SMALL MOLECULE PROBES FOR USE IN THE STREPTOMYCETES

SMALL MOLECULE INTERROGATION OF *S. COELICOLOR* GROWTH, DEVELOPMENT AND SECONDARY METABOLISM

By

ARRYN CRANEY, B.Sc.

University of Waterloo, 2004

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

DOCTOR OF PHILOSOPHY (2012) McMaster University
(Biochemistry & Biomedical Sciences) Hamilton, Ontario

TITLE: Small Molecule Interrogation of *Streptomyces coelicolor* Growth, Development and Secondary Metabolism

AUTHOR: Arryn Craney, B.Sc. (University of Waterloo)

SUPERVISOR: Professor J.R. Nodwell

NUMBER OF PAGES: xiv, 195

ABSTRACT

Secondary metabolites are vital to human health and strategies to improve their production and detection are equally essential. The blue pigmented metabolite actinorhodin produced by *Streptomyces coelicolor*, a genus renowned for their diverse secondary metabolites, provides a unique opportunity to identify small molecules probes of secondary metabolism. Small molecules capable of altering secondary metabolism will have widespread application in the streptomycetes due to their ease of addition to any culture condition. Taking advantage of the phenotypic versatility of the *S. coelicolor* lifecycle, we extended our search for small molecule modulators further to include the entire developmental process. In addition to alterations in secondary metabolism, these processes include growth inhibition, precocious sporulation and alterations in aerial hyphae formation and sporulation. This work provides the foundation for studying *Streptomyces* by chemical manipulation. Those compounds which stimulate secondary metabolism were narrowed down to 19 ARCs (for antibiotic remodeling compounds). From these, a set of 4 structurally related molecules, the ARC2 series, was identified as weak inhibitors of fatty acid biosynthesis and most likely lead to alterations in secondary metabolism through shifting precursors from primary to secondary metabolism. Consistent with the conservation of fatty acid biosynthesis within bacteria, the effect of the ARC2 series extends in general to the actinomycetes. This provides a simple strategy to alter the secondary metabolic profiles of a diverse range of actinomycetes.

ACKNOWLEDGEMENTS

I would like to thank my supervisor Dr. Justin Nodwell for providing guidance over the years and allowing me the opportunity to explore a vast array of projects. I would also like to acknowledge my committee members Dr. Eric Brown and Dr. Marie Elliot. I thank Dr. Fred Capretta and his student Cory Ozimok for the construction of the library of ARC2 analogues - this was vital for understanding the mechanism of action of the ARC2 series. I thank all the Nodwell lab members past and present as you have helped shape the person I am today. In particular, I thank the “Dream Team” Dr. Sheila Pimentel-Elardo, Salman Ahmed and Vanessa Yoon for being the greatest people to share a project with and Charul Jani for her strong dedication to exploring the white molecules. I thank my family, you may not always understand what I do, but you listen to me anyways. Most of all, I thank my husband Addae Jones for his endless support and encouragement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 4

APPENDIX 1

APPENDIX 2

ABBREVIATIONS

(S)-DNPA	4-dihydro-9-hydroxy-1-methyl-10-oxo-3-H-naptho-[2,3-c]-pyran-3 -(S)-acetic acid
aa	amino acid
ACP	acyl carrier protein
ACT	actinorhodin
ARC	antibiotic remodeling compounds
bld	bald
CCC	Canadian Compound Collection
CDA	calcium-dependent antibiotic
cDNA	complementary DNA
CFU	colony forming units
CLF	chain length factor
CoA	coenzyme A
CPK	cryptic polyketide
cps	counts per second
DHK	dihydrokalafungin
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
dNTPs	deoxynucleotide triphosphates
DTT	dithiothreitol
eGFP	enhanced green fluorescent protein
ESI	electrospray ionization
FAS	fatty acid biosynthesis

GBL	gamma-butyrolactone
HBM	4-hydroxy-2,2'-bipyrrole-5-methanol
HEPES	2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid
HTS	high throughput screen
KS	ketosynthase
LC/MS	liquid chromatography/mass spectrometry
MALDI	matrix-assisted laser desorption/ionization
MBC	4-methoxy-2,2'-bipyrrole-5-carbaldehyde
MCAT	malonyl-CoA:ACP transacylase
MCS	multiple cloning site
MMY	methylenomycin
mRNA	messenger RNA
NADH	nicotinamide adenine dinucleotide
NRP	non-ribosomal peptide
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PIPES	1,4-Piperazinediethanesulfonic acid
PK	polyketide
PKS	polyketide synthase
ppGpp	guanosine tetraphosphate
PPTase	phosphopantetheinyl transferase
PVDF	polyvinylidene difluoride

RED	prodiginines
RNA	ribonucleic acid
RR	response regulator
RT-PCR	reverse transcription polymerase chain reaction
SAM	*s*-adenosyl-L-methionine
SARP	*Streptomyces* antibiotic regulatory protein
SDS	sodium dodecyl sulfate
SK(his)	histidine sensor kinase
SK(ser)	serine/threonine sensor kinase
SRP	signal recognition particle
ssDNA	single stranded DNA
str	streptomycin
TAMP	tail anchored membrane protein
TAT pathway	twin arginine repeat pathway
TCA	tricarboxylic acid cycle
TFR	TetR family regulator
tRNA	transfer RNA
UV	ultraviolet
Wbl	whiB-like
whi	white

CHAPTER 1

INTRODUCTION

1.1 The need for new drugs

Antibiotics are vital to human health. Since the identification of penicillin and streptomycin, antibiotics have been our best defense for staving off bacterial infections. However, their misuse has led to bacterial resistance developing to all currently used antibiotics (1). This rise in antibiotic resistance necessitates the need for new antibiotics with more potent activity and novel modes of action (2).

A source of potential new drug leads is bacterial secondary metabolites. These are small molecules produced by bacteria that, while not essential for growth, offer a selective advantage to the producer. These secondary metabolites have been exploited due to their potent biological activity (bioactive) as antibiotics, chemotherapeutics, immunosuppressants and other clinical applications (3-5). The actinomycete bacteria and in particular, the streptomycetes, are rich in bioactive possibilities (6-8).

Two major classes of secondary metabolites that have been developed extensively as antibiotics are polyketides (8) and non-ribosomal peptides (9). They are both produced by biochemical pathways encoded in discrete islands within the genome. In general, each streptomycete is known for the production of only a few secondary metabolites. However, genome sequencing efforts have revealed that streptomycete genomes encode the potential to produce upwards of 30 of these metabolites (10). This represents a large reservoir of potentially worthwhile metabolites. To tap into this metabolite reservoir it is necessary to

understand factors governing their expression and devise strategies to improve their detection.

1.2 *S. coelicolor* as a model for secondary metabolism production

S. coelicolor has emerged as a model for secondary metabolism production due to the production of two pigmented secondary metabolites, actinorhodin (blue) (11) and the prodiginines (red) (**Figure 1.1**) (12). The ease of visual detection of these pigments has enabled both an understanding of their biosynthesis and the regulation governing their expression (13). The production of these pigments has also aided in the understanding of regulators that are common to other streptomycetes (14). In addition, the *S. coelicolor* genome was the first sequenced of the streptomycetes, allowing for the prediction of 29 secondary metabolites of which 16 have been structurally elucidated (10, 13, 15). This makes the secondary metabolites of *S. coelicolor* the most thoroughly analyzed of the streptomycetes (**Figure 1.1/Table 1.1**).

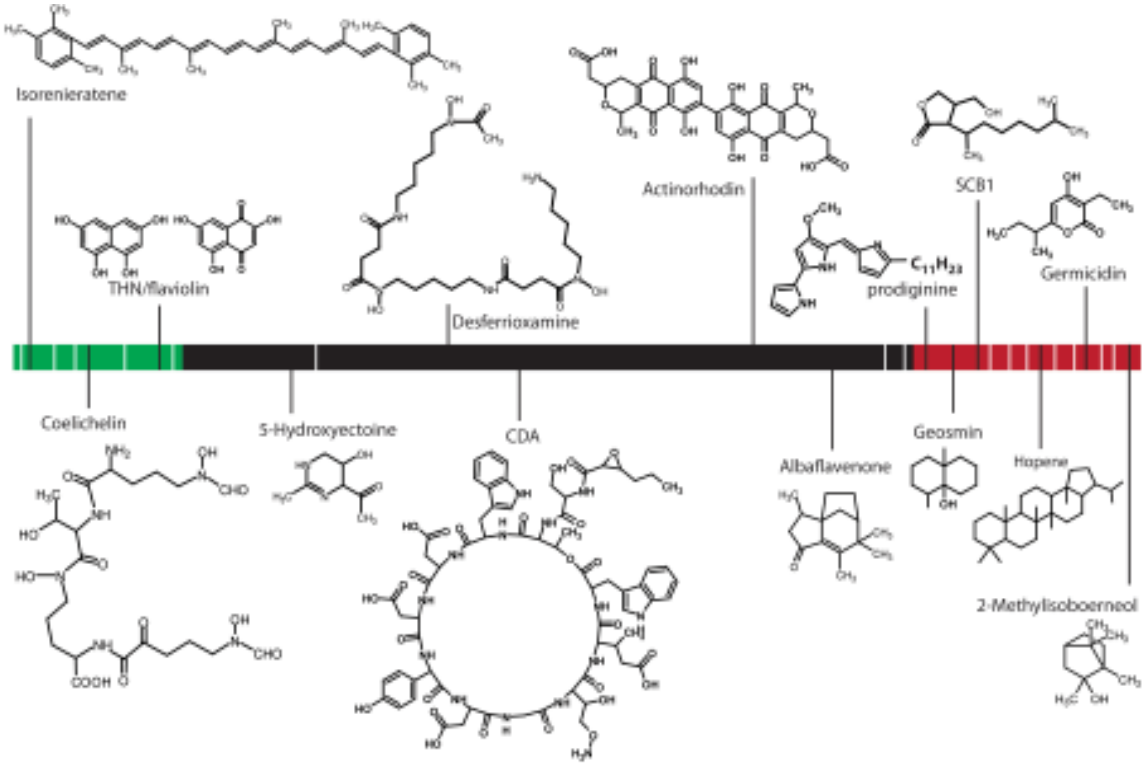

Figure 1.1. *S. coelicolor* secondary metabolites. Secondary metabolites are depicted in relation to their chromosomal location. The left arm is highlighted in green and the right is highlighted in red. Secondary metabolites without structures are shown in white (see **Table 1.1** for details); while those with structures are in black.

