-tune gene expression (*e.g.* St-Onge et al., 2015).

Many biosynthetic clusters employ pathway-specific regulators that are encoded within the cluster they control (Rutledge & Challis, 2015; Takano et al., 2005). For example, the clavulanic acid biosynthetic cluster of *Streptomyces clavuligeris* is activated by the downstream *claR* gene product (Pérez-Redondo et al., 1998), while the jadomycin B cluster in *S. venezuelae* is activated by the upstream *jadR1*-encoded regulator, whose activity is antagonized by the cluster-localized JadR2 repressor (Yang et al., 1995; Zhang, et al., 2013). Enhancing antibiotic production from such clusters, especially those that appear to be transcriptionally silent and whose products are unknown (or cryptic), requires an understanding of the regulators that control these operons.

Previous work in our lab has established that Crp has binding sites in the proximity of eight of the 22 predicted cryptic secondary metabolite clusters in *S. coelicolor*, with binding consensus sequences following a GTG(N)₆GNCAC motif (Gao et al., 2012). In fact, induction of *crp* expression in a Δcrp background resulted in activation of numerous clusters, including those directing the production of actinorhodin and the yellow cryptic polyketide (yCPK), as well as a cryptic biosynthetic operon: *SCO6429-SCO6438* (Gao et al., 2012). However, attempts to detect the product from this

cryptic cluster have been unsuccessful, due in part, perhaps to low product formation in a background full of other complex metabolites (Dr. Erin Carlson, personal communication).

4.1.2 Regulation of the silent *SCO6429-6439* biosynthetic cluster in *S. coelicolor*

The *SCO6429-SCO6439* operon (Figure 4.1) has syntenic homologues in the very closely related species *Streptomyces lividans*, but is not found in other streptomycetes. This cluster is predicted to encode two non-ribosomal peptide synthetases (NRPS, *SCO6431* and *SCO6432*) as well as a suite of tailoring enzymes (*SCO6433-SCO6438*) and a possible secretion channel (*SCO6430*). The presence of the NRPS genes within this cluster has led to the prediction that the operon yields a dipeptide product though it has not been detected in *S. coelicolor* extracts, and RNA-seq data suggest that the cluster is not transcriptionally active during growth on solid medium (Figure 4.2).

There is a predicted Crp consensus binding site within the first 200 nucleotides of the *SCO6429* coding sequence (Figures 4.3 and 4.4). This site was previously identified in chromatin immunoprecipitation-DNA microarray (ChIP chip) experiments and subsequent microarray data demonstrated that an increase in Crp expression activated *SCO6429-6431,6433-6434,6436* (Gao et al., 2012). This presents a unique case: Crp is known as both an activator and a repressor, depending upon where it binds within promoter regions and whether it can interact with a corepressor (Gerlach et al., 1991; Kolb et al., 1993); however, there is nothing known about how it might exert its regulatory effects when binding *within* a gene, as is the case here. Crp binding sites tend to be located no farther than a few hundred nucleotides upstream of its cognate promoter and if bound to a corepressor, typically prohibits open complex formation during transcriptional activation (Kolb et al., 1993). The silent nature of the *SCO6429* operon and the unusual binding site for Crp within this cluster prompted us to investigate the ability of Crp to affect *SCO6429* expression, and the regulation of the predicted dipeptide cluster as a whole.

Given the cryptic nature of the *SCO6429-39* cluster, and its unusual Crp-associated regulation, we sought to probe the regulatory interplay governing the expression of this metabolic cluster. Our aims here were threefold: to determine (1) the native transcriptional expression from $P_{SCO6429}$; (2) the regulatory impact of Crp binding on cluster expression; and (3) the effect of other regulators on *SCO6429* transcription.

4.2 RESULTS

4.2.1 Utilizing a reporter system to assay promoter activity

To evaluate the role of Crp on transcriptional activation at $P_{SCO6429}$, we decided to employ a colorimetric promoter-reporter fusion system in a transgenic strain of *S. coelicolor* that is deficient in all pigmented secondary metabolite production (Δact , Δred , Δcpk , Δcda) (Gomez-Escribano & Bibb, 2011).

We opted to use a β -glucuronidase enzyme reporter system, in which segments of the promoter for *SCO6429* ($P_{SCO6429}$) were cloned upstream of the *Streptomyces*-optimized β -glucuronidase (*gusA*) gene (Myronovskyi et al., 2011). If the promoter segment is transcriptionally active in *S. coelicolor*, then β -glucuronidase will be produced, yielding a blue pigment when exposed to 5-bromo-4-chloro-3-indolyl glucuronide (X-gluc), or a yellow pigment when exposed to 4-Nitrophenyl β -D-glucopyranoside (PNPG). Measuring absorbance values for each of these compounds correlates with transcriptional activity for that promoter segment. Each promoter segment-reporter strain was also further transformed with either an additional empty integrative plasmid or a plasmid harboring a *crp* overexpression construct.

We had anticipated seeing activation of the $P_{SCO6429}$ segment containing the intragenic Crp site in the Crp overexpression strain, compared with the control strain. However, no obvious difference was detectable between these strains. As we had previously observed upregulation of this cluster following induction of Crp (Gao et al., 2012), we were prompted to re-evaluate our understanding of this cluster.

4.2.2 Detection of a second Crp binding site upstream of the *SCO6429-SCO6438* cluster

While examining the nucleotide sequence upstream of *SCO6429*, we identified a second possible Crp binding site located 324 bp upstream of the *SCO6429* start codon. Interestingly, this site consisted of a GTGA-(N)₁₂-TCAC motif, which was similar to the aforementioned GTG-(N)₆-GNCAC Crp consensus motif, only with different spacing between the palindromic repeats (Figure 4.3). As this site was positioned in close proximity to *SCO6429*, we hypothesized that this putative Crp binding site may also influence transcription at this locus. Henceforth, this upstream Crp binding site will be referred to as C1, while the *SCO6429* intragenic site will be referred to as C2.

There was no evidence from the previously conducted ChIP-chip experiment that Crp binds to our predicted C1 site; however, ChIP-chip is known to be somewhat limited in resolution and may miss biologically-relevant binding sites (Park, 2009). The similarity of this upstream sequence to known Crp binding sites was rather striking, however, and this prompted us to assess whether it had a role in regulating *SCO6429* transcription.

4.2.3 P_{SCO6429} reporters were revealed undetectable activity during growth on solid media

To test the relative contribution made by each putative Crp binding site, four different promoter regions were generated, including constructs carrying the C1 site, the C2 site, both sites, and neither site (promoter alone). Each of the resulting products were cloned upstream of *gusA*, prior to being introduced into *S. coelicolor* M1146 (Figure 4.4).

The initial screening for *SCO6429* promoter activity was performed by growing each M1146-pGus:P_{SCO6429} construct on X-gluc-containing agar plates of various media compositions. A range of rich media types (R5, SFM, Bennett's, SAM) was used to ensure that metabolic pathways could be saturated with the appropriate precursor metabolites. None of the selected metabolite-promoting media resulted in obvious blue

pigment formation, suggesting that these different growth conditions were insufficient to activate $P_{SCO6429}$ expression.

However, we could not exclude the possibility that $P_{SCO6429}$ was transcribed at a level below the detectable limit for the *gusA* reporter system during growth on agar plates. Indeed, tight control of $P_{SCO6429}$ or weak transcription may prevent the native end-product from reaching production levels that are energetically costly or inhibitory to cell development during growth on agar plates. To explore this possibility, we sought a more specific and quantitative analysis of β -glucuronidase activity directed by the different $P_{SCO6429}$ segments.

4.2.4 Upstream and intragenic binding sites for Crp influence $P_{SCO6429}$ activity

A quantitative analysis of $P_{SCO6429}$ activity was achieved by growing *S. coelicolor* M1146 harboring the pGUS: $P_{SCO6429}$ constructs in a YEME-TSB liquid medium and collecting whole cell lysates. These lysates were exposed to the β -glucuronidase substrate PNPG and the absorbance of light at 420 nm was measured for each sample over 4 h. Reaction rates were proportional to the quantity of enzyme within the lysate samples but were also dependent on the degree of cell lysis achieved. The measured rates for each $P_{SCO6429}$ promoter construct were normalized to the protein concentration of the lysates so that we could compare the relative enzyme activity for each sample. In addition, P_{ermE}^* , a strong, constitutive promoter in *Streptomyces* M1146 was used as a positive control for *gusA* transcription, while the equivalent empty vector (pGUS)-containing strain served as a negative control.

We harvested samples from these growing strains at 24 h (early exponential), 32 h (mid-exponential), 42 h (late exponential/early stationary), and 72 h (mid-stationary). Transcriptional activity for $P_{SCO6429}$ was only detectable at 72 h, suggesting that the native activation of transcription was under particularly dense cell growth and nutrient-deprived conditions.

At the 72 h timepoint, promoter activity at $P_{SCO6429}$ was low relative to that of P_{ermE^*} , but was still above the background observed for the empty vector controls. Notably, C1- $P_{SCO6429}$ showed a decrease in promoter activity relative to $P_{SCO6429}$ alone.

We found the inclusion of the intragenic C2 site within the reporter fusion also resulted in a modest decrease of reporter activity, while the addition of the upstream C1 site to either the promoter alone, or in association with the downstream C2 site resulted in complete transcriptional inactivation (Figure 4.5A). The promoter activity with the C1- $P_{SCO6429}$ promoter was even lower than the basal transcription from the promoterless *gusA* gene (Figure 4.5A). This suggested that the presence of the C1 site may repress *SCO6429* transcription, and that the C2 site may fine-tune $P_{SCO6429}$ transcription under native Crp background conditions.

We next sought to investigate the role of Crp binding in the control of *SCO6429* expression by introducing Crp overexpression constructs into our reporter strains, in a bid to increase Crp activity. We then repeated the assay for β -glucuronidase expression. Unexpectedly, we observed a 3-fold increase in normalized reaction rates in both the negative control (promoterless *gusA*) and positive control (strong P_{ermE^*} -*gusA*) strains (Figure 4.5B,C). A similar increase of 2.6-fold was also observed for the $P_{SCO6429}$ alone sample when Crp was overexpressed (Figure 4.5B,C). This increase in expression across these three controls may be consistent with the increase in metabolism expected to result from Crp-overexpression.

Interestingly, the promoter regions coupled to either the C1 or C2 sites in the Crp overexpression background exhibited inverse activity levels relative to those in the native background. Overexpressing Crp in strains containing the C1- $P_{SCO6429}$ reporter resulted in enhanced reporter activity. In contrast, the strain carrying the $P_{SCO6429}$ - C2 reporter and the Crp overexpression construct showed little detectable expression (Figure 4.5B). Overexpressing Crp in the strain bearing both Crp binding sites (C1 - $P_{SCO6429}$ - C2) yielded expression levels intermediate to those of C1 - $P_{SCO6429}$ and $P_{SCO6429}$ - C2, suggesting that abundant Crp permitted transcription in the presence of C1, but that this

was tempered by some level of transcription attenuation associated with the presence of the C2 site.

This was particularly notable when examining the fold changes in expression in the wild type background compared with the Crp overexpressing strain. As previously mentioned, overexpressing Crp resulted in a 3-fold increase in expression from the $P_{SCO6429-gus}$ strain and P_{ermE^*-gus} control strain (Figure 4.5C). A 98-fold increase in expression occurred with C1 - $P_{SCO6429}$ when Crp was overexpressed, relative to empty vector, while a 120-fold decrease was observed with $P_{SCO6429}$ - C2, again, relative to empty vector (Figure 4.5C). When both putative Crp binding sites were present, there was a 50-fold increase in reporter activity when Crp was overexpressed, relative to the empty vector-containing strain (Figure 4.5C).

4.2.5 *SCO6439* is a CSD-containing regulator with a potential role in controlling *SCO6429-38* expression.