Table 1.1. *S. coelicolor* Secondary Metabolites

Secondary Metabolite	Location	Type	Detection Method	Reference
Identified Structures				
Actinorhodin	SCO5071-5092	aromatic PK	blue pigment	(11)
Prodiginine	SCO5877-5898	tripyrrole	red pigment	(12)
CDA	SCO3210-3249	NRP	antimicrobial activity	(16)
Germicidin	SCO7221	type III PK	genome mining	(17)
Geosmin	SCO6073	terpenoid	odour/genome mining	(18)
SCB1	SCO6266	ɣ-butyrolactone	genome mining	(19)
Isorenieratene	SCO0185-0191	terpenoid	blue light induction	(20)
Coelichelin	SCO0489-0499	NRP	genome mining	(21)
THN/Flaviolin	SCO1206-1208	type III PK	genome mining	(22)
5-Hydroxyectoine	SCO1864-1867	cyclic amino acid	salt or high temp	(23)
Desferrioxamine	SCO2782-2785	tris-hydroxymate	genome mining	(24)
Albaflavenone	SCO5222-5223	terpenoid	odour/genome mining	(25)
2-Methylisoborneol	SCO7700-7701	terpenoid	odour/genome mining	(26)
Methylenomycin	SCP.228c-246	cyclopentanoid	antimicrobial activity	(27)
Methyl furans		methylfurans	genome mining	(28)
Developmental Secondary Metabolites				
SapB	SCO6681-6685	lantibiotic	mass spec/phenotype	(29)
Predicted Structures (untested)				
Eicosapentaenoic acid	SCO0124-0129	fatty acid	genome prediction	(10)
Melanin	SCO2700-2701	melanin	genome prediction	(10)
Bacteriocin	SCO0753-0756	bacteriocin	genome prediction	(10)
Coelibactin	SCO7681-7691	NRP	genome prediction	(15)
Unable to predict structures				
Lantibiotic	SCO0267-0270	lantibiotic	genome prediction	(10)
Lantibiotic	SCO6927-6932	lantibiotic	genome prediction	(10)
PKS	SCO1265-1273	aromatic PK	genome prediction	(10)
PKS	SCO6826-6827	aromatic PK	genome prediction	(10)
PKS	SCO7669-7671	type III PK	genome prediction	(10)
Cryptic polyketide	SCO6273-6288	PK	yellow pigment/ genome prediction	(30, 31)
Siderophore	SCO5799-5801	-	genome prediction	(10)
Dipeptide	SCO6429-6438	-	genome prediction	(10)
Deoxysugar	SCO0381-0401	-	genome prediction	(10)
Developmental Secondary Metabolites (unknown structure)				
Grey spore pigment	SCO5314-5320	PK	grey pigment	(32)

1.2.1 Biosynthesis of *S. coelicolor*'s blue pigment, actinorhodin

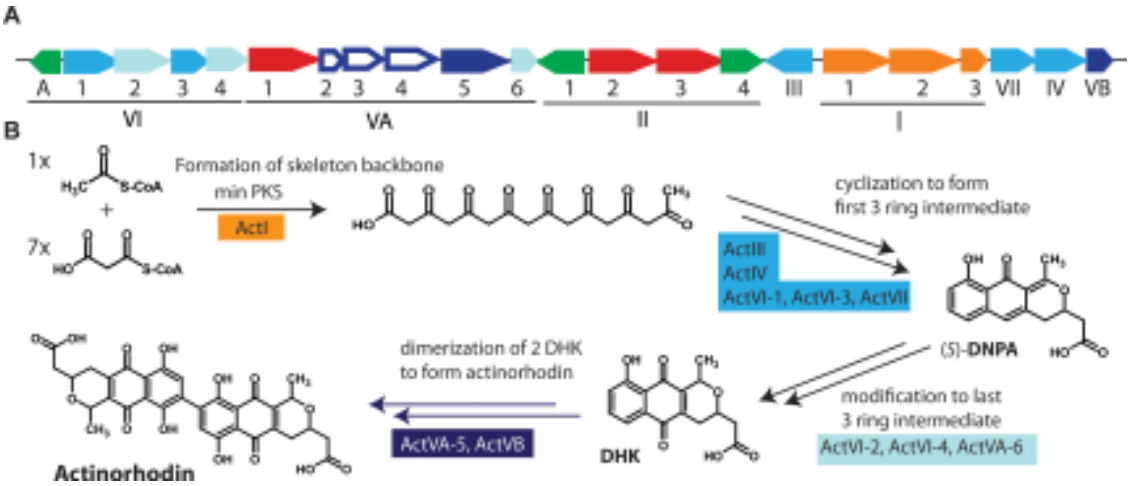

Figure 1.2. Actinorhodin Biosynthesis. (A) Organization of the actinorhodin biosynthetic cluster. Regulatory genes are highlighted in green and putative resistance genes in red. The minimal PKS (ActI) is orange. Tailoring genes are coloured depending on their role in forming actinorhodin. Genes that have not been characterized are filled with white. **(B)** 1 acetyl-CoA and 7 malonyl-CoA are condensed to form the carbon skeleton by ActI. This carbon backbone is cyclized to form a three ring intermediate (*s*)-DNPA (by ActIII, ActVII, ActIV, ACtVI-1 and ActVI-3) followed by modification to DHK (ActVI-2, ActVI-4 and ActVA-6). Dimerization of 2 DHK molecules results in the formation of actinorhodin (by ActVA-5 and ActVB). The involvement ActVA2-4 has yet to be characterized.

Actinorhodin is an aromatic polyketide synthesized from a 22 kb gene cluster encoding all the biosynthetic machinery required for its production (**Figure 1.2**) (33). This includes a pathway specific activator (*act*II-4) which integrates the cellular signals governing actinorhodin expression, resistance genes in the form of efflux pumps (*actII-2*, *actII-3* and *actVA-1*), the minimal polyketide synthase (*actI-orf1/2/3*) which catalyzes the condensation of the carbon skeleton backbone

and tailoring enzymes (the remaining genes in the cluster) responsible for modifying the carbon skeleton to create the final product, actinorhodin (**Figure 1.2**) (33).

Synthesis begins with the formation of the carbon skeleton, created by the condensation of one acetyl-CoA and seven malonyl-CoA precursors supplied from primary metabolism (**Figure 1.2**). This carbon skeleton is created using a type II, or iterative, minimal polyketide synthase (ActI) (34). In iterative condensation, a single set of enzymes is used repeatedly to create the carbon skeleton and consists of only 4 proteins; the acyl carrier protein (ACP – ActI-3), the ketosynthase (KS or KSα – ActI-1), the chain length factor (CLF or KSβ – ActI-2) and malonyl-CoA:ACP transacylase (MCAT). MCAT is responsible for transferring malonyl-CoA precursors to the minimal polyketide synthase. MCAT is borrowed from primary metabolism but, due to its requirement for supplying precursors it is often included in the minimal PKS despite its genomic location being outside the biosynthetic cluster. Incorporation of the initial acetyl-CoA and subsequent 7 malonyl-CoA is carried out by the KS with decarboxylative condensation of each malonyl-CoA unit resulting in the addition of 2 carbon units to the growing chain. The ACP serves as the tether for the growing carbon chain. The length of the carbon skeleton is determined by the CLF which controls the number of condensation reactions; in the case of actinorhodin the length is 16 carbons, corresponding to 8 condensation reactions. Once synthesized, the

carbon skeleton undergoes extensive modification resulting in the cyclization of the backbone into a three ring structure (*S*)-DNPA, followed by modification to dihydrokalafungin (DHK). Finally, two subunits of DHK are dimerized to create the final product, actinorhodin (**Figure 1.2**) (33).

1.2.2 Biosynthesis of *S. coelicolor*'s red pigment, the prodiginines

The red, cell wall-associated pigment produced by *S. coelicolor* is a mixture of prodiginines – undecylprodiginine and the cyclized derivative streptorubin B being the major components (35). Prodiginines are a widespread and structurally related group of tripyrrole antibiotics. Their biosynthesis in *S. coelicolor* is directed by a 30 kb gene cluster, which encodes the proteins involved in regulating its expression (pathway specific activators, *redD* and *redZ*) and the enzymes required for most of its synthesis; however, creation of the lipid region requires enzymes from fatty acid biosynthesis (**Figure 1.3**) (36).

Prodiginine synthesis is a bifurcated process as the final tripyrrole prodiginine products result from the condensation of a dipyrrole moiety (4-methoxy-2,2'-bipyrrole-5-carbaldehyde or MBC) with a monopyrrole moiety (2-undecylpyrrole) created by two separate biosynthetic pathways (**Figure 1.3**). The dipyrrole MBC is created using proline and serine with the incorporation of each resulting in the formation of a pyrrole group. First, proline is modified to pyrrolyl-2-carboxyl by the incorporation of two double bonds in its ring structure (catalyzed by RedMOW). Pyrrolyl-2-carboxyl is subsequently transferred to RedX followed by

decarboxylative condensation of a malonyl unit by RedN. RedN is predicted to also incorporate serine to form the second pyrrole ring, creating 4-hydroxy-2,2'-bipyrrole-5-methanol (HBM). HBM undergoes oxidation by RedV and methylation by RedFI to create the final product MBC (**Figure 1.3)** (36). 2-undecylpyrrole synthesis begins with the production of a 12 carbon lipid, tethered to the ACP, RedQ. Lipid formation requires RedP (ketosynthase that initiates chain formation), RedR (ketosynthase required to elongate the chain) and the fatty acid biosynthetic machinery. The acyl chain is transferred to the ACP domain of RedL. RedL is a multidomain protein encoding a ketosynthase domain for incorporation of another malonyl-CoA and an α-oxoamine synthase domain for glycine addition. Upon release of this intermediate from RedL, RedK catalyzes the final steps (reduction and dehydration of the intermediate) to create the monopyrrole, 2-undecylpyrrole (37). One unit of MBC and one unit of 2-undecylpyrrole are condensed by RedH to form undecylprodiginine. Approximately 1/3 of the undecylprodiginine undergoes oxidative cyclization performed by RedG to create the cyclized prodiginine, streptorubin B (**Figure 1.3**) (38).

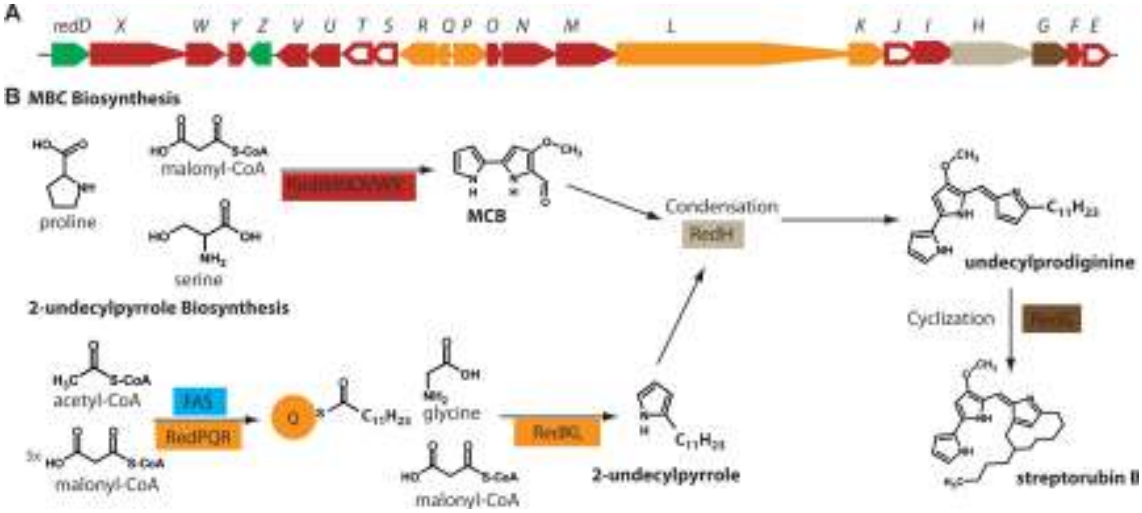

Figure 1.3. Prodiginine Biosynthesis. (**A**) Organization of the prodiginine biosynthetic cluster. Regulatory genes are in green, genes for MBC synthesis are red and genes for 2-undecylpyrrole are orange. Genes for condensation of MBC and 2-undecylpyrrole and subsequent cyclization are grey and brown, respectively. Genes with unknown function have white centers. **(B)** Prodiginine synthesis requires the production of a dipyrrole, MBC, and a monopyrrole, 2 undecylpyrrole, from separate enzymatic reactions which are subsequently condensed together to form the final tripyrrole. MBC synthesis requires proline, malonyl-CoA and serine as substrates and is catalyzed by RedMNOVWX. 2-undecylpyrrole begins with the formation of a 12 carbon lipid, synthesized by RedPQR with the aid of the enzymes from fatty acid biosynthesis (FAS). This lipid is transferred to RedL where glycine and another malonyl-CoA are added to the chain. Once released from RedL, RedK performs the final modifications to form 2-undecylpyrrole. MBC and 2-undecylpyrrole are condensed by RedH to form undecylprodiginine. Further cyclization by RedG occurs to ~1/3 to form streptorubin B.

1.3 Complex regulatory network governing secondary metabolism

While the regulatory network controlling secondary metabolism in the streptomycetes has yet to be fully elucidated the factors controlling secondary metabolism include sensing nutrients levels, environmental stresses and growth rate. To cope with responding to these diverse signals, the cell has evolved a complex regulatory network to ensure energy is not wasted and the desired end product can be achieved. This network can be divided into two layers of regulation – pathway specific and global.

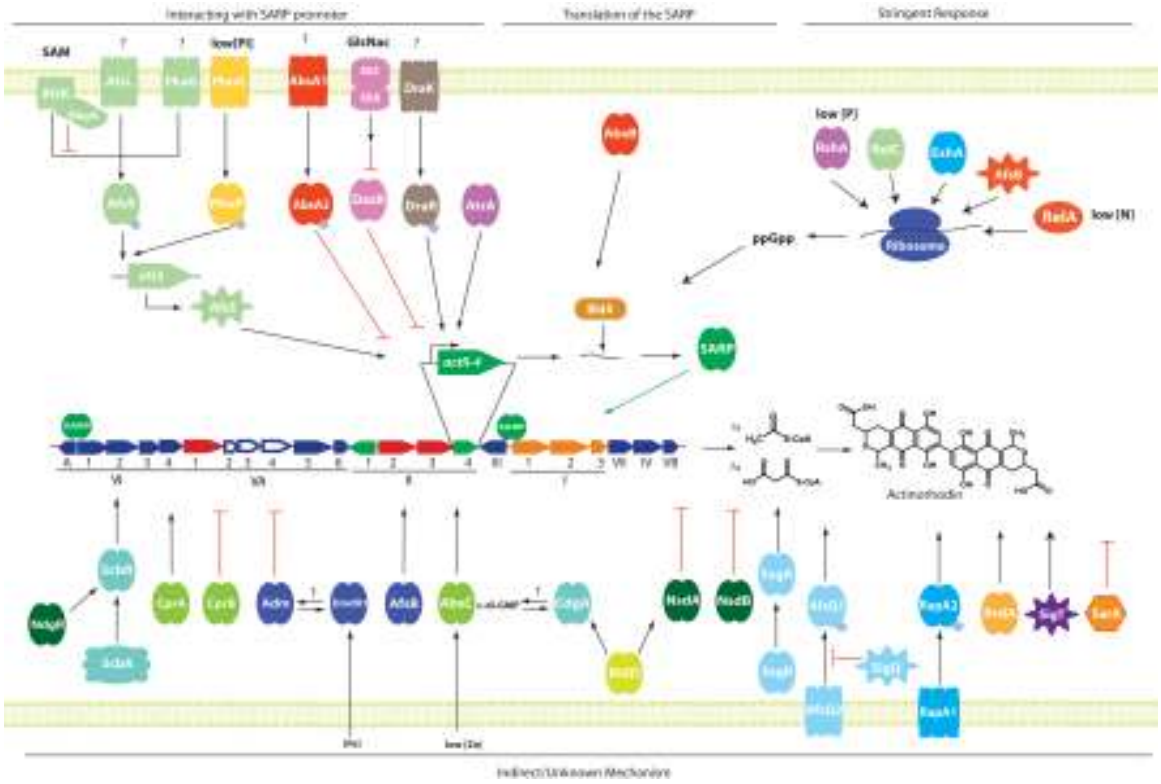

Figure 1.4. Complex regulation of actinorhodin production in *S. coelicolor*. The pathway specific activator ActII-4 integrates many of the global regulators. Repression in the network is illustrated in red and activation in black. Regulators with experimental evidence of binding to the promoter region of *actII*-4 are grouped at the top; while regulators without direct evidence or which do not act via ActII-4 are grouped on the bottom. BldA acts at the level of translation as it is a developmentally regulated tRNA encoding for leucine TTA codon (*actII-4* contains 1 TTA codon). Ribosomal modification improves expression of the actinorhodin cluster through the stringent response (ppGpp synthesis). A complete list of regulators can be seen in **Table 1.3**.

1.3.1 Pathway Specific Regulators

Regulation of individual biosynthetic clusters often occurs by the controlled expression of a pathway specific activator embedded within the cluster (**Table 1.2**). The most common type of pathway specific regulator is the SARP (*Streptomyces* antibiotic regulatory protein) family, characterized by a winged helix-turn-helix motif at their N-terminus (39, 40). These regulators integrate the global signals for the control of a particular metabolite. Regulation can occur by direct binding to the promoter controlling transcription as demonstrated by AbsA2 (41), by controlling translation as demonstrated by BldA (42) or by indirect mechanisms which improved SARP expression without evidence of binding to the promoter region of the SARP as in the case of ppGpp synthesis (43) and the deletion of *scbA* (**Figure 1.4**) (44).

Table 1.2. Pathway specific regulators in *S. coelicolor*

Regulator	Type	Metabolite	Reference
ActII-4	SARP	actinorhodin	(45)
RedD	SARP	prodiginine	(46)
RedZ	RR	prodiginine	(47)
EcrA1/A2	SK(his)/RR	prodiginine	(48)
CdaR	SARP	CDA	(49)
CpkO/KasO	RR	CPK	(50)
ScbR	TFR	CPK	(50)
MmyR	TFR	methylenomycin	(51)
MmfR	TFR	methylenomycin	(51)
MmfB	Xre-like	methylenomycin	(51)

1.3.2 Global Regulators

Regulators that function to alter the production of many secondary metabolites are referred to as global, or pleiotropic, regulators. *S. coelicolor*'s pigment production has led to the identification of a vast array of the global regulators. These regulators sense and respond to the cellular environment, competing in their regulation of secondary metabolites as some activate expression, while others repress (**Table 1.3** and **Figure 1.4**).

One of the best characterized global regulators is the AfsK/R/S system, a global regulator of actinorhodin and prodiginine production (**Figure 1.4**). AfsK is a serine/threonine sensor kinase, found loosely associated with the membrane (52). To date the only activating signal identified for AfsK is *s*-adenosyl-L-methionine (SAM) (53, 54). SAM is an important metabolite as it is the major methyl donor in all living organisms and is essential for many biological processes, including providing the methyl group during prodiginine synthesis (final step of MBC synthesis) (36). Signals sensed by AfsK result in autophosphorylation which is modulated by KbpA (AfsK's ability to autophosphorylate is inhibited when bound to KpbA) (55). AfsR is subsequently phosphorylated by AfsK (56). AfsR can also be phosphorylated by two other sensor kinases, AfsL and PkaG (57); however, signalling in these cases is unknown. The resulting active form AfsR~P is capable of binding DNA. The target of AfsR~P is the promoter region of *afsS*, binding to the -35 region and

promoting *afsS* transcription (58). Expression of AfsS, a small sigma-like protein, results in increased expression of the pathway specific activators for actinorhodin and prodiginine; ultimately increasing yields of these metabolites (59).

AfsS expression is also regulated by phosphate limitation (**Figure 1.4**) (60). Phosphate limitation is sensed by the sensor kinase PhoR which autophosphorylates, and transfers the phosphate to the response regulator PhoP. Responding to limited phosphate, PhoP~P activates expression of the genes in the *pho* regulon by binding to the PHO boxes within their promoter regions. *afsS* is a member of the Pho regulon, and expression of AfsS is increased in phosphate limiting conditions, resulting in increased production of actinorhodin and prodiginine (60).

The AfsK/R/S system demonstrates the true complexity in eliciting secondary metabolite production as AfsS and AfsR are both integration points for multiple signals and involve at least 2 signaling molecule (SAM and phosphate). Despite this system being one of the best characterized global regulators, many details are still required for complete understanding of regulation by this system and the numerous other global regulators.

1.3.3 Gamma-butyrolactones – both pathway specific and global regulators

Signalling molecules play an important role in regulating secondary metabolism. In particular, γ-butyrolactones (GBL) have a widespread presence in the streptomycetes. Their effects can be pathway specific or global depending on the evolved regulatory network for the strain. For example, in *S. griseus*, both sporulation and secondary metabolism are controlled by the production of a single GBL, A-factor, making it a global regulator (61). In *S. avermitilis*, its GBL avenolide is responsible for eliciting only the production of the avermectins and thus is a pathway specific regulator (62). The GBL of *S. lavendulae* IM-2 is also a global regulator but has differing effects on secondary metabolism; GBL production reduces D-cycloserine and increases nucleoside antibiotics and a blue pigment (63).

In a few cases the regulation is less clear, in *S. coelicolor*, the GBL SCB1, synthesized by ScbA is a pathway specific regulator of CPK (50). However, deletion of ScbA causes a strong stimulation of both prodiginine and actinorhodin through an unknown mechanism, making ScbA a global regulator (64) (**Table 1.3** lists ScbA as a global regulator; while in **Table 1.2**, ScbR the receptor for SCB1 is a pathway specific regulator). Many streptomycetes encode multiple γ-butyrolactone synthases; however, regulation by multiple GBLs has yet to be explored.

Table 1.3. Regulators involved in *S. coelicolor* secondary metabolism

Regulator	Type	Notes	Reference
Multi Component Systems			
afsK regulon			
AsfK	SK(ser)	Phosphorylation stimulated by SAM	(53, 54)
		Phosphorylates AfsR	(52)
		Δ*asfK* ↓ACT	(52)
AfsL	SK(ser)	Phosphorylates AfsR	(57)
PkaG	SK(ser)	Phosphorylates AfsR	(57)
AfsR	RR(SARP-like)	Integrates signals from AsfK/AfsL/PkaG	(57)
		Δ*asfR* ↓ACT	(65)
KpbA		Represses AfsK phosphorylation	(55)
		Δ*kpbA* ↑ACT	
AfsS/AfsR2	σ-like	Δ*asfS* ↓ACT	(59)
		Integrates signals from AfsR~P and PhoR~P	
Pho regulon			
PhoR	SK(his)	Phosphorylated in low phosphate	(60)
		Phosphorylates PhoP	
PhoP	RR	PhoP~R binds promoters of *pho* regulon	(60)
		afsS is part of the *pho* regulon	
		↑ACT/RED in ↓P_i	
absA regulon			
AbsA1	SK(his)	Phosphorylates AbsA2	(66)
		Δ*absA1* ↑ACT/RED	(67)
		↑*absA1*↓ACT/RED	(68)
AbsA2	RR	AbsA2~P global repressor	(41)
		Directly binds to pathway specific promoter	
		Δ*absA2* ↑ACT/RED	(68)
afsQ regulon			
AfsQ2	SK(his)	↑afsQ1/Q2↑ACT/RED/CDA in *S. lividans*	(69)
AfsQ1	RR	MM+glutamate required to see mutant phenotypes	
		Δ*afsQ1/Q2* ↓ACT/RED/CDA	
σ^Q	sigma factor	Antagonizes AfsQ1-Q2	
RapA1/A2	RR/ SK(his)	Δ*rapA1*/A2 ↓ACT/CPK=RED	(70)