To bring in an additional level of complexity, we considered the possibility that *SCO6429*, and its associated cluster, may also be controlled by the product of *SCO6439*, which is found immediately downstream of the dipeptide biosynthetic cluster (Figure 4.1). *SCO6439* was annotated as a putative DNA binding protein-encoded gene with its amino acid sequence suggesting the presence of an N-terminal domain with homology to DNA or RNA binding cold shock proteins. Cold shock proteins in eukaryotes have variable auxiliary domains following the cold shock domain (CSD), such as the Y-box that directs the CSD to a specific target; however, in bacteria, most homologues possess only a single CSD, without additional auxiliary domains. Interestingly, *SCO6439* contains some homology to the Y-box domain of eukaryotic cold shock proteins (Chaikam & Karlson, 2010; Lasonderet al., 2002). *SCO6439* is effectively transcribed under laboratory growth conditions, as suggested by RNA-seq data from wild type *S. coelicolor*, and thus may have the potential to act as a pathway-specific repressor for the *SCO6429* operon (Figure 4.2).

To determine the effect of *SCO6439* on $P_{SCO6429}$ expression, we sought to delete the gene from the chromosome. We opted to use CRISPR-Cas9 to delete a 22 bp segment within the N-terminally annotated DNA-binding region (Figure 4.6). This deletion would also lead to a frameshift mutation in *SCO6439*. Interestingly, we were unable to generate this deletion in the wildtype *S. coelicolor* M145 background, although we successfully deleted this gene in the M1146 strain.

To assess the effect of the *SCO6439* mutation on the expression of the *SCO6429-38* operon, each of the reporter constructs from $P_{SCO6429}$ assays, as well as the construct harboring the Crp O/E plasmid or its empty negative control, were introduced into the $\Delta SCO6439$ strain by conjugation.

Unexpectedly, we observed no promoter activity for any strain apart from the P_{ermE^*} positive control. As the loss of a functional *SCO6439* gene appeared to lead to a complete abolition of promoter activity for $P_{SCO6429}$, it would seem that contrary to our expectations, *SCO6439* may act as an activator, rather than a repressor. We also tested the reporter activity of these different strains on plates in wild type and *SCO6439* mutant strains, after exposure to a range of temperatures (4°C, 12°C, or -20°C), given the cold-shock domain associated with *SCO6439*. We found there was no reporter activity detected after exposure to these different cold shocks.

4.3 DISCUSSION

4.3.1. *SCO6429-SCO6438* operon is regulated by Crp

Overall, our results suggested that transcription from $P_{SCO6429}$ was most rigorous in the presence of abundant Crp and the C1 binding site, while it was attenuated somewhat by the addition of the intragenic C2 site. While we do not have direct evidence for Crp binding at C1, we know that it is capable of binding C2, based on our ChIP-chip analyses (Gao et al., 2012). It is possible that Crp is binding to a corepressor at the C2 site to repress *SCO6429* repression. A similar mechanism has been shown to affect the *cdd* promoter in *E. coli*, where two tandem Crp binding sites function as antagonistic activator/repressors depending on the association of Crp with these sites when in

association with the CytR corepressor (Gerlach et al., 1991). What appears to be multi-level control of the *SCO6429* operon by Crp suggests that sensitive control of this operon must be necessary for *S. coelicolor* viability and/or fitness.

It is worth noting that the predicted C1 and C2 Crp binding sites were not identical in sequence (Figure 4.3), and it is possible that Crp may have a higher affinity for one of the two sites. It is also conceivable that alternative growth conditions may yield lower levels of intracellular cAMP, which may in turn limit the availability of high-affinity cAMP-Crp and consequently skew transcription towards a C1-on or C2-off state (Süsstrunk et al., 1998). Additionally, cAMP-Crp bound at C1 may affect the transcription of the divergently transcribed *SCO6428*.

4.3.2 *SCO6439* as a putative activator of *SCO6429-SCO6438* expression

We considered the possibility that additional factors may be influencing the expression of *SCO6429*, and tested the regulatory potential of the downstream-encoded *SCO6439* regulator. Deleting the gene failed to activate expression of any of our reporter genes, consistent with a role for this protein as a transcriptional activator for this cluster. Typically, CSD-containing proteins like *SCO6439* are involved in the cold-shock response and act as activators (Barria et al., 2016; Chaikam & Karlson, 2010). These regulators comprise the family of CspA-like proteins and bind to DNA and in some cases, RNA. It is thought that CspA, the founding member of this family, recognizes a CCAAT-motif within or upstream of its target promoter, with binding increasing transcription (Nakashima et al., 1996). This family of regulators can also exert post-transcriptional control by binding to mRNAs and stabilizing them in cold temperatures (Rennella et al., 2017). CspA-like proteins also have extended 5' UTRs in their own mRNA, which can act as a thermosensor, whereby the mRNA is only stable in cold conditions where the 5' UTR adopts a structured configuration that effectively occludes an RNase cleavage site (Barria et al., 2016).

In contrast, *SCO6439* does not have an extended 5' UTR, and in fact, transcription appears to start very near to the coding sequence based on RNA-seq data (M. Moody,

personal communication). If the gene and its product do respond to cold stress, then they would have to do so through some means other than an associated thermosensor.

4.4 FUTURE DIRECTIONS

Here we have preliminary results that suggest Crp plays a role in regulating a silent and cryptic biosynthetic cluster. There is significant progress to be made in understanding its regulation such that we can activate cluster expression within *S. coelicolor*, and characterize its associated product.

A major unknown at this point is whether the product itself is toxic to *Streptomyces*. Endogenous toxicity would limit how successful we could be in activating cluster expression and activity, and may require us to introduce a conditional promoter *in situ* for further study of its regulation and determination of its product.

In the short term, work must be done to confirm our observations that Crp overexpression affects transcription of the *SCO6429-6438* operon. This can be accomplished by making mutations to the Crp binding sites C1 and C2, and conducting our reporter assays in the presence and absence of Crp overexpression, to reveal whether the C1 and C2 motifs are specifically necessary for Crp-mediated control of operon gene expression to observe the reported fold changes in the presence of abundant Crp. We could complement this work by performing the assays in a Δcrp strain and test whether expression is altered by the presence/absence of the C1 and C2 sites. It would be challenging to use mobility shift assays to confirm Crp binding at these sites due to the unique isoelectric point of the *Streptomyces* Crp protein (pI of 5.8) that prohibits traditional shifts.

Our next steps could also include generating reporter strains harboring a *SCO6439* overexpression construct to further test its function as an activator. Its overexpression may drive higher reporter activity from the *SCO6429* promoter, and follow-up combinatorial reporter assays with Crp overexpression could elucidate how these proteins, both individually and collectively, tune the transcription of *SCO6429*. It will also be important to validate these studies with RT-qPCR targeting *SCO6429* in

background strains in the presence and absence of Crp overexpression and/or *SCO6439* overexpression. These assays would provide us with insight into regulatory interplay between Crp and *SCO6439* in the control of this cryptic metabolic cluster.

If we are successful at activating the cluster, we will endeavour to identify and characterize the peptide product and assay it for activity against pathogens, as other dipeptides produced by *Streptomyces* and *Bacillus* species are known to have antimicrobial properties (Boeck, Higgs, & Kastner, n.d.; E. Katz & Demain, 1977).

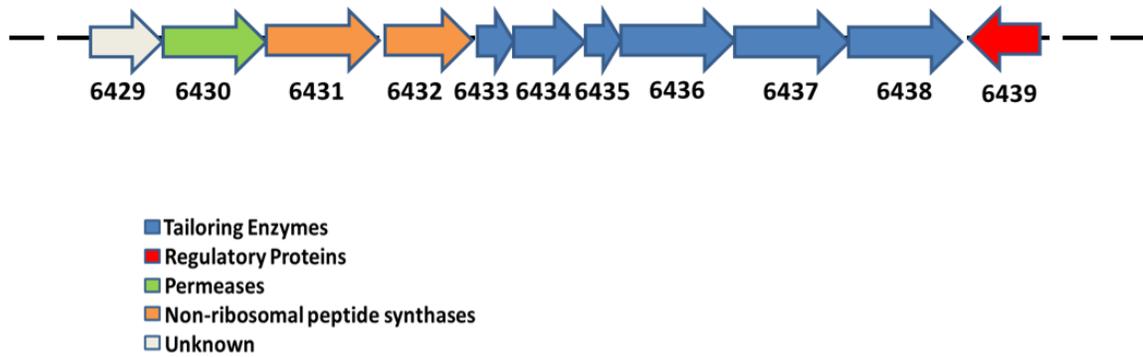


Figure 4.1: *SCO6429* operon encoding a predicted dipeptide. The eleven gene cluster from *S. coelicolor* is predicted to encode a dipeptide compound. Gene numbers are indicated below each gene, and gene colours correspond to *in silico* predicted gene activity as annotated in StrepDB (strepdb.streptomyces.org.uk/).

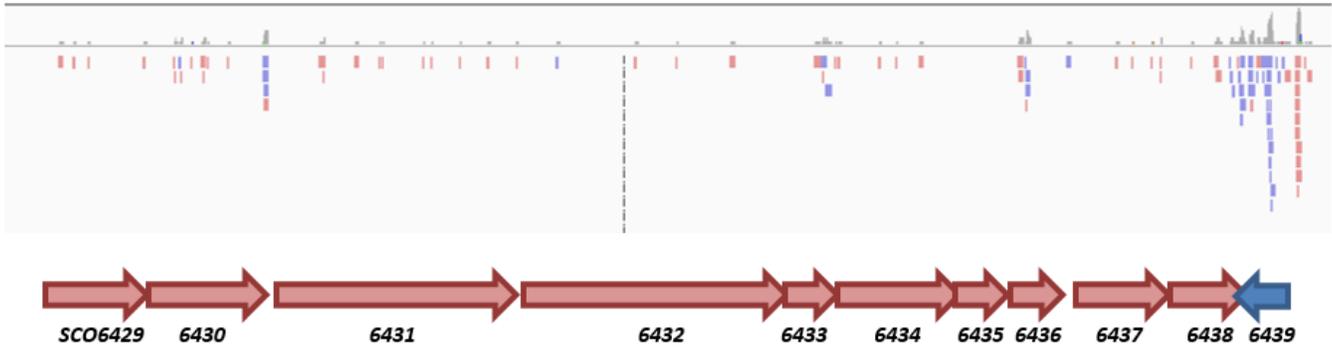


Figure 4.2: RNA-seq reads covering *SCO6429-6439*. RNA-seq data obtained from pooled timepoints from RNA extracted from *S. coelicolor*. Blue represents genes/reads oriented in the reverse direction, while pink genes/reads indicate genes going in the forward direction.

<u><i>nuoA</i></u>	<u>GTGAAAATGTCAC</u>
<u><i>cdaR</i></u>	<u>GCGGAGGCACTCAC</u>
<i>accA2</i>	<u>GTGTGGGCAAGCTCAC</u>
C1	<u>GTGAACGCTCCGCCATCAC</u>
C2	<u>GTGAAGGCGCCCTG</u>

Figure 4.3: Crp binding sites C1 and C2 consensus comparison. All sequences have been detected from experimental Crp ChIP-chip microarray data with the exception of C1. Underlined nucleotides matched consensus sequence from previous experiments (Gao et al., 2012)

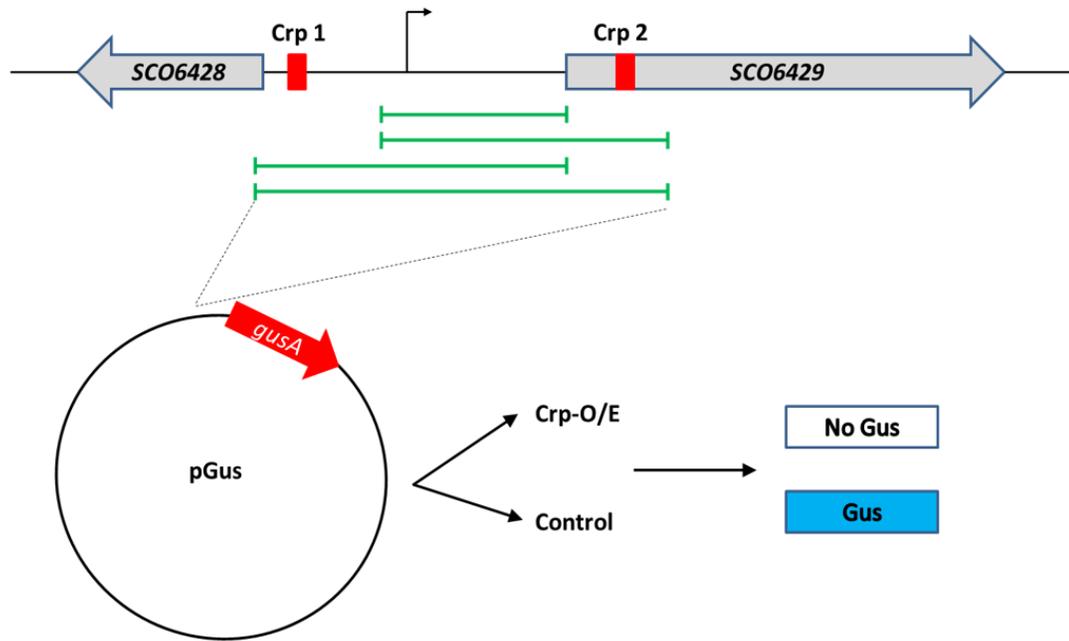


Figure 4.4: Summary of the different reporter fusions created to monitor the effect of two Crp binding sites on *SCO6429* expression. *S. coelicolor* M1146 overexpressing Crp (and its corresponding control strain) was transformed with the integrative vector pGus. pGus harbored the *gusA* gene, whose expression was driven by segments of the *SCO6429* promoter region (shown in green). Strains carrying promoter regions capable of directing the transcription of *gusA* will produce β -glucuronidase, which can cleave the substrates X-gluc or PNPG to yield a blue or yellow pigment, respectively. The pigment level was then used as a proxy for promoter strength.

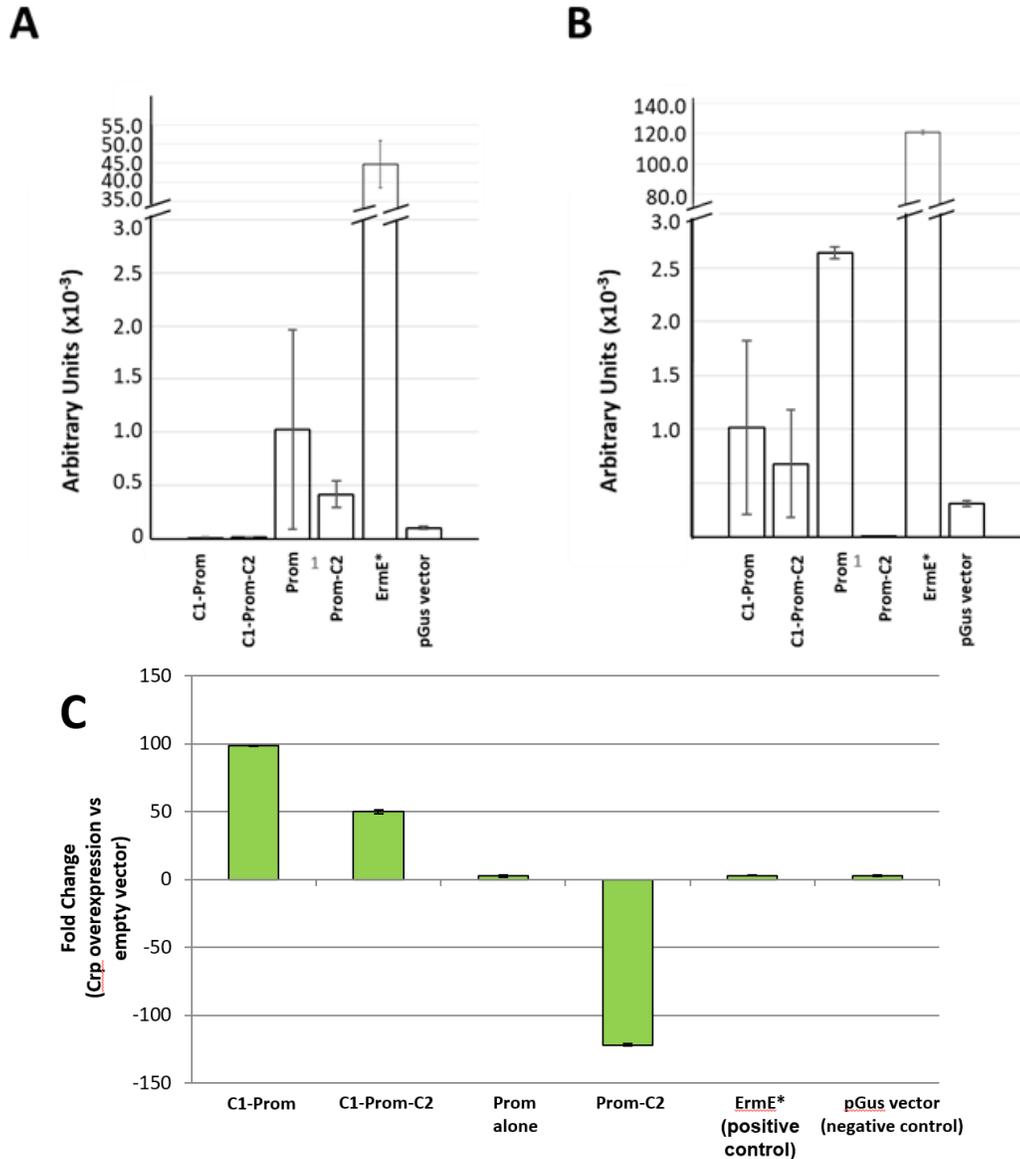


Figure 4.5: Reporter assays assessing the transcriptional activity of different promoter fragments (from Fig. 4.3) of *SCO6429*. Absorbance measurements at 405 nm were used to measure PNPG cleavage/ β -glucuronidase production over 4 h at 5 min intervals and linear segments of absorbance rate used to calculate ΔA_{405nm} . Error bars indicate standard error. **A)** Average ΔA_{405nm} for each promoter segment without Crp overexpression (+ empty vector). **B)** Average ΔA_{405nm} for each promoter segment with Crp overexpression. **C)** Fold changes comparing expression levels of empty vector (control) and Crp overexpression strains.

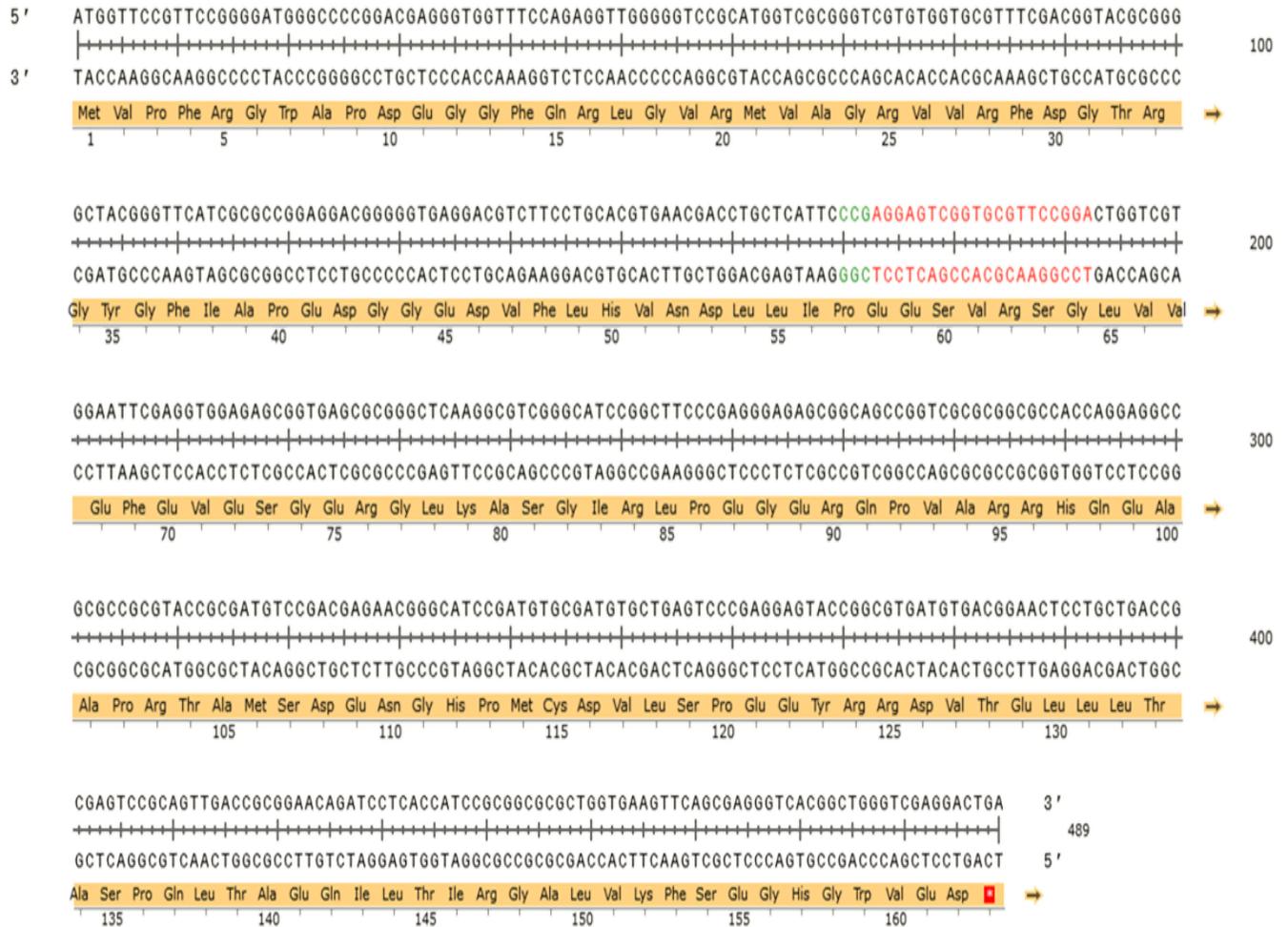


Figure 4.6: CRISPR-directed mutational strategy for *SCO6439*. *S. coelicolor* *SCO6439* gene targeted for deletion/frameshift mutation by CRISPR. Red nucleotides represent the 22 bp deleted region (which would also shift the frame of all sequences downstream), while green nucleotides mark the protospacer adjacent motif (PAM), necessary for promoting Cas9 cleavage.

Chapter 5: Induction of antibiotic production in a library of actinomycetes

5.A ACKNOWLEDGEMENTS

I would like to thank Dr. Emma Sherwood for performing the high-throughput conjugations discussed in this section. Further, I would like to thank David Crisante and Nicole Revie for their assistance in carrying out conjugations to create full sets of negative control/test transgenic WAC strains and screening them for antibiotic activity. I would also like to thank Dr. Gerry Wright for providing a portion of his lab's actinomycete collection for these experiments.

5.1 INTRODUCTION

5.1.1 Antibiotic screenings are limited by metabolome coverage

Classic strategies for identifying novel pharmaceuticals have involved taking cell extracts or purified chemical compounds and evaluating their ability to inhibit the growth of bacterial pathogens (Baltz, 2008). Attempts are then made to purify the compound, determine its structure, and analyze it biochemically to identify its cellular target(s). Some groups have established libraries of actinobacterial species (including *Streptomyces* and *Micromonospora*) known to be productive sources of bioactive molecules (Baltz, 2008). These libraries include a wide-range of species and strain isolates, comprising an abundance of genetic and metabolic diversity (Brown & Wright, 2016). High throughput strategies involve culturing these strains in micro-fermentation broths, extracting metabolites (using *e.g.* methanol or ethyl acetate solvents), and screening for bioactivity against pathogens by measuring inhibition of indicator organisms (Brown & Wright, 2016; Kinashi et al., 1988).