DmdR1/Adm	DmdR	Antiparallel overlapping genes Senses Fe2+ Δ*dmdR1/adm* ↓ACT/RED Δ*adm* ↑ACT/RED Δ*dmdR1* =ACT/RED	(71)
DraK/R	SK(his)/RR	Δ*draK/R* ↓ACT/↑RED/CPK Directly with pathway specific promoter ACT/CPK Indirect effect on RED	(72)
One component Systems			
AtrA	TFR	Δ*atrA* ↓ACT	(73)
DasR	GntR	DasR represses pathway specific promoters GlcNac relieves DasR repression	(74)
RrdA	TFR	Δ*rrdA* ↑RED ↓ACT ↑*rrdA* ↓RED ↑ACT	(75)
AbsC	MarR-like	Δ*absC* + low Zn no ACT/RED	(76)
CprA	TFR	Δ*cprA* ↓ACT/RED ↑*cprA* ↑ACT/RED	(77)
CprB	TFR	Δ*cprB* ↑ACT = RED	(77)
NdgR	IclR-like	Δ*ndgR* ↑ACT Binds intergenic region of ScbA/R	(78)
BldD regulated			
BldD	DNA binding	Δ*bldD*↓ACT	(79, 80)
SsgR/SsgA	SALP	Δ*ssgA*↓ACT *ssgR* positively regulated ssgA	(81) (82)
NsdA	TPR-like	Δ*nsdA* ↑ACT/CDA/MMY	(83)
NsdB	TPR-like	Δ*nsdB* ↑ACT/CDA	(84)
CdgA	c-di-GMP	↑*cdgA* ↓ACT	(79)
CdgB	c-di-GMP	↑*cdgB* ↓ACT	(85)
Translation			
AdpA/bldA	tRNA –leu	Developmentally regulated Prevents translation of genes with TTA codons until expressed *redD*, *actII*-4 and *cdaR* contain TTA	(42)
AbsB	RNAseIII	Δ*absB* ↓ACT/RED Degrades double stranded mRNA	(86-88)

Stringent Response			
	**ppGpp synthesis ↑ACT/RED		
RelA	ppGpp	Nitrogen limited ppGpp production	(89, 90)
RshA	ppGpp	Phosphate limited ppGpp production	(91)
RelC/RplK	ppGpp	Mutations reduce ppGpp production	(92)
EshA	ppGpp	Δ*eshA* reduce ppGpp production	(93)
AfsB/HrdB	sigma factor	Mutations interfere with ppGpp synthesis resulting in lowering pigment production	(94)
Miscellaneous			
ScbA	GBL	Δ*scbA* ↑ACT/RED	(44, 64)

1.4 Strategies to improve secondary metabolism production and detection

Many strategies have been taken in order to tap into the diversity of secondary metabolites within the streptomycetes. These strategies involve altering global processes in the cell (cell-based), manipulating individual biosynthetic clusters (cluster-based), or through bioinformatic analysis of the genome (genome-based) (**Figure 1.5**).

1.4.1 Cell-based methods

Cell-based methods work by altering processes in the cell generally (not the biosynthetic cluster), resulting in the improved production of secondary metabolites. These can be non-invasive such as mutagenesis, changing media components, engineering ribosomes and using small molecules, or invasive, involving genetic engineering of the strain (**Figure 1.5A**).

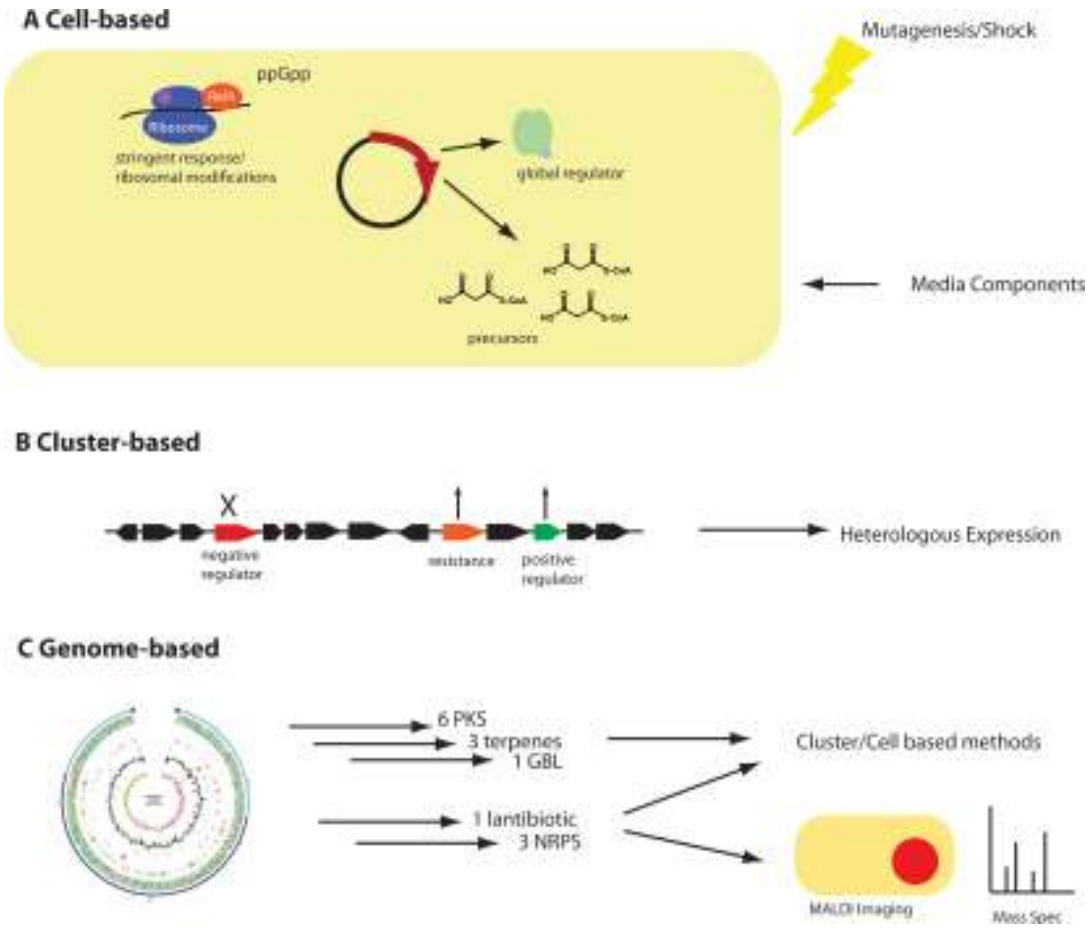

Figure 1.5. Strategies to improve secondary metabolism. (**A**) Cell-based methods include eliciting the stringent response (through ribosomal modifications), over-expressing global regulators or precursor metabolites, changing media composition or by stressing the cell through mutagenesis or shock. (**B**) Cluster-based methods involve either over-expression of a pathway specific activator or resistance determinant, deletion of a pathway specific repressor, or heterologous expression of the biosynthetic cluster. (**C**) Genome sequences can be analyzed for secondary metabolites based on homology to genes known for the production of different classes of metabolites. Following genomic identification cell-based and cluster-based methods can be used for detection and structural elucidation.

Classic Methods (media components, shock and mutagenesis)

A simple, but effective method for expressing secondary metabolites is changing the growth medium and culture conditions. While difficult to predict the outcome, changing media conditions, shock or growth at high temperature can elicit secondary metabolite production in some strains. Daptomycin, the clinically relevant form of a family of lipopeptides produced by *S. roseosporus* requires feeding decanoic acid to ensure that daptomycin is the major product (95). Jadomycin B production in *S. venezulae* requires induction by ethanol shock or growth at high temperature (42^{o}C) and can be improved by combining these two growth conditions (96, 97).

Many secondary metabolites offer selective advantages to the producers and are only produced during specific conditions. Siderophores are secondary metabolites that sequester iron and are expressed in low iron conditions (21, 24, 98). The carotenoids of *S. coelicolor* are expressed in the presence of blue light, consistent with their protective role against photodamage (20). Production of ectoine and 5-hydroxyectoine protect against dehydration and thus are expressed under high salt or high temperature conditions in *S. coelicolor* (23).

Mutagenesis is a simple technique that can also be used to improve or alter secondary metabolite production. It is most often used once a desired metabolite has been identified and the producer organism is being optimized for industrial production, but may also have application in inducing new secondary

metabolites. The producer is subject to rounds of mutagenesis involving either UV or chemical mutagens, with surviving cells screened for improved activity. For example, yields of the clavulanic acid from *S. clavuligerus* (99) and rapamycin from *S. hygroscopicus* (100) have both been improved through mutagenesis.

Ribosomal Engineering

Modifications to the ribosome that result in stringent response induction have been demonstrated to improve production of many secondary metabolites. The stringent response mediates the cell's response to harsh environments and is regulated by the production of ppGpp (guanosine tetraphosphate) which occurs when translation stalls (uncharged tRNA at the A site of the ribosome) during amino acids limitation (101). ppGpp synthesis and the activation of the stringent response results in changes in gene expression, including increased production of secondary metabolites in the streptomycetes (90). The stringent response can be mimicked by developing resistance to antibiotics that target the ribosome (e.g. streptomycin, rifampicin, paromycin and gentamicin). The resulting mutations generally map to the ribosomal protein S12 of the 30S subunit, *rpsL,* (89) or in the case of rifampicin, to the β subunit of RNA polymerase (102). The effects of these mutations can be combined for increased effects on secondary metabolism and has been demonstrated by developing stepwise resistance of up to eight ribosomal antibiotics in *S. coelicolor* with a concomitant increase in production of

actinorhodin (103). One advantage to this method of strain improvement is that there is no requirement for genetic engineering. Resistant mutants arise at a frequency conducive to improving secondary metabolism in a wide-range of strains making ribosomal engineering amenable to high throughput analysis (104).

Small molecule manipulation

To date, only one synthetic small molecule has been demonstrated to improve secondary metabolism, an inhibitor of phosphopantetheinyl transferase (PPTase) (105). PPTases activate the ACP of fatty acid biosynthesis and secondary metabolism by "priming" the carrier protein. Priming occurs by the addition of a phosphopantetheinyl group to the ACP, providing the reactive sulfhydryl group that tethers the new metabolites and is essential for these processes to occur. A PPTase inhibitor developed against *Bacillus subtilis* resulted in improved actinorhodin production when added to *S. coelicolor*. The mechanism for this increase is unclear as *S. coelicolor* encodes many PPTases (105). Two possible mechanisms have been proposed: inhibition of fatty acid biosynthesis improves precursors for actinorhodin production, or a stress response from addition of the PPTase inhibitor results in improved actinorhodin production (105). However, there is no experimental evidence to support either mechanism.

Metabolic Engineering

Metabolic engineering differs from the classic method of altering media conditions as it involves genetically modifying the organism to produce more of certain metabolic precursors. Acyl-CoA precursors are important for the production of a number of secondary metabolites; acetyl-CoA, malonyl-CoA and methylmalonyl-CoA are common building blocks of polyketide synthesis and the production of these metabolites often require the primary metabolic pool for precursors. This precursor pool can be improved by manipulating the biochemical pathways which produce or consume them; fatty acid biosynthesis, fatty acid degradation, branched chain amino acid degradation, and glucose metabolism are prevalent examples of these pathways (106). For example, over-expression of methylmalonyl-CoA mutase pathway (*mutAB*) improves production of methylmalonyl-CoA (by isomerization of succinyl-CoA from the TCA cycle) and results in improved production of FK606 in *S. clavuligerus* (107). Disruption of *zwf1 or zwf2* from the pentose phosphate pathway improves production of acetyl-CoA and malonyl-CoA, resulting in increased production of actinorhodin in *S. coelicolor* (108) and oxytetracycline in *S. ambofaciens* (109).

Engineering global regulators

Strains can be genetically engineered to over-express global regulators and elicit overall changes in secondary metabolites within the host organism or when heterologously expressed in other streptomycetes (110). For example, over-expression of a mutant allele of AbsA2 in which its repressive function has been eliminated, results in overproduction of actinorhodin, prodiginines and CDA in *S. coelicolor* (111). This same allele has the capacity to enhance secondary metabolites in other streptomycetes, demonstrated by increased production of streptomycin in *S. griseus* and blasticidin S in *S. griseochromogenes*. As well, introduction into *S. flavopersicus* resulted in production of pulvomycin, previously unreported in this strain and undetectable in the absence of the AbsA2 mutant allele (111).