However, this strategy is constrained by the metabolites produced during growth on/in the selected culture medium. While different culture medias have been used to stimulate antibiotic production, these screenings are plagued by a high rate of rediscovery (repeatedly identifying the same antibiotic), and an abundance of metabolic clusters that

remain silent and/or cryptic (Elibol, 2004; Ramadhar et al., 2013). It is inferred that these silent clusters are under particularly specific regulatory control, where the conditions for activating transcription have not been met, or where the end products are in too low abundance to exhibit bioactivity. ‘Mining’ of these silent and cryptic clusters can involve a number of different approaches, including predicting their products through bioinformatics and attempting to activate their expression in a heterologous host (Scherlach & Hertweck, 2009).

5.1.2 Development of a “genetic toolkit” for enhancing antibiotic yield from existing *Streptomyces* libraries

Our third aim was to leverage our collective knowledge of *Streptomyces* regulation to develop a strategy that could increase antibiotic production in a library of streptomycetes. We opted for an integrative plasmid vector that could deliver a global antibiotic regulatory gene under the control of a strong promoter. Conventional vectors for introducing genetic elements into *Streptomyces* are stably integrated into the chromosome through the use of distinct – and broadly-conserved - *attB* integration sites for the phage ϕ C31 and ϕ BT1 (Combes et al., 2002; Gregory et al., 2003). The corresponding plasmid vectors also contain an integrase gene and *attP* sites specific to either ϕ C31 or ϕ BT1. However, integration of any plasmid into these sites can impact secondary metabolism, as the integration sites are within the coding region of well-conserved genes. ϕ C31-vectors integrate into a gene encoding a predicted chromosome condensation protein in *S. coelicolor* (*SCO3798*), while ϕ BT1-vectors integrate into a predicted transmembrane protein-encoding gene (*SCO4848*). In order to understand whether any new antibiotic production is a result of plasmid integration, or overexpression of an antibiotic regulator, it is necessary to generate a plasmid-alone negative control strain for comparison with the regulator overexpression strain.

A prime candidate for enhancing antibiotic production is Crp, which primarily functions as an activator of genes in many species and has been implicated in binding near promoters of biosynthetic clusters in *Streptomyces* (Gao et al., 2012). The

overexpression of Crp has also led to enhanced and novel antibiotic production in *S. coelicolor* (Gao et al., 2012).

We sought to deploy an integrative plasmid vector harboring *crp* under a strong promoter (Crp O/E) through conjugation into a large set of streptomycetes, and then screen these for the ability to inhibit the growth of two clinically-relevant pathogens: the Gram-positive methicillin-resistant *Staphylococcus aureus* (MRSA) and the Gram-negative *Pseudomonas aeruginosa* (Figure 5.1).

MRSA is a pathogen that can cause infections of the skin, but is notable for its prevalence in infecting the blood and tissues of individuals undergoing surgery (Hiramatsu et al., 2001). It is resistant to the β -lactam class of molecules, making it difficult to treat with front-line antibiotics (Hiramatsu et al., 2001). *P. aeruginosa* is an opportunistic pathogen that primarily affects immunocompromised individuals and those with cystic fibrosis (Canton et al., 2000). It is renowned for its antibiotic resistance, which it accomplishes via a variety of means, including: forming impermeable biofilms; activating multi-drug efflux pumps; and acquiring mutations in porin-encoding genes whose products facilitate antibiotic transit across the outer membrane (Li et al., 1994; Mah et al., 2003; Poole et al., 2011). Clinical isolates have revealed an abundance of drug-resistant *P. aeruginosa* within hospitals and studies have established a relationship between the presence of these bacteria and elevated morbidity of patients (Estrellado et al., 2010; Pai et al., 2001).

Gram-negative pathogens such as *P. aeruginosa* possess lipopolysaccharides (LPS) and glycopeptides in their outer membrane that make it impermeable to many antibiotics, including the peptidoglycan-targeting vancomycin (Al-tahhan et al., 2000). In contrast, the Gram-positive cell wall lacks this outer membrane, and instead possesses a thick, exposed peptidoglycan layer that can be readily accessed by cell wall-active antibiotics (Rice, 2006). As a result, the available treatments for Gram-negative pathogens are more limited, and antibiotics that are effective against Gram-negative bacteria also tend to be active against Gram positive bacteria (Al-tahhan et al., 2000; Nikaido, 1998). By targeting MRSA and *P. aeruginosa*, we can distinguish between

antibiotic activity that is Gram-positive specific, Gram-negative specific, and broad-spectrum (or lethal to all cells).

Introducing the empty vector and Crp overexpression plasmid to 200 strains from the Wright Actinomycete Collection, also known as the ‘WAC library’, were carried out previously via high-throughput conjugation, with 48 strains taking up at least one of the plasmids. Here, we generated a complete set of empty vector/Crp overexpression constructs for each of the 48 strains, and evaluated the impact of Crp overexpression on antibiotic production for each of these. As the majority of the initial 200 WAC strains did not take up the vector, we sought to assess all WAC antibiotic activity (both plasmid-transformed and naive) against MRSA and *P. aeruginosa*.

5.2 RESULTS

5.2.1 Select *Streptomyces* strains show activity against known pathogens

We grew individual colonies of each WAC strain for 4 days on solid Bennett’s medium and then overlaid cultures with a thin layer of Bennett’s medium inoculated with either MRSA or *P. aeruginosa*. After 24 h incubation, secreted antibiotics from select WAC strains would yield circular zones of inhibition in the otherwise dense growth of each pathogen. Positive hits were subjected to a second round of testing performed in duplicate, using the same conditions. In the absence of any insertional plasmids, four WAC strains demonstrated antibiotic activity against *P. aeruginosa* while 21 exhibited activity against MRSA. An example of activity against *P. aeruginosa* is in Figure 5.2.

The insertion of the empty vector alone was sufficient to stimulate/enhance antibiotic activity in four of the WAC strains, specifically against MRSA (Table 5.1). Notably, three of these had no detectable antibiotic activity on their own; however, for two of these WAC strains, antibiotic activity was abolished when transformed with the Crp overexpression plasmid (Table 5.1). This suggested that Crp may somehow attenuate the increased antibiotic production resulting from plasmid insertion in these strains.

Of the 48 strains harboring the integrative plasmids, three of them had newly emergent antibiotic production that was correlated with Crp overexpression (Table 5.1).

In these cases (modified WAC5514, WAC7779 and WAC8217), the new antibiotic activity was specific for MRSA. WAC5514 and WAC7779 exhibited total clearing within their zone of inhibition, suggesting that their new bioactive molecule(s) had bactericidal activity. WAC8217 elicited a zone that was noticeably less dense than the surrounding growth, but did not kill MRSA within that region, and thus may indicate bacteriostatic activity.

5.2.2 Media conditions have a profound effect on native and Crp-mediated antibiotic production

Growth medium is known to have significant effects on the secondary metabolites produced by streptomycetes, and this is expected to hold for strains expressing elevated levels of Crp. Sodium butyrate (Moore et al., 2012), manganese sulfate (Butterworth et al., 1979), and sodium citrate (Sánchez et al., 2010) are known to enhance antibiotic production in wild type *Streptomyces*. As a result, we sought to determine if alterations to media conditions would significantly impact the effect of Crp-overexpression on antibiotic production. Given that we detected only four strains that were capable of inhibiting *P. aeruginosa* growth, we further sought to probe whether the antibiotic production by these strains could be enhanced through medium optimization.

We explored a number of different growth conditions, summarized in Table 5.2. We found that the WAC strains with activity against *P. aeruginosa* in the first set of trials did not show any elevated activity following growth under these modified conditions. Moreover, experiments using minimal medium were rendered inconclusive due to an unexpected response from the *P. aeruginosa* overlay that turned an opaque blue colour and did not show any zone of growth inhibition.

Among the Crp-overexpression/empty vector strains subjected to either manganese sulfate or sodium butyrate exposure, we observed three additional strains containing the empty vector alone to have activity against MRSA. The combination of manganese sulfate and sodium butyrate supplementation further resulted in the identification of two additional WAC strains overexpressing Crp that gained the ability to

inhibit the growth of MRSA. Neither treatment alone or in combination resulted in any additional activity against *P. aeruginosa*.

5.3 DISCUSSION

Our process for identifying new or increased antibiotic activity involved growing a WAC strain alone, or transformed with either an empty plasmid vector or one from which Crp was overexpressed, on solid medium. The resulting colonies were then overlaid with a thin agar slice inoculated with the indicator organism of choice as described previously. Any nascent antibiotics made by WAC strains would prevent indicator growth immediately adjacent to the colony, and this would result in a zone of clearing (of varying sizes and degrees of clarity) around the WAC strain.

While our integrative vector was able to integrate into approximately one quarter of the total strains we attempted to conjugate into, we would ultimately like to expand our screen to include a larger pool of WAC strains. Why the majority of the strains we tested were failed to yield stable exconjugants is not clear. One possibility is that some strains may not be conjugation competent, they may lack a suitable ϕ C31 *attB* site, or they may not have been amenable to plasmid DNA uptake under the conditions used for the high-throughput conjugations were performed.

The resultant data suggested that of the wild type WAC strains tested, ~10% exhibited antibiotic activity against MRSA. In contrast, activity against *P. aeruginosa* was far less common. This was consistent with the known propensity of Gram-negative bacteria to resist antibiotic agents through the endogenous protection provided by their outer membranes, and of the particular capacity of *P. aeruginosa* to tolerate/resist antibiotic treatment.

Our strategy of incorporating a Crp-overexpressing construct into the WAC strains in a bid to increase antibiotic production led to the identification of several strains exhibiting new antibiotic production. However, these data also suggest that some strains increased their antibiotic production in response to gaining the plasmid vector alone. Interestingly, Crp-overexpression attenuated this increase in antibiotic production in

several strains. It is possible that the empty plasmid and the Crp O/E plasmid integrated at different places in the chromosome in these strains. While *Streptomyces* do contain a ϕ C31 site located within a chromosome condensation gene, the actual ϕ C31 sequence itself is poorly conserved and the integrative vector used here (pSET152) has been observed to integrate occasionally at pseudo-*attB* sites elsewhere in the chromosome (Combes et al., 2002). The empty vector could thus be integrating into a protein that regulates the production of some antibiotic, while the integration of the Crp-O/E plasmid is benign. Alternatively, each plasmid may be integrating into the same site, but the role of the condensation protein in these particular WAC species negatively regulates antibiotic production. Insertion of the empty vector into this gene relieves the repression, while insertion of Crp-O/E restores repression through an alternate pathway. Indeed, while Crp is known for its capacity to increase antibiotic production, particularly in *S. coelicolor*, we cannot exclude the possibility that it may transcriptionally activate genes that inhibit biosynthetic pathways in other streptomycetes.

The inability of Crp to enhance antibiotic production across a broad host range suggests that it may not be an ideal candidate, at least in isolation, for a general genetic construct to increase antibiotic production for screening purposes. It is possible that the WAC strains differ in their expression of native Crp. If there are saturating levels of Crp already being produced in a cell, then overexpression of exogenous Crp would not exhibit any effect. These strains may have also lacked adequate levels of cAMP (a co-factor of Crp) to sufficiently enhance the activity of Crp in the overexpression strains. We cannot exclude the possibility that the native Crp in each of these species may have a recognition site that differs from that of the *S. coelicolor* Crp, which was the protein being overexpressed. Finally, it is possible that of those WAC strains that could take up our plasmid vector, that these strains simply do not make many antibiotics, and may instead make other biologically active metabolites.

It is important to note that the bactericidal and bacteriostatic activities reported here are not necessarily of any clinical relevance. It is conceivable that the molecules responsible for conferring this observed activity against MRSA and *P. aeruginosa* have

already been identified. The bactericidal effects, in particular, may be due to molecules that are toxic to all living organisms, which could prove problematic for host cells in treating bacterial infections, although may then have potential chemotherapeutic utility.

5.4 FUTURE DIRECTIONS

Here we have described our progress in evaluating Crp as a global activator of antibiotic production across a broad range of *Streptomyces* species. Further developments stemming directly from this work should focus on the WAC strains that showed activity only when Crp-overexpression was present, and the WAC strains that were effective against *P. aeruginosa*. Those that demonstrated novel antibiotic activity when Crp was overexpressed present further evidence of Crp enhancing antibiotic production in another streptomycete. This suggests that Crp has utility as an antibiotic stimulatory tool, and could be further exploited in an effort to tap into the reservoir of cryptic and silent metabolites.

It would be worth narrowing the type of antibiotic that is associated with this activity, by screening extracts against otherwise susceptible bacteria harboring a range of resistances to known antibiotics. If the extracts are still able to kill these resistant bacteria, these compounds would be excellent candidates for purification and characterization. An effective strategy that could be employed to identify the molecule of interest, would be to compare the metabolic profile of empty vector-containing strains, with Crp-overexpressing strains using liquid chromatography/mass spectrometry, in association with bioassay-based fractionation. For those species for which a chromosome sequence is available, it could be useful to perform a bioinformatics analysis to identify potential Crp binding sites throughout their chromosomes, and assess their proximity to biosynthetic gene clusters. It is also possible to combine Crp activity modulation with the manipulation of another global regulator, with the goal of yielding secondary metabolite enhancements in the desired broad host range.

In addition to Crp, it will be important to identify other regulatory candidates that could be used in a plasmid vector to manipulate antibiotic production in *Streptomyces*. Our lab is currently investigating the HN-S-like protein known as Lsr2. Lsr2 acts as a

homodimeric putative repressor in *Streptomyces* and the related *Mycobacterium*. Investigations into Lsr2 have revealed that it binds DNA in the vicinity of biosynthetic clusters in *S. venezuelae* and that it represses the expression of many of these clusters, including the one for the well-known chloramphenicol antibiotic. We predict that overexpressing a dominant negative mutant Lsr2 variant may inhibit the activity of Lsr2 and stimulate antibiotic production in any *Streptomyces* species of interest. We are in the process of generating such Lsr2 ‘knock-down’ strains using our small WAC library, and will be testing these strains for new antibiotic activity.

Collectively, these experiments will provide further information on how we can manipulate Crp, how we can expand our strategies to include other global regulators, and how we can ensure that our genetic strategy will be tractable in as many streptomycetes as possible.

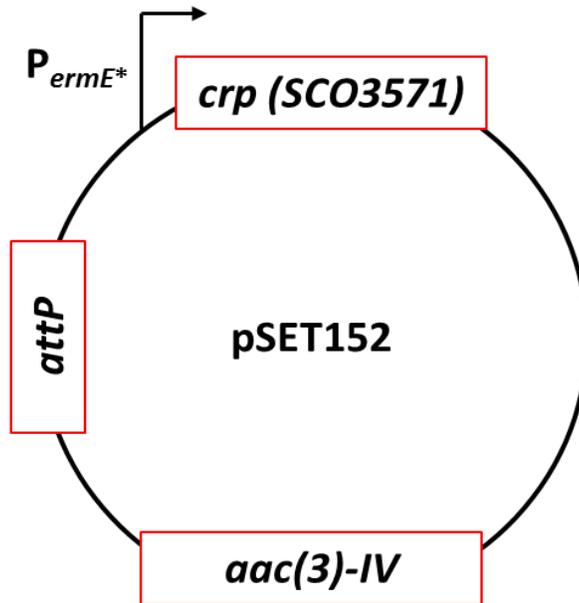


Figure 5.1: Integrating plasmid vector used in overexpressing the antibiotic regulator Crp. WAC strains were transformed with either an empty integrative plasmid (pSET152) or plasmid with *crp* under the control of a strong promoter (P_{ermE^*} -*crp*).

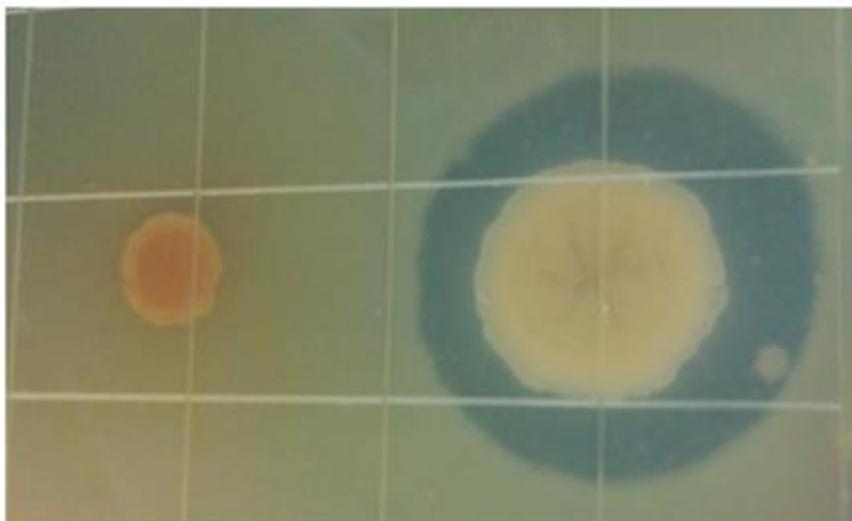


Figure 5.2: WAC strain antibiotic activity assay. Image depicts *P. aeruginosa* screened against WAC 7911 (LEFT) and WAC 8452 (RIGHT). WAC strains were grown for 4 days and overlaid with an inoculated thin agar sheet of Bennett's medium and grown an additional 24 hours at 30°C.

MRSA

WAC alone	WAC + empty	WAC + Crp O/E
572	-	-
1312	-	-
1351	-	-
1510	-	-
1561	-	-
2713	-	-
4143	4143 + empty	-
-	4718 + empty	-
4777	-	-
5344	-	-
5475	-	-
-	-	5514 + Crp
-	5836 + empty	5836 + Crp
-	6536 + empty	6536 + Crp
6833	-	-
7019	-	-
7072	-	-
7151	-	-
7166	-	-
7480	-	-
7509	-	-
7515	-	-
7573	-	-
-	-	7779 + Crp
7854	-	-
8133	-	-
-	-	8217 + Crp

P. aeruginosa

WAC alone	WAC + empty	WAC + Crp O/E
1312	-	-
1353	-	-
7540	-	-
8452	8452 + empty	8452 + Crp

Table 5.1: WAC strains hits for bacterial growth inhibition. WAC strains were transformed with either an empty integrative plasmid (pSET152) or plasmid with Crp under a strong promoter (*PermE*-crp*). Spores stocks of each strain were spotted on agar plates (Bennett's medium) for 96 h and overlaid with a thin agar sheet inoculated with either *MRSA* or *P. aeruginosa*. Zones of clearing were measured as clear (bacteriocidal, red) or cloudy (bacteriostatic, green), or in between (black). Zones were >3mm in size (colony sizes ranged from 4mm to 8mm). Dashes indicate strain variants with no hit.

Table 5.2: Media types used to enhance antibiotic production in WAC strains, as determined by bioassay.

Media	Condition	Concentration	Effect on WACs vs <i>P. aeruginosa</i>	Effect on WACs + empty/ <u>Crp-O/E</u> vs <i>MRSA</i>	Effect on WACs + empty/ <u>Crp-O/E</u> vs <i>P. aeruginosa</i>
Bennett's	None	-	-	-	-
	Ferrous sulfate	100 μ M	None	-	-
	Magnesium sulfate	2 <u>mM</u>	None	-	-
	Sodium citrate	20 μ M	None	-	-
	Sodium sulfate	3 <u>mM</u>	None	-	-
	Manganese sulfate	3 <u>mM</u>	None	WAC2599 + empty, WAC7072 + empty	None
	Sodium butyrate	150 μ M	None	WAC2599 + empty, WAC7072 + empty	None
	Sodium butyrate + Manganese sulfate	150 μ M, 3 <u>mM</u>	-	None	None
Minimal	None	-	-	-	-
	Ferrous sulfate	100 μ M	N/A	-	-
	Magnesium sulfate	2 <u>mM</u>	N/A	-	-
	Sodium citrate	20 μ M	N/A	-	-
	Sodium sulfate	3 <u>mM</u>	N/A	-	-
	Manganese sulfate	3 <u>mM</u>	N/A	WAC2599 + empty, WAC7072 + empty	None
	Sodium butyrate	150 μ M	N/A	None	None
	Sodium butyrate + Manganese sulfate	150 μ M, 3 <u>mM</u>	-	WAC2599 + empty, WAC7072 + empty, WAC6320 + <u>Crp O/E</u> , WAC7510 + <u>Crp O/E</u> ,	None

Chapter 6: Summary

6.1 Summary of work to date

Here we describe a multi-faceted approach to understanding and manipulating the regulatory systems of the genus *Streptomyces*.

We have examined the impact of the proposed DNA cleft binding pocket as well as two auxiliary interfaces on sIHF activity by assessing whether it can complement a Δ *sihf* phenotype.

Additionally, we have examined a silent promoter that governs expression of an unknown biosynthetic cluster and established that upstream nucleotide sequences and that the global regulator Crp may impact expression. We also have reason to believe that the downstream *SCO6439* gene may encode a pathway-specific regulatory protein that acts as an activator of our cluster.

Crp was then overexpressed in a range of actinomycetes in order to determine whether we could elicit novel bioactive metabolites.

This work has formed the basis for future investigations into these systems and how they may be manipulated in order to improve the coverage of the *Streptomyces* metabolome when screening for new pharmaceuticals.

APPENDIX**A1. Script for processing RNA-seq data (BASH + R)**

```

#The 12 Fastq files were derived from Illumina sequencing
#This script will be used to produce Fastqc files for evaluating the quality of each of the
12 datafiles.
mkdir sIHF_RNA-seq
mkdir ~/sIHF_RNA-seq/{Fastqc_quality_Untrimmed,Fastqc_quality_Trimmed}
#Files of reads uploaded to ~/sIHF_RNA-seq/Fastqc_quality_Trimmed using
MobaXterm upload command
#This function FastQC will produce a quality assessment for each of the sequencing read
files
/usr/local/fastqc/fastqc T*.fastq
#This script will trim the reads using Trimmomatic and yield new fastqc results files
for infile in ~/sIHF_RNA-seq/Fastqc_quality_Untrimmed/T*.fastq
>do
>outfile=${infile}_trimout.fastq
>java -jar /usr/local/trimmomatic/trimmomatic-0.33.jar SE $infile $outfile
MAXINFO:20:0.2
>done
mv ~/sIHF_RNA-seq/Fastqc_quality_Untrimmed/*out*.fastq ~/sIHF_RNA-
seq/Fastqc_quality_Trimmed
/usr/local/fastqc/fastqc T*.fastq
#Align single end reads to genome, pulled from NCBI Genbank database using
Tophat/Bowtie2
mkdir ~/sIHF_RNA-seq/{reads,mapped,counts,references}
cd ~/sIHF_RNA-seq/references
#Pull genomes and annotation file from ncbi GENBANK

```

```

curl -O
ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000253235.1_ASM25323v1/GCA_0002532
35.1_ASM25323v1_genomic.fna.gz
gunzip GCA_000253235.1_ASM25323v1_genomic.fna.gz
mv GCA_000253235.1_ASM25323v1_genomic.fna
Streptomyces_venezuelae_ATCC10712.fa
curl -O
ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000253235.1_ASM25323v1/GCA_0002532
35.1_ASM25323v1_genomic.gff.gz
gunzip GCA_000253235.1_ASM25323v1_genomic.gff.gz
mv GCA_000253235.1_ASM25323v1_genomic.gff
Streptomyces_venezuelae_ATCC10712.gff
#remove exon lines from the GFF file, a known issue with this genome build
sed '/exon/d' Streptomyces_venezuelae_ATCC10712.gff >
Streptomyces_venezuelae_ATCC10712_EXONREMOVED.gff
rm -rf Streptomyces_venezuelae_ATCC10712.gff
#Rename Exon_removed file as the main base name for the genome build
mv Streptomyces_venezuelae_ATCC10712_EXONREMOVED.gff
Streptomyces_venezuelae_ATCC10712.gff
#Build indexing files using Bowtie2-build
bowtie2-build -f Streptomyces_venezuelae_ATCC10712.fa
Streptomyces_venezuelae_ATCC10712
samtools faidx Streptomyces_venezuelae_ATCC10712.fa
reference=~/.sIHF_RNA-seq/references/Streptomyces_venezuelae_ATCC10712
annotation=~/.sIHF_RNA-seq/references/Streptomyces_venezuelae_ATCC10712.gff
#Map all reads using Bowtie2
#first create an array for all 12 sequencing files
samples[1]=T1d1080v1__good_kS1n
samples[2]=T1d1080v2__good_7ziY