1.4.2 Cluster-based methods

The biosynthetic gene cluster for a desired metabolite can be altered to improve its expression. Genes that enhance production can be over-expressed (resistance genes and pathway specific activators) and those genes that repress production can be deleted. As well, clusters can be moved to alternate hosts for improved expression (heterologous expression) **(Figure 1.5B)**.

Engineering Self-Resistance

Self-resistance is an important factor in the production of secondary metabolites with antimicrobial properties. Increasing the expression of the resistance mechanism can improve production of the metabolite by removing feedback inhibition and reducing toxicity in the presence of the metabolite. This strategy has been successful in improving production of doxorubicin and daunorubicin by over-expressing the resistance genes *drrABC* in the producer *S. peucetius* (112), production of avermectin by over-expressing AvtAB in *S. avermitilis* (113) and production of actinorhodin by over-expressing ActAB *S. coelicolor* (Ye Xu and Justin Nodwell, manuscript in preparation).

Regulatory Engineering

Since many biosynthetic clusters encode pathway-specific activators, over-expression of these activators can improve production of a desired metabolite and induce expression of cryptic clusters. Over-expression of the pathway specific activators for actinorhodin and prodiginine in *S. coelicolor* improve production of their respective metabolites (114); as was also seen for AveR and avermectin production in *S. avermitilis* (115) and StrR in Streptomyces production in *S. griseus* (116) also improve production of their metabolite when over-expressed. Recently, over-expression of the predicted pathway specific activator, SamR0484 resulted in activation of a silent cluster resulting in the

production of stambomycin A-D, a family of 51-membered glycosylated macrolides in *S. ambofaciens* (117).

Often biosynthetic clusters encode pathway specific repressors which when deleted improve production. Deletion of the pathway specific repressor *cmmRII* in *S. griseus* results in overproduction of chromomycin (118) and deletion of AlpW in *S. ambofaciens* results in constitutive expression of alpomycin (119).

Heterologous expression

Heterologous expression of biosynthetic clusters has been successful for improving production of many secondary metabolites and many strains have been used as hosts for over-expression (**Table 1.4**). *S. lividans* and *S. albus* J0174 were originally the main hosts for heterologous metabolite expression due to their low expression of endogenous secondary metabolites and ease of introduction of exogenous DNA. *S. albus* J1074 is a mutant strain of *S. albus* G which has lost its restriction system and as a result can efficiently take up foreign DNA (**Table 1.4**). *S. venezulae* has been engineered for improved heterologous expression of flavonoids by improving precursor availability and removing endogenous metabolites (**Table 1.4**).

Development of an empty host, or chassis, for expressing a given metabolite offers an improved strategy for heterologous expression. By removing competing biosynthetic clusters the cell puts more resources to the production of the

heterologous metabolite. As well, detection of the metabolite is easier in a strain devoid of other secondary metabolites. Two new chassis have been recently created using *S. avermitilis* and *S. coelicolor*. These strains were chosen as starting strains as they already produce high yields of their respective secondary metabolites and thus are optimized for production. The *S. avermitilis* chassis was created by removing approximately 1.4 Mb of its 9.02 Mb genome containing two of *S. avermitilis*' major secondary metabolites, the avermectins and filipin. Oligomycin and terpene biosynthetic genes were also removed in subsequent deletion strains (SUKA17). Heterologous expression of streptomycin resulted in a 4 fold increase in expression compared to expression in wild type *S. avermitilis* (**Table 1.5**) (123).

A similar approach was taken in creating the *S. coelicolor* chassis in which the prominent biosynthetic clusters (actinorhodin, prodiginine, CDA and CPK) were selectively deleted, resulting in the loss of 1.73 Mb. Mutations in RpoB [C1298T] and RpsL[A262G] were added to improve secondary metabolite production creating the *S. coelicolor* chassis M1154. M1154 was tested using heterologous expression of chloramphenicol from *S. venezuelae* and congocidine from *S. ambofaciens* with production 20-40 times greater than in wild type *S. coelicolor* M145 (120).

Table 1.4. Heterologous Hosts

Strain	Relevant Modifications	Reference
S. coelicolor		
M1154	Δ*act*/*red*/*cda*/*cpk* *rpoB[C1298T]rpsL[A262G]*	(120)
M512	Δ*redD* /*act*II-4	(121)
CH999	Δ*act*	(122)
S. avermitilis		
SUKA17	Δ*sav71-1286(ave/pte)/olm* Δterpenes (sav2161-2168, *sav2990-3002*, *sav7456-7491)*	(123)
S. venezuelae		
DHS2001	Δ*pikAI–IV*	(124)
YJ309	DHS2001 + *matBC*	(125)
S. albus		
J1074	Δ*salGI* (restriction system)	(126)
S. lividans		
66/1326	wild type	(127)
TK24	1326 derivative *str*^{R}	(127)

1.4.3 Genome-based Methods

Perhaps the most successful approach to identifying new secondary metabolites has been through mining genomes for metabolites (**Figure 1.5C**). While there is a great structural diversity in secondary metabolites, the main enzymes for their production are highly conserved making it possible to identify the presence of the classes of metabolites within the genome (21, 128, 129). While genome mining can identify biosynthetic cluster by scanning for these conserved regions, structural elucidation can be difficult depending on the metabolite class. Non-iterative assembly which occurs in non-ribosomal peptide synthesis and type I polyketides biosynthesis can be predicted as each enzymatic reaction is carried

out by an individual module (130-132). Iterative processes such as type II and type III polyketides production are difficult to predict as the carbon skeleton is created by a single module which repeatedly adds carbon units, followed by tailoring of the skeleton backbone (133). The genomes of many streptomycetes are now available and have been mined for their secondary metabolites; *S. coelicolor* is predicted to encode 29, *S. avermitilis* 37, and *S. griseus* 36 potential secondary metabolites (134). Drawbacks to this method are that it requires a sequenced genome and while the presence of a metabolite can be determined through bioinformatics, biochemical analysis is still required for structural elucidation.

The first predicted metabolite through genome mining was the non-ribosomal peptide siderophore, coelichelin in *S. coelicolor* (21). Genome prediction aided greatly in the structural elucidation of coelichelin, as it suggested culture conditions and detection methods. Its prediction as a siderophore (due to the presence of hydroxamic acid functional groups) meant it would have best expression in iron-deficient media. While a complete prediction could not be made from the genomic information, accurate prediction of the substrate specificity was achieved, and paved the way for genome mining for non-ribosomal peptide clusters.

A recent advancement in genome-based structural elucidation is the aid of mass spectrometry (135). This method is well suited for peptide natural products

(ribosomal and non-ribosomal peptides) as structural prediction by genome analysis can be matched to the masses detected by mass spectrometry. Detection of metabolites can occur by either analysis of extracted metabolites by mass spectrometry or by direct detection of the growing strain using MALDI imaging (135, 136). The mass of interest is fragmented and these fragments, or sequence tags, are used to deduce the identity of the amino acid sequence and ultimately, the identity of the peptide natural product (135). While this technique shows particular promise in identifying peptide natural products; it is limited to peptide natural products and requires expression of the metabolite for detection. Initial proof of principle work demonstrated correct predictions for the previously identified ribosomal peptide AmfS from *S. griseus*, the non-ribosomal peptide stendomycin from *S. hygroscopicus* and nine new ribosomal proteins including their biosynthetic clusters (135). As well, the production of the non-ribosomal peptide arylomycin was detected in the daptomycin producer *S. roseosporus* using imaging mass spectrometry (136).

1.4.4 Need for new detection methods

The variation in structural, nutrient and regulatory requirements for secondary metabolites in the streptomycetes requires a multitude of methods in order to thoroughly explore the available chemical diversity. The various strategies currently available have been vital for improving the detection and yields of secondary metabolites. However, there are still many more metabolites to be identified, so new methods are always needed.

1.5 A chemical biology approach to understanding and exploiting secondary metabolism

The use of chemistry to understand biological processes is referred to as chemical biology or chemical genetics. It is similar to classical genetic techniques except manipulation of the cell occurs at the level of the protein by small molecules. Instead of manipulating the gene itself the small molecule acts like a mutation to alter protein function (**Figure 1.6A**) (137-139).

To identify small molecules capable of altering protein function, large libraries of small molecules are screened for the desired effect. These small molecule libraries can be made up of synthetic molecules created by chemical synthesis, natural products created by bacterial secondary metabolism or a mixture of both (138). Screening of these libraries can be performed in a forward or reverse chemical genetic approach. Akin to classic genetics, forward chemical screens

are phenotype-based and require follow up work to identify the target protein; while reverse chemical screens are protein or process based and require follow up to determine *in vivo* effects (**Figure 1.6B**) (137).

There are significant advantages to a chemical biological strategy for understanding cellular processes (140). The small molecule effect is titratable, which can be difficult to achieve genetically. Small molecules act quickly and often reversible, exerting their effect on a timescale that is much faster than genetic changes, allowing less chance of cellular complementation by compensatory pathways and may result in a more pronounced effect. Small molecule manipulation also negates the need for genetic manipulation which is essential for organisms that are difficult to study genetically. Also, once identified, these small molecule effectors can be easily applied to other organisms to assess their effect on the same process in a wide range of organisms.

A chemical biological approach for studying secondary metabolism in the streptomycetes can be of benefit as genetic manipulation in the streptomycetes is exceedingly difficult. For example, there are extensive restriction barriers that can prevent the introduction of foreign DNA (141-146). So, while information learned from *S. coelicolor* and related streptomycetes may be applicable generally, difficulty lies in applying this genetically to a wide range of streptomycetes, especially in high throughput. Small molecules which alter

secondary metabolism can be immediately applied to all streptomycetes to determine their effects without genetic manipulation. As well, identifying novel bioactive metabolites within the streptomycetes requires high throughput screening to which small molecule manipulation is readily amenable. A chemical biological approach can be used to identify small molecule modulators of secondary metabolism that can be applied in a wide range of streptomycetes to improved secondary metabolism.