```

```

samples[3]=T1WTv1__good_sBrb
samples[4]=T1WTv2__good_6tpN
samples[5]=T2d1080v1__good_Xc5I
samples[6]=T2d1080v2__good_i6Uk
samples[7]=T2WTv1__good_HMG9
samples[8]=T2WTv2__good_RZ75
samples[9]=T3d1080v1__good_qIJx
samples[10]=T3d1080v2__good_8xol
samples[11]=T3WTv1__good_WGIq
samples[12]=T3WTv2__good_tV5J
for i in {1..12}
do
sample=${samples[$i]}
echo ${sample}
bowtie2 -p 6 -x ${reference} -U ~/sIHF_RNA-seq/reads/${sample}.fastq_trimout.fastq -
S ~/sIHF_RNA-seq/mapped/${sample}.sam
done
#Convert .sam file to .bam files
for i in {1..12}
do
sample=${samples[$i]}
echo ${sample}
samtools view -bS ~/sIHF_RNA-seq/mapped/${sample}.sam > ~/sIHF_RNA-
seq/mapped/${sample}.bam
done
#sort bam files by name
for i in {1..12}
do
sample=${samples[$i]}

```

```

echo ${sample}
samtools sort -n ~/sIHF_RNA-seq/mapped/${sample}.bam ${sample}_nsorted.bam
done
for i in {1..12}
do
sample=${samples[${i}]}
echo ${sample}
#Count the number of reads mapping to each feature using HTSeq
htseq-count --format=bam --stranded=yes --type=gene --idattr=ID --order=name
~/sIHF_RNA-seq/mapped/${sample}_nsorted.bam ${annotation} > ~/sIHF_RNA-
seq/counts/${sample}.htseq_counts.txt
done
#Check count numbers make sense with total # of mapped reads
for i in {1..12}
do
sample=${samples[${i}]}
echo ${sample}
awk '{s+= $2} END{print s}' ~/sIHF_RNA-seq/counts/${sample}
done
#####
Appendix A2 R Script for Statistical analysis
#Use R-script on HT-seq output in order to process it through DESeq2
library(DESeq2)
library("RColorBrewer")
library("gplots")
#code below partly adapted from http://dwheelerau.com/2014/02/17/how-to-use-deseq2-to-analyse-RNA-seq-data/
#files at each timepoint renamed to be "sihf" or "delsihf" with replicate number
#####T1#####

```

```

setwd("C:/Users/Chris/Desktop/HTseq_Files/T1")
library("DESeq2")
directory <- "C:/Users/Chris/Desktop/HTseq_Files/T1"
sampleFiles <- grep("sihf",list.files(directory),value=TRUE)
sampleCondition <- c("delsihf","delsihf","sihf","sihf")
sampleTable <- data.frame(sampleName=sampleFiles, fileName=sampleFiles,
condition=sampleCondition)
ddsHTSeq<-DESeqDataSetFromHTSeqCount(sampleTable=sampleTable,
directory=directory, design=~condition)
colData(ddsHTSeq)$condition <- factor(colData(ddsHTSeq)$condition,
levels=c("sihf","delsihf"))
dds<-DESeq(ddsHTSeq)
res<-results(dds)
#order according to padj value
res<-res[order(res$padj),]
head(res)
#MA plot
plotMA(dds,ylim=c(-8,8),main="DESeq2")
#write csv file for export
write.csv(as.data.frame(res),file="T1_analysis.csv")
#dispersion of data
plotDispEsts(dds)
#variance stabilized transformation of data
vsd <- varianceStabilizingTransformation(dds, blind=TRUE)
#variance stabilized effect mapping
select <- order(rowMeans(counts(dds,normalized=TRUE)),decreasing=TRUE)[1:30]
hmc col <- colorRampPalette(brewer.pal(9, "GnBu"))(100)
heatmap.2(counts(dds,normalized=TRUE)[select,], col = hmc col,
Rowv = FALSE, Colv = FALSE, scale="none",

```

```

dendrogram="none", trace="none", margin=c(10,6))
heatmap.2(assay(vsd)[select,], col = hmccl,
Rowv = FALSE, Colv = FALSE, scale="none",
dendrogram="none", trace="none", margin=c(10, 6))
# heat map of data sample clustering
distsRL <- dist(t(assay(vsd)))
mat <- as.matrix(distsRL)
rownames(mat) <- colnames(mat) <- with(colData(dds),
paste(condition,sampleFiles , sep=" : "))
hc <- hclust(distsRL)
heatmap.2(mat, Rowv=as.dendrogram(hc),
symm=TRUE, trace="none",
col = rev(hmccl), margin=c(15, 15))
#PCA plot of vsd data
print(plotPCA(vsd, intgroup=c("condition")))
#####T2#####
setwd("C:/Users/Chris/Desktop/HTseq_Files/T2")
directory <- "C:/Users/Chris/Desktop/HTseq_Files/T2"
sampleFiles <- grep("sihf",list.files(directory),value=TRUE)
sampleCondition <- c("delsihf","delsihf","sihf","sihf")
sampleTable <- data.frame(sampleName=sampleFiles, fileName=sampleFiles,
condition=sampleCondition)
ddsHTSeq<-DESeqDataSetFromHTSeqCount(sampleTable=sampleTable,
directory=directory, design=~condition)
colData(ddsHTSeq)$condition <- factor(colData(ddsHTSeq)$condition,
levels=c("sihf","delsihf"))
dds<-DESeq(ddsHTSeq)
res<-results(dds)
#order according to padj value

```

```

res<-res[order(res$padj),]
head(res)
#MA plot
plotMA(dds,ylim=c(-8,8),main="DESeq2")
#write csv file for output
write.csv(as.data.frame(res),file="T2_analysis.csv")
#dispersion of data
plotDispEsts(dds)
#variance stabilized transformation of data
vsd <- varianceStabilizingTransformation(dds, blind=TRUE)
#variance stabilized effect heat maps
select <- order(rowMeans(counts(dds,normalized=TRUE)),decreasing=TRUE)[1:30]
hmcol <- colorRampPalette(brewer.pal(9, "GnBu"))(100)
heatmap.2(counts(dds,normalized=TRUE)[select,], col = hmcol,
Rowv = FALSE, Colv = FALSE, scale="none",
dendrogram="none", trace="none", margin=c(10,6))
heatmap.2(assay(vsd)[select,], col = hmcol,
Rowv = FALSE, Colv = FALSE, scale="none",
dendrogram="none", trace="none", margin=c(10, 6))
# heat map of data sample clustering
distsRL <- dist(t(assay(vsd)))
mat <- as.matrix(distsRL)
rownames(mat) <- colnames(mat) <- with(colData(dds),
paste(condition,sampleFiles , sep=" : "))
hc <- hclust(distsRL)
heatmap.2(mat, Rowv=as.dendrogram(hc),
symm=TRUE, trace="none",
col = rev(hmcol), margin=c(15, 15))
#PCA plot of vsd data

```

```

print(plotPCA(vsd, intgroup=c("condition")))
#####T3#####
setwd("C:/Users/Chris/Desktop/HTseq_Files/T3")
directory <- "C:/Users/Chris/Desktop/HTseq_Files/T3"
sampleFiles <- grep("sihf",list.files(directory),value=TRUE)
sampleCondition <- c("delsihf","delsihf","sihf","sihf")
sampleTable <- data.frame(sampleName=sampleFiles, fileName=sampleFiles,
condition=sampleCondition)
ddsHTSeq<-DESeqDataSetFromHTSeqCount(sampleTable=sampleTable,
directory=directory, design=~condition)
colData(ddsHTSeq)$condition <- factor(colData(ddsHTSeq)$condition,
levels=c("sihf","delsihf"))
dds<-DESeq(ddsHTSeq)
res<-results(dds)
#order according to padj value
res<-res[order(res$padj),]
head(res)
#MA plot
plotMA(dds,ylim=c(-8,8),main="DESeq2")
#write csv file
write.csv(as.data.frame(res),file="T3_analysis.csv")
#dispersion of data
plotDispEsts(dds)
#variance stabilized transformation of data
vsd <- varianceStabilizingTransformation(dds, blind=TRUE)
#variance stabilized effect
select <- order(rowMeans(counts(dds,normalized=TRUE)),decreasing=TRUE)[1:30]
hmc col <- colorRampPalette(brewer.pal(9, "GnBu"))(100)
heatmap.2(counts(dds,normalized=TRUE)[select,], col = hmc col,

```

```
Rowv = FALSE, Colv = FALSE, scale="none",
dendrogram="none", trace="none", margin=c(10,6))
heatmap.2(assay(vsd)[select,], col = hmcol,
Rowv = FALSE, Colv = FALSE, scale="none",
dendrogram="none", trace="none", margin=c(10, 6))
# heat map of data sample clustering
distsRL <- dist(t(assay(vsd)))
mat <- as.matrix(distsRL)
rownames(mat) <- colnames(mat) <- with(colData(dds),
paste(condition,sampleFiles , sep=" : "))
hc <- hclust(distsRL)
heatmap.2(mat, Rowv=as.dendrogram(hc),
symm=TRUE, trace="none",
col = rev(hmcol), margin=c(15, 15))
#PCA plot of vsd data
print(plotPCA(vsd, intgroup=c("condition")))
```

A2. Comparison of RNA-seq processing methods

Function	Initial Processing	Second Processing
Read Trimming	TrimR	Trimmomatic
Genome database	Bowtie	Bowtie2
Mapping/Counting	eXpress	HTSeq
Data Analysis	DESeq2	DESeq2

REFERENCES

- Al-tahhan, R. A., Sandrin, T. R., Bodour, A. A., & Maier, R. M. (2000). Rhamnolipid-induced removal of lipopolysaccharide from *Pseudomonas aeruginosa*: effect on cell surface properties and interaction with hydrophobic substrates. *Applied and Environmental Microbiology*. 66(8), 3262–3268.
- Altshul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*. 215(215), 403–410.
- Anders, S., Pyl, P. T., & Huber, W. (2015). Genome analysis HTSeq — a Python framework to work with high-throughput sequencing data. *Bioinformatics*. 31(2), 166–169.
- Anzai, Y., Iizaka, Y., Li, W., Idemoto, N., Kazuo, S. T., & Kenji, K. (2009). Production of rosamicin derivatives in *Micromonospora rosaria* by introduction of D -mycinose biosynthetic gene with C31-derived integration vector pSET152. *Journal of Industrial Microbiology and Biotechnology*. 1013–1021.
- Baltz, R. H. (2008). Renaissance in antibacterial discovery from actinomycetes. *Current Opinion in Pharmacology*, 8, 557–563.
- Barka, E. A., Vatsa, P., Sanchez, L., Gaveau-Vaillant, N., Jacquard, C., Klenk, H.-P., ... van Wezel, G. P. (2016). Taxonomy, physiology, and natural products of actinobacteria. *Microbiology and Molecular Biology Reviews*. 80(1), 1–43.
- Barria, C., Malecki, M., & Arraiano, C. M. (2013). Bacterial adaptation to cold. *Microbiology 159*: 2437–2443.
- Becker, N. A., Kahn, J. D., & Maher, L. J. (2007). Effects of nucleoid proteins on DNA repression loop formation in *Escherichia coli*. *Nucleic Acids Research*, 35(12), 3988–4000.
- Bentley, S., Chater, K., Cerdeño-Tárraga, A., Challis, G., Thomson, N., James, K., Harris,

- D., Quail, M., Kieser, H., Harper, D., Bateman, A., Brown, S., Chandra, G. (2002). Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature*. 417(6885), 141–147.
- Bernhardt, T. G., & De Boer, P. A. J. (2005). SlmA, a nucleoid-associated, FtsZ binding protein required for blocking septal ring assembly over chromosomes in *E. coli*. *Molecular Cell*, 18(5), 555–564.
- Beylot, C., Auffret, N., Poli, F., Claudel, J. P., Leccia, M. T., Del Giudice, P., & Dreno, B. (2014). *Propionibacterium acnes*: An update on its role in the pathogenesis of acne. *Journal of the European Academy of Dermatology and Venereology*. 28(3), 271–278.
- Bibb, M. J. (2005). Regulation of secondary metabolism in streptomycetes. *Current Opinion in Microbiology*, 8(2), 208–215.
- Bierman, M., Logan, R., O'Brien, K., Seno, E. T., Rao, R. N., & Schoner, B. E. (1992). Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp. *Gene*, 116, 43–49.
- Boeck, L. D., Higgins, G. E., & Kastner, R. E. (1971). A new dipeptide antibiotic from *Streptomyces collinus*. *The Journal of Antibiotics*. XXV(2) 137-141.
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114–2120.
- Brown, E. D., & Wright, G. D. (2016). Antibacterial drug discovery in the resistance era. *Nature*. 529:336-343
- Browning, D. F., & Busby, S. J. W. (2004). The regulation of bacterial transcription initiation. *Nature Reviews Microbiology*. 2(1): 57–65.
- Bush, M. J., Chandra, G., Bibb, M. J., Findlay, K. C., & Buttner, M. J. (2016). Genome-wide chromatin immunoprecipitation sequencing analysis shows that WhiB is a

transcription factor that cocontrols its regulon with WhiA to initiate developmental cell division in *Streptomyces*. *mBio*. 7(2): 1–13.

- Butterworth, D., Cole, M., Hanscomb, G., & Rolinson, G. N. (1979). Olivanic acids, a family of β -lactam antibiotics with β -lactamase inhibitory properties produced by *Streptomyces* species. I. Detection, properties and fermentation studies. *The Journal of Antibiotics*. 32(4): 287–294.
- Canton, R., Universitario, H., & Ram, F. B. (2000). High frequency of hypermutable *Pseudomonas aeruginosa* in Cystic Fibrosis Lung Infection. *Science*. 288(1251):1251-1253
- Chaikam, V., & Karlson, D. T. (2010). Comparison of structure, function and regulation of plant cold shock domain proteins to bacterial and animal cold shock domain proteins. *BMB Reports*, 43(1), 1–8.
- Challis, G. L. (2008). Mining microbial genomes for new natural products and biosynthetic pathways. *Microbiology*, 154(6), 1555–1569.
- Chater, K. F. (2006). *Streptomyces* inside-out: a new perspective on the bacteria that provide us with antibiotics. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 361(1469), 761–768.
- Chen, C. W., Huang, C. H., Lee, H. H., Tsai, H. H., & Kirby, R. (2002). Once the circle has been broken: Dynamics and evolution of *Streptomyces* chromosomes. *Trends in Genetics*, 18(10), 522–529.
- Chen, J. M., Ren, H., Shaw, J. E., Wang, Y. J., Li, M., Leung, A. S., ... Liu, J. (2008). Lsr2 of Mycobacterium tuberculosis is a DNA-bridging protein. *Nucleic Acids Research*, 36(7), 2123–2135.
- Cobb, R. E., Wang, Y., & Zhao, H. (2015). High-efficiency multiplex genome editing of *Streptomyces* species using an engineered crispr/cas system. *American Chemical Society Synthetic Biology*, 4(6), 723–728.

- Combes, P., Till, R., Bee, S., & Smith, M. C. M. (2002). The *Streptomyces* genome contains multiple pseudo- attB sites for the c31-encoded site-specific recombination system. *Journal of Bacteriology*, 184(20), 5746–5752.
- Craig, N. L., & Nash, H. A. (1984). E. coli integration host factor binds to specific sites in DNA. *Cell*, 39(3), 707–716.
- Demain, A. L. (1974). How do antibiotic-producing microorganisms avoid suicide? *Annals of the New York Academy of Sciences*, 235(1), 601–612.
- Dillon, S. C., & Dorman, C. J. (2010). Bacterial nucleoid-associated proteins , nucleoid structure and gene expression. *Nature Reviews Microbiology*, 8(3), 185–195.
- Elibol, M. (2004). Optimization of medium composition for actinorhodin production by *Streptomyces coelicolor* A3 (2) with response surface methodology. *Process Biochemistry*. 39, 1057–1062.
- Esposito, D., Petrovic, A., Harris, R., Ono, S., Eccleston, J. F., Mbabaali, A., Ladbury, J. E. (2002). H-NS oligomerization domain structure reveals the mechanism for high order self-association of the intact protein. *Journal of Molecular Biology*, 324(4), 841–850.
- Estrellado, W., Baraboutis, I. G., & Sriram, P. (2010). NIH Public Access, 64(3), 311–319.
- Fiedler, H. (1993). Screening for secondary metabolites by hplc and uv-visible absorbance spectral libraries. *Natural Product Letters*. 2, 119–128.
- French, G. L. (2010). The continuing crisis in antibiotic resistance. *International Journal of Antimicrobial Agents*. 36, S3–S7.
- Gao, C., Mulder, D., Yin, C., & Elliot, M. A. (2012). Crp Is a Global Regulator of Antibiotic Production in *Streptomyces*. *mBio*. 3(6), 1–12.
- Gerlach, P., Sogaard-Andersen, L., Pedersen, H., Martinussen, J., Valentin-Hansen, P., &

- Bremer, E. (1991). The cyclic AMP (cAMP)-cAMP receptor protein complex functions both as an activator and as a corepressor at the *tsx-p2* promoter of *Escherichia coli* K-12. *Journal of Bacteriology*, *173*(17), 5419–5430.
- Gilboa, R., Zharkov, D. O., Golan, G., Fernandes, A. S., Gerchman, S. E., Matz, E., Shoham, G. (2002). Structure of formamidopyrimidine-DNA glycosylase covalently complexed to DNA. *Journal of Biological Chemistry*, *277*(22), 19811–19816.
- Glazebrook, M. a, Doull, J. L., Stuttard, C., & Vining, L. C. (1990). Sporulation of *Streptomyces venezuelae* in submerged cultures. *Journal of General Microbiology*. *136*(1990), 581–588.
- Gomez-escribano, J. P., & Bibb, M. J. (2011). Engineering *Streptomyces coelicolor* for heterologous expression of secondary metabolite gene clusters. *Biotechnology*. *4*, 207–215.
- Gordon, B. R. G., Imperial, R., Wang, L., Navarre, W. W., & Liu, J. (2008). Lsr2 of *Mycobacterium* represents a novel class of H-NS-like proteins. *Journal of Bacteriology*, *190*(21), 7052–7059.
- Gregory, M. A., Till, R., & Smith, M. C. M. (2003). Integration site for *Streptomyces* phage bt1 and development of site-specific integrating vectors. *Journal of Bacteriology*. *185*(17), 5320–5323.
- Grunstein, M. (1997). Histone acetylation in chromatin structure and transcription. *Nature*, *389*(6649), 349–352.
- Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *Journal of Molecular Biology*, *166*(4), 557–580.
- Hassan, R., Shaaban, M. I., Abdel Bar, F. M., El-Mahdy, A. M., & Shokralla, S. (2016). Quorum sensing inhibiting activity of *Streptomyces coelicoflavus* isolated from soil. *Frontiers in Microbiology*. *7*, 1–12.

- Helmann, J. D., & Chamberlin, M. J. (1988). Structure and function of bacterial sigma factors. *Annual Reviews Biochemistry*, (8).
- Hendlin, A. D., Stapley, E. O., Jackson, M., Wallick, H., Miller, A. K., Wolf, F. J., ... Mochales, S. (2017). Phosphonomycin, a New Antibiotic Produced by Strains of *Streptomyces*. *Science*. 166(3901):122-123.
- Hiramatsu, K., Cui, L., Kuroda, M., & Ito, T. (2001). The emergence and evolution of methicillin-resistant *Staphylococcus aureus*. *Trends in Microbiology*. 9(10):486-493
- Hobbs, G., Frazer, C. M., Gardner, D. C. J., Flett, F., & Oliver, S. G. (1990). Pigmented antibiotic production by *Streptomyces coelicolor* A3(2): kinetics and the influence of nutrients. *Journal of General Microbiology*, 136(11), 2291–2296.
- Holden, M. T. G., McGowan, S. J., Bycroft, B. W., Stewart, G. S. A. B., Williams, P., & Salmond, G. P. C. (1998). Cryptic carbapenem antibiotic production genes are widespread in *Erwinia carotovora*: Facile trans activation by the *carR* transcriptional regulator. *Microbiology*, 144(6), 1495–1508.
- Hong, H. J., Hutchings, M. I., Hill, L. M., & Buttner, M. J. (2005). The role of the novel fem protein VanK in vancomycin resistance in *Streptomyces coelicolor*. *Journal of Biological Chemistry*, 280(13), 13055–13061.
- Hopwood, D. A. (2006). Soil to genomics: the *Streptomyces* chromosome. *Annual Review of Genetics*, 40(1), 1–23.
- Horikoshi, M., Bertuccioli, C., Takada, R., Wang, J., Yamamoto, T., & Roeder, R. G. (1992). Transcription factor *tfiid* induces dna bending upon binding to the tata element. *Proceedings of the National Academy of Sciences of the United States of America*, 89(3), 1060–1064.
- Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2008). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, 4(1), 44–57.

- Huang, J., Lih, C., Pan, K., & Cohen, S. N. (2001). Global analysis of growth phase responsive gene expression and regulation of antibiotic biosynthetic pathways in *Streptomyces coelicolor* using DNA microarrays. *Genes and Development*, *15*: 3183–3192.
- Katz, E., & Demain, A. L. (1977). The peptide antibiotics of *Bacillus*: chemistry, biogenesis, and possible functions. *Bacteriological Reviews*, *41*(2), 449–74.
- Katz, M., Hover, B. M., & Brady, S. F. (2016). Culture - independent discovery of natural products from soil metagenomes. *Journal of Industrial Microbiology & Biotechnology*, *43*(2), 129–141.
- Kinashi, H., Otten, S. L., Duncann, J. S., & Hutchinson, C. R. (1988). Frequent loss and restoration of antibiotic production by *Streptomyces lasaliensis*. *The Journal of Antibiotics*. *XLI*(5): 624–637.
- Kolb, A., Busby, S., & Buc, I. (1993). Transcriptional regulation by cAMP and its receptor protein. *Annual Review of Biochemistry*, *1993*(62), 749–795.
- Kumar, P. S., Duraipandiyar, V., & Ignacimuthu, S. (2014). Isolation, screening and partial purification of antimicrobial antibiotics from soil *Streptomyces* sp. SCA 7. *Kaohsiung Journal of Medical Sciences*, *30*(9), 435–446.
- Lam, K. S. (2006). Discovery of novel metabolites from marine actinomycetes. *Current Opinion in Microbiology*, *9*(3), 245–251.
- Langmead. (2013). Bowtie2. *Nature Methods*, *9*(4), 357–359.
- Lasonder, E., Hoffmann, A., Vuister, G. W., Grzesiek, S., & Hilbers, C. W. (2002). The solution structure and dna-binding properties of the cold-shock domain of the human y-box protein yb-1. *Journal of Molecular Biology* *316*: 317-326.
- Lawrence, J. G., & Roth, J. R. (1996). Selfish operons: Horizontal transfer may drive the evolution of gene clusters. *Genetics*, *143*(4), 1843–1860.