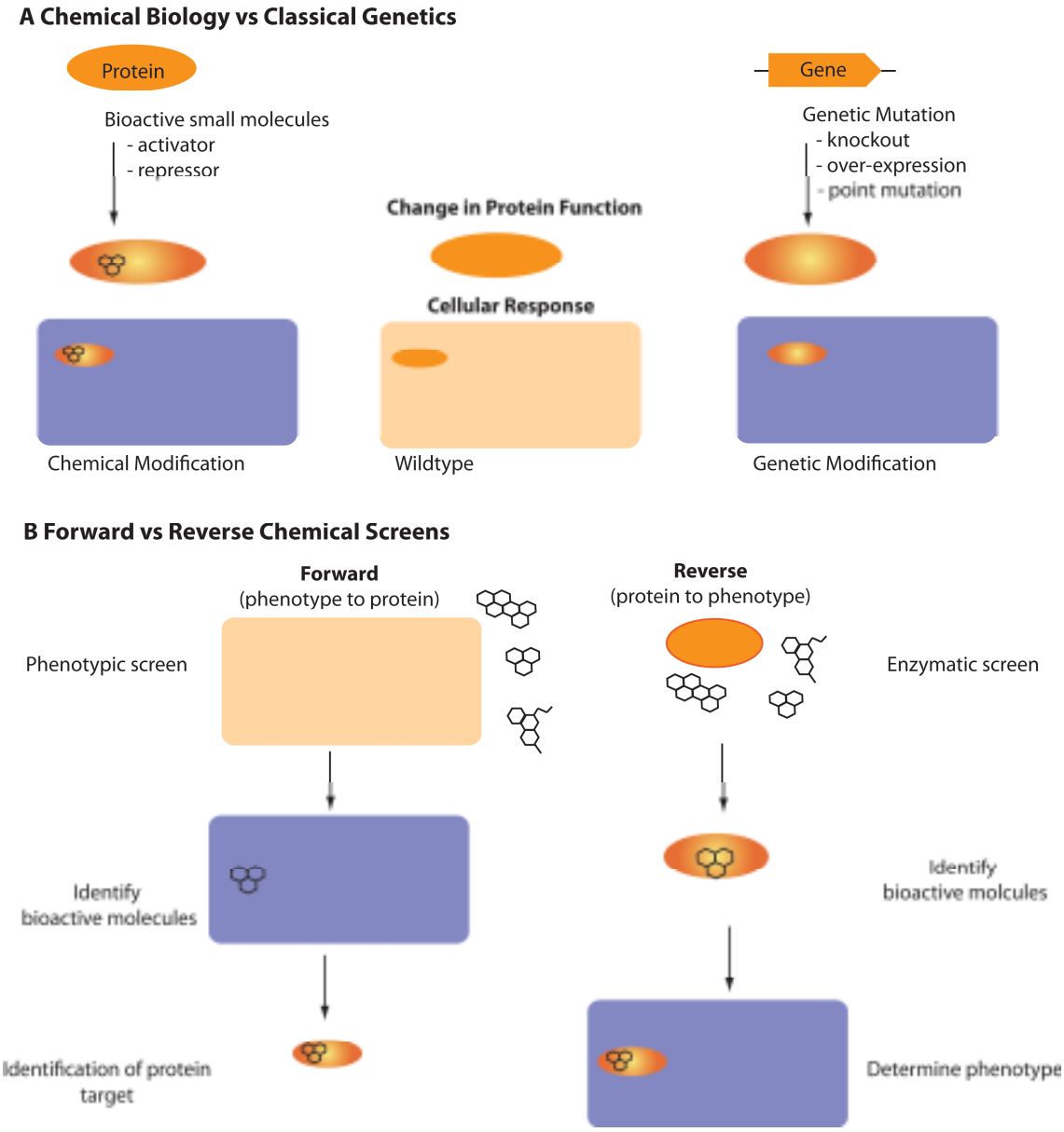

Figure 1.6. An overview of chemical biology. **(A)** Chemical biology is similar to classic genetics except mutations occur at the level of the protein by the action of small molecules which either enhance or repress the activity of the protein. **(B)** Small molecules with biological activity are identified in forward or reverse screens. Forward screens identify small molecules through phenotypic screens. Once the bioactive small molecule is identified further work is necessary to identify the protein target responsible for the effect. Reverse screens identify small molecules that effect a particular process or protein first; followed by determining the resulting phenotype.

1.6 The visual production of *S. coelicolor*'s secondary metabolism warrants a chemical screen

S. coelicolor is particularly well suited for a chemical screen against secondary metabolism due to the production of its two pigments, actinorhodin and prodiginine. These pigments offer an easy, non-invasive screening method to identify small molecules capable of improving their production. Screening for changes in the production of these pigments has been instrumental genetically for understanding their biosynthesis, as well as the regulatory network governing their expression. Small molecules capable of altering pigment production can be used to further understand the complex regulatory network controlling secondary metabolism and offer a chemical strategy for improving, or altering, secondary metabolism in other streptomycetes.

1.7 References

1. Davies J & Davies D (2010) Origins and evolution of antibiotic resistance. *Microbiol Mol Biol Rev* 74: 417-433.
2. Wright GD (2012) Antibiotics: A new hope. *Chem Biol* 19: 3-10.
3. Bhatnagar I & Kim SK (2010) Immense essence of excellence: Marine microbial bioactive compounds. *Mar Drugs* 8: 2673-2701.
4. Mann J (2001) Natural products as immunosuppressive agents. *Nat Prod Rep* 18: 417-430.
5. Zhong JJ & Xiao JH (2009) Secondary metabolites from higher fungi: Discovery, bioactivity, and bioproduction. *Adv Biochem Eng Biotechnol* 113: 79-150.
6. Challis GL & Hopwood DA (2003) Synergy and contingency as driving forces for the evolution of multiple secondary metabolite production by *Streptomyces* species. *Proc Natl Acad Sci U S A* 100 Suppl 2: 14555-14561.
7. Watve MG, Tickoo R, Jog MM & Bhole BD (2001) How many antibiotics are produced by the genus *Streptomyces*?. *Arch Microbiol* 176: 386-390.
8. Weber T, Welzel K, Pelzer S, Vente A & Wohlleben W (2003) Exploiting the genetic potential of polyketide producing streptomycetes. *J Biotechnol* 106: 221-232.
9. Witting K & Sussmuth RD (2011) Discovery of antibacterials and other bioactive compounds from microorganisms-evaluating methodologies for discovery and generation of non-ribosomal peptide antibiotics. *Curr Drug Targets* 12: 1547-1559.
10. Nett M, Ikeda H & Moore BS (2009) Genomic basis for natural product biosynthetic diversity in the actinomycetes. *Nat Prod Rep* 26: 1362-1384.
11. Wright LF & Hopwood DA (1976) Actinorhodin is a chromosomally-determined antibiotic in *Streptomyces coelicolor* A3(2). *J Gen Microbiol* 96: 289-297.
12. Feitelson JS, Malpartida F & Hopwood DA (1985) Genetic and biochemical characterization of the red gene cluster of *Streptomyces coelicolor* A3(2). *J Gen Microbiol* 131: 2431-2441.
13. McCormick JR & Flardh K (2012) Signals and regulators that govern *Streptomyces* development. *FEMS Microbiol Rev* 36: 206-231.

14. Chen L, *et al* (2012) High-throughput screening for *Streptomyces* antibiotic biosynthesis activators. *Appl Environ Microbiol* 78: 4526-4528.

15. Bentley SD, *et al* (2002) Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* 417: 141-147.

16. Hojati Z, *et al* (2002) Structure, biosynthetic origin, and engineered biosynthesis of calcium-dependent antibiotics from *Streptomyces coelicolor*. *Chem Biol* 9: 1175-1187.

17. Song L, *et al* (2006) Type III polyketide synthase beta-ketoacyl-ACP starter unit and ethylmalonyl-CoA extender unit selectivity discovered by *Streptomyces coelicolor* genome mining. *J Am Chem Soc* 128: 14754-14755.

18. Cane DE & Watt RM (2003) Expression and mechanistic analysis of a germacradienol synthase from *Streptomyces coelicolor* implicated in geosmin biosynthesis. *Proc Natl Acad Sci U S A* 100: 1547-1551.

19. Takano E, *et al* (2000) Purification and structural determination of SCB1, a gamma-butyrolactone that elicits antibiotic production in *Streptomyces coelicolor* A3(2). *J Biol Chem* 275: 11010-11016.

20. Takano H, Obitsu S, Beppu T & Ueda K (2005) Light-induced carotenogenesis in *Streptomyces coelicolor* A3(2): Identification of an extracytoplasmic function sigma factor that directs photodependent transcription of the carotenoid biosynthesis gene cluster. *J Bacteriol* 187: 1825-1832.

21. Lautru S, Deeth RJ, Bailey LM & Challis GL (2005) Discovery of a new peptide natural product by *Streptomyces coelicolor* genome mining. *Nat Chem Biol* 1: 265-269.

22. Austin MB, *et al* (2004) Crystal structure of a bacterial type III polyketide synthase and enzymatic control of reactive polyketide intermediates. *J Biol Chem* 279: 45162-45174.

23. Bursy J, *et al* (2008) Synthesis and uptake of the compatible solutes ectoine and 5-hydroxyectoine by *Streptomyces coelicolor* A3(2) in response to salt and heat stresses. *Appl Environ Microbiol* 74: 7286-7296.

24. Barona-Gomez F, Wong U, Giannakopulos AE, Derrick PJ & Challis GL (2004) Identification of a cluster of genes that directs desferrioxamine biosynthesis in *Streptomyces coelicolor* M145. *J Am Chem Soc* 126: 16282-16283.

25. Zhao B, *et al* (2008) Biosynthesis of the sesquiterpene antibiotic albaflavenone in *Streptomyces coelicolor* A3(2). *J Biol Chem* 283: 8183-8189.

26. Komatsu M, Tsuda M, Omura S, Oikawa H & Ikeda H (2008) Identification and functional analysis of genes controlling biosynthesis of 2-methylisoborneol. *Proc Natl Acad Sci U S A* 105: 7422-7427.

27. Wright LF & Hopwood DA (1976) Identification of the antibiotic determined by the SCP1 plasmid of *Streptomyces coelicolor* A3(2). *J Gen Microbiol* 95: 96-106.

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

29. Kodani S, *et al* (2004) The SapB morphogen is a lantibiotic-like peptide derived from the product of the developmental gene ramS in *Streptomyces coelicolor. Proc Natl Acad Sci U S A* 101: 11448-11453.

30. Pawlik K, Kotowska M, Chater KF, Kuczek K & Takano E (2007) A cryptic type I polyketide synthase (*cpk*) gene cluster in *Streptomyces coelicolor* A3(2). *Arch Microbiol* 187: 87-99.

31. Pawlik K, Kotowska M & Kolesinski P (2010) *Streptomyces coelicolor* A3(2) produces a new yellow pigment associated with the polyketide synthase *cpk. J Mol Microbiol Biotechnol* 19: 147-151.

32. Davis NK & Chater KF (1990) Spore colour in *Streptomyces coelicolor* A3(2) involves the developmentally regulated synthesis of a compound biosynthetically related to polyketide antibiotics. *Mol Microbiol* 4: 1679-1691.

33. Okamoto S, Taguchi T, Ochi K & Ichinose K (2009) Biosynthesis of actinorhodin and related antibiotics: Discovery of alternative routes for quinone formation encoded in the act gene cluster. *Chem Biol* 16: 226-236.

34. Beltran-Alvarez P, Cox RJ, Crosby J & Simpson TJ (2007) Dissecting the component reactions catalyzed by the actinorhodin minimal polyketide synthase. *Biochemistry* 46: 14672-14681.

35. Mo S, *et al* (2008) Elucidation of the *Streptomyces coelicolor* pathway to 2-undecylpyrrole, a key intermediate in undecylprodiginine and streptorubin B biosynthesis. *Chem Biol* 15: 137-148.

36. Williamson NR, Fineran PC, Leeper FJ & Salmond GP (2006) The biosynthesis and regulation of bacterial prodiginines. *Nat Rev Microbiol* 4: 887-899.

37. Mo S, *et al* (2008) Elucidation of the streptomyces coelicolor pathway to 2-undecylpyrrole, a key intermediate in undecylprodiginine and streptorubin B biosynthesis. *Chem Biol* 15: 137-148.

38. Williamson NR, Fineran PC, Leeper FJ & Salmond GP (2006) The biosynthesis and regulation of bacterial prodiginines. *Nat Rev Microbiol* 4: 887-899.

39. Martin JF & Liras P (2010) Engineering of regulatory cascades and networks controlling antibiotic biosynthesis in streptomyces. *Curr Opin Microbiol* 13: 263-273.

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

41. McKenzie NL & Nodwell JR (2007) Phosphorylated AbsA2 negatively regulates antibiotic production in streptomyces coelicolor through interactions with pathway-specific regulatory gene promoters. *J Bacteriol* 189: 5284-5292.

42. Chater KF & Chandra G (2008) The use of the rare UUA codon to define "expression space" for genes involved in secondary metabolism, development and environmental adaptation in streptomyces. *J Microbiol* 46: 1-11.

43. Hesketh A, Sun J & Bibb M (2001) Induction of ppGpp synthesis in *Streptomyces coelicolor* A3(2) grown under conditions of nutritional sufficiency elicits *actII*-ORF4 transcription and actinorhodin biosynthesis. *Mol Microbiol* 39: 136-144.

44. D'Alia D, *et al* (2011) Deletion of the signalling molecule synthase ScbA has pleiotropic effects on secondary metabolite biosynthesis, morphological differentiation and primary metabolism in *Streptomyces coelicolor* A3(2). *Microb Biotechnol* 4: 239-251.

45. Fernandez-Moreno MA, Caballero JL, Hopwood DA & Malpartida F (1991) The act cluster contains regulatory and antibiotic export genes, direct targets for translational control by the *bldA* tRNA gene of streptomyces. *Cell* 66: 769-780.

46. Narva KE & Feitelson JS (1990) Nucleotide sequence and transcriptional analysis of the *redD* locus of streptomyces coelicolor A3(2). *J Bacteriol* 172: 326-333.

47. White J & Bibb M (1997) *bldA* dependence of undecylprodigiosin production in *Streptomyces coelicolor* A3(2) involves a pathway-specific regulatory cascade. *J Bacteriol* 179: 627-633.

48. Li YQ, Chen PL, Chen SF, Wu D & Zheng J (2004) A pair of two-component regulatory genes *ecrA1/A2* in *S. coelicolor. J Zhejiang Univ Sci* 5: 173-179.

49. Ryding NJ, Anderson TB & Champness WC (2002) Regulation of the *Streptomyces coelicolor* calcium-dependent antibiotic by *absA*, encoding a cluster-linked two-component system. *J Bacteriol* 184: 794-805.

50. Takano E, *et al* (2005) A bacterial hormone (the SCB1) directly controls the expression of a pathway-specific regulatory gene in the cryptic type I polyketide biosynthetic gene cluster of *Streptomyces coelicolor. Mol Microbiol* 56: 465-479.

51. O'Rourke S, *et al* (2009) Extracellular signalling, translational control, two repressors and an activator all contribute to the regulation of methylenomycin production in *Streptomyces coelicolor. Mol Microbiol* 71: 763-778.

52. Matsumoto A, Hong SK, Ishizuka H, Horinouchi S & Beppu T (1994) Phosphorylation of the AfsR protein involved in secondary metabolism in streptomyces species by a eukaryotic-type protein kinase. *Gene* 146: 47-56.

53. Jin YY, *et al* (2011) S-adenosyl-L-methionine activates actinorhodin biosynthesis by increasing autophosphorylation of the Ser/Thr protein kinase AfsK in *Streptomyces coelicolor* A3(2). *Biosci Biotechnol Biochem* 75: 910-913.

54. Lee Y, Kim K, Suh JW, Rhee S & Lim Y (2007) Binding study of AfsK, a Ser/Thr kinase from *Streptomyces coelicolor* A3(2) and S-adenosyl-L-methionine. *FEMS Microbiol Lett* 266: 236-240.

55. Umeyama T & Horinouchi S (2001) Autophosphorylation of a bacterial serine/threonine kinase, AfsK, is inhibited by KbpA, an AfsK-binding protein. *J Bacteriol* 183: 5506-5512.

56. Horinouchi S & Beppu T (1992) Regulation of secondary metabolism and cell differentiation in *Streptomyces*: A-factor as a microbial hormone and the AfsR protein as a component of a two-component regulatory system. *Gene* 115: 167-172.

57. Horinouchi S (2003) AfsR as an integrator of signals that are sensed by multiple serine/threonine kinases in *Streptomyces coelicolor* A3(2). *J Ind Microbiol Biotechnol* 30: 462-467.

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

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

60. Santos-Beneit F, Rodriguez-Garcia A & Martin JF (2011) Complex transcriptional control of the antibiotic regulator *afsS* in *Streptomyces*: PhoP and AfsR are overlapping, competitive activators. *J Bacteriol* 193: 2242-2251.

61. Horinouchi S (2002) A microbial hormone, A-factor, as a master switch for morphological differentiation and secondary metabolism in *Streptomyces griseus. Front Biosci* 7: d2045-57.

62. Kitani S, *et al* (2011) Avenolide, a *Streptomyces* hormone controlling antibiotic production in *Streptomyces avermitilis. Proc Natl Acad Sci U S A* 108: 16410-16415.

63. Kitani S, Doi M, Shimizu T, Maeda A & Nihira T (2010) Control of secondary metabolism by *farX*, which is involved in the gamma-butyrolactone biosynthesis of *Streptomyces lavendulae* FRI-5. *Arch Microbiol* 192: 211-220.

64. Butler MJ, *et al* (2003) Deletion of *scbA* enhances antibiotic production in *Streptomyces lividans. Appl Microbiol Biotechnol* 61: 512-516.

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

66. Sheeler NL, MacMillan SV & Nodwell JR (2005) Biochemical activities of the *absA* two-component system of *Streptomyces coelicolor. J Bacteriol* 187: 687-696.

67. Brian P, Riggle PJ, Santos RA & Champness WC (1996) Global negative regulation of *Streptomyces coelicolor* antibiotic synthesis mediated by an *absA*-encoded putative signal transduction system. *J Bacteriol* 178: 3221-3231.

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

69. Shu D*, et al* (2009) *afsQ1-Q2-sigQ* is a pleiotropic but conditionally required signal transduction system for both secondary metabolism and morphological development in *Streptomyces coelicolor. Appl Microbiol Biotechnol* 81: 1149-1160.

70. Lu Y*, et al* (2007) Characterization of a novel two-component regulatory system involved in the regulation of both actinorhodin and a type I polyketide in *Streptomyces coelicolor. Appl Microbiol Biotechnol* 77: 625-635.

71. Tunca S, Barreiro C, Coque JJ & Martin JF (2009) Two overlapping antiparallel genes encoding the iron regulator DmdR1 and the adm proteins control siderophore [correction of sedephore] and antibiotic biosynthesis in *Streptomyces coelicolor* A3(2)*. FEBS J* 276: 4814-4827.

72. Yu Z*, et al* (2012) Differential regulation of antibiotic biosynthesis by DraR-K, a novel two-component system in *Streptomyces coelicolor. Mol Microbiol*

73. Uguru GC*, et al* (2005) Transcriptional activation of the pathway-specific regulator of the actinorhodin biosynthetic genes in *Streptomyces coelicolor. Mol Microbiol* 58: 131-150.

74. Rigali S*, et al* (2008) Feast or famine: The global regulator DasR links nutrient stress to antibiotic production by *Streptomyces. EMBO Rep* 9: 670-675.

75. Ou X, Zhang B, Zhang L, Zhao G & Ding X (2009) Characterization of *rrdA*, a TetR family protein gene involved in the regulation of secondary metabolism in *Streptomyces coelicolor. Appl Environ Microbiol* 75: 2158-2165.

76. Hesketh A, Kock H, Mootien S & Bibb M (2009) The role of *absC*, a novel regulatory gene for secondary metabolism, in zinc-dependent antibiotic production in *Streptomyces coelicolor* A3(2)*. Mol Microbiol* 74: 1427-1444.

77. Onaka H, Nakagawa T & Horinouchi S (1998) Involvement of two A-factor receptor homologues in *Streptomyces coelicolor* A3(2) in the regulation of secondary metabolism and morphogenesis*. Mol Microbiol* 28: 743-753.

78. Yang YH*, et al* (2009) NdgR, an IclR-like regulator involved in amino-acid-dependent growth, quorum sensing, and antibiotic production in *Streptomyces coelicolor. Appl Microbiol Biotechnol* 82: 501-511.

79. den Hengst CD*, et al* (2010) Genes essential for morphological development and antibiotic production in *Streptomyces coelicolor* are targets of BldD during vegetative growth*. Mol Microbiol* 78: 361-379.

80. Elliot MA, Locke TR, Galibois CM & Leskiw BK (2003) BldD from *Streptomyces coelicolor* is a non-essential global regulator that binds its own promoter as a dimer. *FEMS Microbiol Lett* 225: 35-40.

81. van Wezel GP, *et al* (2000) *ssgA* is essential for sporulation of *Streptomyces coelicolor* A3(2) and affects hyphal development by stimulating septum formation. *J Bacteriol* 182: 5653-5662.

82. Traag BA, Kelemen GH & Van Wezel GP (2004) Transcription of the sporulation gene *ssgA* is activated by the IclR-type regulator SsgR in a whi-independent manner in *Streptomyces coelicolor* A3(2). *Mol Microbiol* 53: 985-1000.

83. Li W, *et al* (2006) Identification of a gene negatively affecting antibiotic production and morphological differentiation in *Streptomyces coelicolor* A3(2). *J Bacteriol* 188: 8368-8375.

84. Zhang L, Li WC, Zhao CH, Chater KF & Tao MF (2007) NsdB, a TPR-like-domain-containing protein negatively affecting production of antibiotics in *Streptomyces coelicolor* A3 (2). *Wei Sheng Wu Xue Bao* 47: 849-854.

85. Tran NT, Den Hengst CD, Gomez-Escribano JP & Buttner MJ (2011) Identification and characterization of CdgB, a diguanylate cyclase involved in developmental processes in *Streptomyces coelicolor. J Bacteriol* 193: 3100-3108.

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

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

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

89. Shima J, Hesketh A, Okamoto S, Kawamoto S & Ochi K (1996) Induction of actinorhodin production by *rpsL* (encoding ribosomal protein S12) mutations that confer streptomycin resistance in *Streptomyces lividans* and *Streptomyces coelicolor* A3(2). *J Bacteriol* 178: 7276-7284.

90. Chakraburtty R & Bibb M (1997) The ppGpp synthetase gene (*relA*) of *Streptomyces coelicolor* A3(2) plays a conditional role in antibiotic production and morphological differentiation. *J Bacteriol* 179: 5854-5861.

91. Sun J, Hesketh A & Bibb M (2001) Functional analysis of *relA* and *rshA*, two *relA/spoT* homologues of *Streptomyces coelicolor* A3(2). *J Bacteriol* 183: 3488-3498.

92. Ochi K & Hosoya Y (1998) Genetic mapping and characterization of novel mutations which suppress the effect of a *relC* mutation on antibiotic production in *Streptomyces coelicolor* A3(2). *J Antibiot (Tokyo)* 51: 592-595.

93. Saito N, *et al* (2006) EshA accentuates ppGpp accumulation and is conditionally required for antibiotic production in *Streptomyces coelicolor* A3(2). *J Bacteriol* 188: 4952-4961.

94. Wang G, Tanaka Y & Ochi K (2010) The G243D mutation (*afsB* mutation) in the principal sigma factor sigmaHrdB alters intracellular ppGpp level and antibiotic production in *Streptomyces coelicolor* A3(2). *Microbiology* 156: 2384-2392.

95. Yu G, *et al* (2011) Strain improvement of *Streptomyces roseosporus* for daptomycin production by rational screening of he-ne laser and NTG induced mutants and kinetic modeling. *Appl Biochem Biotechnol* 163: 729-743.

96. Jakeman DL, Graham CL, Young W & Vining LC (2006) Culture conditions improving the production of jadomycin B. *J Ind Microbiol Biotechnol* 33: 767-772.

97. Doull JL, Ayer SW, Singh AK & Thibault P (1993) Production of a novel polyketide antibiotic, jadomycin B, by *Streptomyces venezuelae* following heat shock. *J Antibiot (Tokyo)* 46: 869-871.

98. Barona-Gomez F, *et al* (2006) Multiple biosynthetic and uptake systems mediate siderophore-dependent iron acquisition in *Streptomyces coelicolor* A3(2) and *Streptomyces ambofaciens* ATCC 23877. *Microbiology* 152: 3355-3366.

99. Medema MH, Alam MT, Breitling R & Takano E (2011) The future of industrial antibiotic production: From random mutagenesis to synthetic biology. *Bioeng Bugs* 2: 230-233.