- Li, C., Wally, H., Miller, S. J., & Lu, C. D. (2009). The multifaceted proteins MvaT and MvaU, members of the H-NS family, control arginine metabolism, pyocyanin synthesis, and prophage activation in *Pseudomonas aeruginosa* PAO1. *Journal of Bacteriology*, *191*(20), 6211–6218.
- Li, X., Livermore, D. M., & Nikaido, H. (1994). Role of efflux pump(s) in intrinsic resistance of *Pseudomonas aeruginosa* : Resistance to tetracycline , chloramphenicol , and norfloxacin. *Antimicrobial agents and Chemotherapy*. *38*(8), 1732–1741.
- Liao, G., Xie, L., Li, X., Cheng, Z., & Xie, J. (2014). Unexpected extensive lysine acetylation in the trump-card antibiotic producer *Streptomyces roseosporus* revealed by proteome-wide profiling. *Journal of Proteomics*, *106*, 260–269.
- Lin, J., Chen, H., Dröge, P., & Yan, J. (2012). Physical Organization of DNA by Multiple Non-Specific DNA-Binding Modes of Integration Host Factor (IHF). *PLoS ONE*, *7*(11).
- Lisser, S., & Margalit, H. (1993). Compilation of *E. coli* mRNA promoter sequences. *Nucleic Acids Research*. *21*(7), 1507–1516.
- Liu-Johnson, H. N., Gartenberg, M. R., & Crothers, D. M. (1986). The DNA binding domain and bending angle of *E. coli* CAP protein. *Cell*, *47*(6), 995–1005.
- Liu, Y., Chen, H., & Kenney, L. J. (2010). A divalent switch drives H-NS / DNA-binding conformations between stiffening and bridging modes service A divalent switch drives conformations between stiffening and bridging modes. *Genes & Development*.. *346* 339–344.
- Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 550.
- MacNeil, D. J., Gewain, K. M., Ruby, C. L., Dezeny, G., Gibbons, P. H., & MacNeil, T. (1992). Analysis of *Streptomyces avermitilis* genes required for avermectin biosynthesis utilizing a novel integration vector. *Gene*, *111*(1), 61–68.

- Mah, T., Pitts, B., Pellock, B., Walker, G. C., Stewart, P. S., & Toole, G. A. O. (2003). A genetic basis for *Pseudomonas aeruginosa* biofilm antibiotic resistance. *Nature*, *426*, 1–5.
- Manteca, A., Alvarez, R., Salazar, N., Yagüe, P., & Sanchez, J. (2008). Mycelium differentiation and antibiotic production in submerged cultures of *Streptomyces coelicolor*. *Applied and Environmental Microbiology*, *74*(12), 3877–3886.
- Martínez-Antonio, A., & Collado-Vides, J. (2003). Identifying global regulators in transcriptional regulatory networks in bacteria. *Current Opinion in Microbiology*, *6*(5), 482–489.
- Mazurakova, V., Sevcikova, B., Rezuchova, B., & Kormanec, J. (2006). Cascade of sigma factors in streptomycetes: Identification of a new extracytoplasmic function sigma factor that is under the control of the stress-response sigma factor in *Streptomyces coelicolor* A3(2). *Archives of Microbiology*, *186*(6), 435–446.
- Mcgregor, J. F. (1954). Nuclear division and the life cycle in a *Streptomyces* species. *Journal of General Microbiology*, *11*(1), 52–56.
- Miller, J. H. (1972). Experiments in molecular genetics. *Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.*
- Mishra, A., Viji, M., Kumar, D., Taneja, V., Mondal, A. K., Bothra, A., Taneja, B. (2013). Integration host factor of mycobacterium tuberculosis, mihf, compacts dna by a bending mechanism. *PLoS ONE*, *8*(7)
- Miyadoh, S. (1993). Research on antibiotic screening in japan over the last decade: a producing microorganism approach. *Actinomycetologica*. *7*:100-106.
- Moore, J. M., Bradshaw, E., Seipke, R. F., Hutchings, M. I., & McArthur, M. (2012). Use and discovery of chemical elicitors that stimulate biosynthetic gene clusters in streptomyces bacteria. *Methods in Enzymology*. *517*, 367–385.

- Myronovskyi, M., Welle, E., Fedorenko, V., & Luzhetskyy, A. (2011). B-Glucuronidase as a sensitive and versatile reporter in Actinomycetes. *Applied and Environmental Microbiology*. 77(15): 5370–5383.
- Nakashima, K., Kanamaru, K., Mizuno, T., & Horikoshi, K. (1996). A novel member of the cspA family of genes that is induced by cold shock in *Escherichia coli*. *Journal of Bacteriology*. 178(10): 2994–2997.
- Nanji, T. (2014). The structure and function of the sihf protein. Thesis. McMaster university.
- Nanji, T. et al. (2017). *Streptomyces* IHF uses multiple interfaces to bind DNA and condense nucleoids. *Under Review*.
- Nikaido, H. (1998). Antibiotic Resistance Caused by Gram-Negative Multidrug Efflux Pumps. *Biochemical Mechanisms*. 27: 32–41.
- Niu, W., Kim, Y., Tau, G., Heyduk, T., & Ebright, R. H. (1996). Transcription activation at class II CAP-dependent promoters: Two interactions between CAP and RNA polymerase. *Cell*. 87(6): 1123–1134.
- Oshima, T., Ishikawa, S., Kurokawa, K., Aiba, H., & Ogasawara, N. (2006). *Escherichia coli* histone-like protein H-NS preferentially binds to horizontally acquired DNA in association with RNA polymerase. *DNA Research*. 13(4): 141–153.
- Pai, H., Kim, J., Kim, J., Lee, J. I. H., Choe, K. W. O. N., & Gotoh, N. (2001). Carbapenem resistance mechanisms in *Pseudomonas aeruginosa* clinical isolates. *Antimicrobial Agents and Chemotherapy*. 45(2): 480–484.
- Park, P. J. (2009). ChIP-seq: advantages and challenges of a maturing technology. *Nature Reviews Genetics*. 10(10): 669–680.
- Pedulla, M. L., & Hatfull, G. F. (1998). Characterization of the mIHF gene of *Mycobacterium smegmatis*. *Journal of Bacteriology*. 180(20): 5473–5477.

- Pérez-Redondo, R., Rodríguez-García, A., Martín, J. F., & Liras, P. (1998). The *clbR* gene of *Streptomyces clavuligerus*, encoding a LysR-type regulatory protein controlling clavulanic acid biosynthesis, is linked to the clavulanate-9-aldehyde reductase (*car*) gene. *Gene*. 211(2): 311–321.
- Poole, K., Hazlett, L. D., State, W., & Greenberg, E. P. (2011). *Pseudomonas aeruginosa* : resistance to the max. *Frontiers in Microbiology*. 2: 1–13.
- Rahman, S.A., Singh, Y. (2014). Pathogenic mycobacteria reveal genomic and biochemical variabilities. *mBio*. 5(6): 1-2.
- Ramadhar, T. R., Beemelmans, C., Currie, C. R., & Clardy, J. (2013). Bacterial symbionts in agricultural systems provide a strategic source for antibiotic discovery. *The Journal of Antibiotics*. 67(1): 53–58.
- Rennella, E., Sára, T., Juen, M., Wunderlich, C., Imbert, L., Solyom, Z., ... Brutscher, B. (2017). RNA binding and chaperone activity of the *E. coli* cold-shock protein CspA. *Nucleic Acids Research*. 45(7): 4255-4268.
- Rice, L. B. (2006). Antimicrobial Resistance in Gram-Positive Bacteria. *The American Journal of Medicinie*. 119: 11–19.
- Ruddick, S. M., & Williams, S. T. (1972). Studies on the ecology of actinomycetes in soil V. Some factors influencing the dispersal and adsorption of spores in soil. *Soil Biology and Biochemistry*. 4(1): 93–103.
- Rutledge, P. J., & Challis, G. L. (2015). Discovery of microbial natural products by activation of silent biosynthetic gene clusters. *Nature Reviews Microbiology*. 13(8): 509–523.
- Sánchez, S., Chávez, A., Forero, A., García-Huante, Y., Romero, A., Sánchez, M., ... Ruiz, B. (2010). Carbon source regulation of antibiotic production. *The Journal of Antibiotics*. 63(8): 442–459.

- Sangal, V., & Hoskisson, P. A. (2016). Evolution, epidemiology and diversity of *Corynebacterium diphtheriae*: New perspectives on an old foe. *Infection, Genetics and Evolution*. 43: 364–370.
- Sasseti, C. M., Boyd, D. H., & Rubin, E. J. (2003). Genes required for mycobacterial growth defined by high density mutagenesis. *Molecular Microbiology*. 48(1): 77–84.
- Scherlach, K., & Hertweck, C. (2009). Triggering cryptic natural product biosynthesis in microorganisms. *Organic & Biomolecular Chemistry*. 7(9): 1753–1760.
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., ... Cardona, A. (2012). Fiji: an open-source platform for biological-image analysis. *Nature Methods*. 9(7): 676–82.
- Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*. 9(7): 671–675.
- Sherwood, E. J., Hesketh, A. R., & Bibb, M. J. (2013). Cloning and Analysis of the Planosporicin Lantibiotic Biosynthetic Gene Cluster of *Planomonospora alba*. *Journal of Bacteriology*. 195(10): 2309–2321.
- Smits, W. K., & Grossman, A. D. (2010). The transcriptional regulator Rok binds A+T-rich DNA and is involved in repression of a mobile genetic element in *Bacillus subtilis*. *PLoS Genetics*. 6(11).
- St-Onge, R. J., Haiser, H. J., Yousef, M. R., Sherwood, E., Tschowri, N., Al-Bassam, M., & Elliot, M. A. (2015). Nucleotide second messenger-mediated regulation of a muralytic enzyme in *Streptomyces*. *Molecular Microbiology*. 96(4): 779–795.
- Süsstrunk, U., Pidoux, J., Taubert, S., Ullmann, A., & Thompson, C. J. (1998). Pleiotropic effects of cAMP on germination, antibiotic biosynthesis and morphological development in *Streptomyces coelicolor*. *Molecular Microbiology*. 30(1): 33–46.

- Swiercz, J. P., Nanji, T., Gloyd, M., Guarné, A., & Elliot, M. A. (2013). A novel nucleoid-associated protein specific to the actinobacteria. *Nucleic Acids Research*. 41(7): 4171–4184.
- Takano, E., Kinoshita, H., Mersinias, V., Bucca, G., Hotchkiss, G., Nihira, T., 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*. *Molecular Microbiology*. 56(2): 465–479.
- Ventola, C. L. (2015). The antibiotic resistance crisis: part 1: causes and threats. *P & T: A Peer-Reviewed Journal for Formulary Management*. 40(4): 277–83.
- Wang, G., & Maier, R. J. (2015). Bacterial histone-like proteins: roles in stress resistance. *Current Genetics*. 61(4): 489–492.
- Wang, W., Li, G.-W., Chen, C., Xie, S. X., & Zhuang, X. (2012). Chromosome organization by a nucleoid-associated protein in live bacteria. *Science*. 333(6048): 1445–1449.
- Watve, M. G., Tickoo, R., Jog, M. M., & Bhole, B. D. (2001). How many antibiotics are produced by the genus *Streptomyces*? *Archives of Microbiology*. 176(5): 386–390.
- Wezel, G. P. Van, White, J., Hoogvliet, G., Bibb, M. J., Laboratories, G., & Centre, J. I. (2000). Application of redd, the transcriptional activator gene of the undecylprodigiosin biosynthetic pathway, as a reporter for transcriptional activity in *Streptomyces* further reading. *Journal of Molecular Microbiology*. 2: 551–556.
- Williams, S. T., Bradshaw, R. M., Costerton, J. W., & Forge, A. (1972). Fine structure of the spore sheath of some *Streptomyces* species. *Journal of General Microbiology*, 72(2): 249–258.
- Yang, K., Han, L., & Vining, L. C. (1995). Regulation of jadomycin B production in *Streptomyces venezuelae* ISP5230: Involvement of a repressor gene, jadR2. *Journal of Bacteriology*. 177(21): 6111–6117.

Zhang, Y. Y., Zou, Z. Z., Niu, G. Q., & Tan, H. R. (2013). *jadR** and *jadR2* act synergistically to repress jadomycin biosynthesis. *Science China Life Sciences*. 56(7): 584–590.

Zimmerman, S. B., & Murphy, L. D. (1996). Macromolecular crowding and the mandatory condensation of DNA in bacteria. *Diabetes*. 390: 245–248.