100. Cheng YR, Huang J, Qiang H, Lin WL & Demain AL (2001) Mutagenesis of the rapamycin producer *Streptomyces hygroscopicus* FC904. *J Antibiot (Tokyo)* 54: 967-972.

101. Strauch E, Takano E, Baylis HA & Bibb MJ (1991) The stringent response in *Streptomyces coelicolor* A3(2). *Mol Microbiol* 5: 289-298.

102. Xu J, Tozawa Y, Lai C, Hayashi H & Ochi K (2002) A rifampicin resistance mutation in the *rpoB* gene confers ppGpp-independent antibiotic production in *Streptomyces coelicolor* A3(2). *Mol Genet Genomics* 268: 179-189.

103. Wang G, Hosaka T & Ochi K (2008) Dramatic activation of antibiotic production in *Streptomyces coelicolor* by cumulative drug resistance mutations. *Appl Environ Microbiol* 74: 2834-2840.

104. Hosaka T, *et al* (2009) Antibacterial discovery in actinomycetes strains with mutations in RNA polymerase or ribosomal protein S12. *Nat Biotechnol* 27: 462-464.

105. Foley TL, Young BS & Burkart MD (2009) Phosphopantetheinyl transferase inhibition and secondary metabolism. *FEBS J* 276: 7134-7145.

106. Olano C, Lombo F, Mendez C & Salas JA (2008) Improving production of bioactive secondary metabolites in actinomycetes by metabolic engineering. *Metab Eng* 10: 281-292.

107. Mo S, Ban YH, Park JW, Yoo YJ & Yoon YJ (2009) Enhanced FK506 production in *Streptomyces clavuligerus* CKD1119 by engineering the supply of methylmalonyl-CoA precursor. *J Ind Microbiol Biotechnol* 36: 1473-1482.

108. Ryu YG, Butler MJ, Chater KF & Lee KJ (2006) Engineering of primary carbohydrate metabolism for increased production of actinorhodin in *Streptomyces coelicolor. Appl Environ Microbiol* 72: 7132-7139.

109. Liu Z, Guo M, Qian J, Zhuang Y & Zhang S (2008) Disruption of *zwf2* gene to improve oxytetracycline biosynthesis in *Streptomyces rimosus* M4018. *Wei Sheng Wu Xue Bao* 48: 21-25.

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

111. McKenzie NL, *et al* (2010) Induction of antimicrobial activities in heterologous streptomycetes using alleles of the *Streptomyces coelicolor* gene *absA1. J Antibiot (Tokyo)* 63: 177-182.

112. Malla S, Niraula NP, Liou K & Sohng JK (2010) Self-resistance mechanism in *Streptomyces peucetius*: Overexpression of *drrA*, *drrB* and *drrC* for doxorubicin enhancement. *Microbiol Res* 165: 259-267.

113. Qiu J, *et al* (2011) Overexpression of the ABC transporter AvtAB increases avermectin production in *Streptomyces avermitilis. Appl Microbiol Biotechnol* 92: 337-345.

114. McKenzie NL & Nodwell JR (2007) Phosphorylated AbsA2 negatively regulates antibiotic production in *Streptomyces coelicolor* through interactions with pathway-specific regulatory gene promoters. *J Bacteriol* 189: 5284-5292.

115. Guo J, *et al* (2010) The pathway-specific regulator AveR from *Streptomyces avermitilis* positively regulates avermectin production while it negatively affects oligomycin biosynthesis. *Mol Genet Genomics* 283: 123-133.

116. Retzlaff L & Distler J (1995) The regulator of streptomycin gene expression, StrR, of *Streptomyces griseus* is a DNA binding activator protein with multiple recognition sites. *Mol Microbiol* 18: 151-162.

117. Laureti L, *et al* (2011) Identification of a bioactive 51-membered macrolide complex by activation of a silent polyketide synthase in *Streptomyces ambofaciens*. *Proc Natl Acad Sci U S A* 108: 6258-6263.

118. Menendez N, Brana AF, Salas JA & Mendez C (2007) Involvement of a chromomycin ABC transporter system in secretion of a deacetylated precursor during chromomycin biosynthesis. *Microbiology* 153: 3061-3070.

119. Bunet R, *et al* (2011) Characterization and manipulation of the pathway-specific late regulator AlpW reveals *Streptomyces ambofaciens* as a new producer of kinamycins. *J Bacteriol* 193: 1142-1153.

120. Gomez-Escribano JP & Bibb MJ (2011) Engineering *Streptomyces coelicolor* for heterologous expression of secondary metabolite gene clusters. *Microb Biotechnol* 4: 207-215.

121. Eustaquio AS, *et al* (2005) Heterologous expression of novobiocin and clorobiocin biosynthetic gene clusters. *Appl Environ Microbiol* 71: 2452-2459.

122. Cane DE, Luo G, Khosla C, Kao CM & Katz L (1995) Erythromycin biosynthesis. highly efficient incorporation of polyketide chain elongation intermediates into 6-deoxyerythronolide B in an engineered *Streptomyces* host. *J Antibiot (Tokyo)* 48: 647-651.

123. Komatsu M, Uchiyama T, Omura S, Cane DE & Ikeda H (2010) Genome-minimized *Streptomyces* host for the heterologous expression of secondary metabolism. *Proc Natl Acad Sci U S A* 107: 2646-2651.

124. Jung WS, *et al* (2006) Heterologous expression of tylosin polyketide synthase and production of a hybrid bioactive macrolide in *Streptomyces venezuelae*. *Appl Microbiol Biotechnol* 72: 763-769.

125. Park SR, Ahn MS, Han AR, Park JW & Yoon YJ (2011) Enhanced flavonoid production in *Streptomyces venezuelae* via metabolic engineering. *J Microbiol Biotechnol* 21: 1143-1146.

126. Chater KF & Wilde LC (1980) *Streptomyces albus* G mutants defective in the SalGI restriction-modification system. *J Gen Microbiol* 116: 323-334.

127. Hopwood DA, Kieser T, Wright HM & Bibb MJ (1983) Plasmids, recombination and chromosome mapping in *Streptomyces lividans* 66. *J Gen Microbiol* 129: 2257-2269.

128. Cane DE & Ikeda H (2012) Exploration and mining of the bacterial terpenome. *Acc Chem Res* 45: 463-472.

129. Challis GL (2008) Mining microbial genomes for new natural products and biosynthetic pathways. *Microbiology* 154: 1555-1569.

130. Koglin A & Walsh CT (2009) Structural insights into nonribosomal peptide enzymatic assembly lines. *Nat Prod Rep* 26: 987-1000.

131. Rottig M, *et al* (2011) NRPSpredictor2--a web server for predicting NRPS adenylation domain specificity. *Nucleic Acids Res* 39: W362-7.

132. Chen AY, Cane DE & Khosla C (2007) Structure-based dissociation of a type I polyketide synthase module. *Chem Biol* 14: 784-792.

133. Yadav G, Gokhale RS & Mohanty D (2009) Towards prediction of metabolic products of polyketide synthases: An in silico analysis. *PLoS Comput Biol* 5: e1000351.

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

135. Kersten RD, *et al* (2011) A mass spectrometry-guided genome mining approach for natural product peptidogenomics. *Nat Chem Biol* 7: 794-802.

136. Liu WT, Kersten RD, Yang YL, Moore BS & Dorrestein PC (2011) Imaging mass spectrometry and genome mining via short sequence tagging identified the anti-infective agent arylomycin in *Streptomyces roseosporus*. *J Am Chem Soc* 133: 18010-18013.

137. Spring DR (2005) Chemical genetics to chemical genomics: Small molecules offer big insights. *Chem Soc Rev* 34: 472-482.

138. O'Connor CJ, Laraia L & Spring DR (2011) Chemical genetics. *Chem Soc Rev* 40: 4332-4345.

139. Barker CA, Farha MA & Brown ED (2010) Chemical genomic approaches to study model microbes. *Chem Biol* 17: 624-632.

140. Spring DR (2005) Chemical genetics to chemical genomics: Small molecules offer big insights. *Chem Soc Rev* 34: 472-482.

141. Alvarez MA, Gomez A, Gomez P, Brooks JE & Rodicio MR (1996) Comparative analysis of expression of the SalI restriction-modification system in *Escherichia coli* and *Streptomyces*. *Mol Gen Genet* 253: 74-80.

142. Alvarez MA, Gomez A, Gomez P & Rodicio MR (1995) Expression of the SalI restriction-modification system of *Streptomyces albus* G in escherichia coli. *Gene* 157: 231-232.

143. Godany A, Farkasovska J, Bukovska G & Timko J (2001) Connection between foreign DNA replication and induced expression of the restriction-modification system in *Streptomyces aureofaciens. Folia Microbiol (Praha)* 46: 193-196.

144. Lyutzkanova D, Stoilova-Disheva M & Peltekova V (2001) The restriction-modification system in *Streptomyces flavopersicus. Folia Microbiol (Praha)* 46: 119-122.

145. Matseliukh AB (2001) Genetic transformation of *Streptomyces globisporus* strain 1912: Restriction barrier and plasmid compatibility. *Mikrobiol Z* 63: 15-22.

146. Zotchev SB, Schrempf H & Hutchinson CR (1995) Identification of a methyl-specific restriction system mediated by a conjugative element from *Streptomyces bambergiensis. J Bacteriol* 177: 4809-4812.

CHAPTER 2

A SYNTHETIC *LUXCDABE* GENE CLUSTER OPTIMIZED FOR EXPRESSION IN HIGH-GC BACTERIA

Portions of this chapter were published in:

Arryn Craney, Tobias Hohenauer, Ye Xu, Naveen Kumar, Yingfu Li and Justin Nodwell. A synthetic *luxCDABE* gene cluster optimized for expression in high-GC bacteria. Nucleic Acids Res. 2007; 35(6):e46.

Author Contributions:

The design of the high-GC *luxCDABE* cluster and the initial amplification of the individual *lux* genes were carried out by Tobias Hohenaur. Xe Yu aided in correcting the errors in the individual *luxCDABE* genes. Salman Ahmed created the pMUs* plasmid under my supervision. I carried out correcting, assembly and testing of the initial *luxCDABE* plasmid as well as all other corrections (pMU1* and pF*lux*).

2.1 Abstract

The *luxCDABE* operon of the bioluminescent bacterium *Photorhabdus luminescens* has proven to be a superb transcriptional reporter. It encodes a luciferase (LuxA and LuxB) and the enzymes that produce its substrate (LuxC, LuxD and LuxE) so cells that express the cluster emit 490 nm light spontaneously. The sequence of these genes is AT-rich (>69%) and for this and other reasons, they are not expressed efficiently in high-GC bacteria like *Streptomyces coelicolor.* We therefore constructed a synthetic *luxCDABE* operon encoding the *P. luminescens* Lux proteins optimized for expression in high-GC bacteria. We tested the genes using transcriptional fusions to *S. coelicolor* promoters having well-established expression profiles during this organism's lifecycle. *hrdB* encodes a housekeeping sigma factor; while *ramC* is important for the formation of the spore-forming cells called aerial hyphae and *whiE* is required for the production of a grey, spore-associated pigment that is deposited in the walls of developing spores. Using these fusions we demonstrate that our synthetic *lux* genes are functional in *S. coelicolor* and that they accurately report complex developmental gene expression patterns. We suggest that this *lux* operon will be widely useful for research in high-GC bacteria and envision its application to understanding chemical manipulation in *S. coelicolor*.

2.2 Introduction

Reporter genes that permit facile detection of promoter activity are of central importance to molecular genetic research. One especially useful transcriptional reporter is the *luxCDABE* operon which encodes both luciferase (a heterodimer of LuxA and LuxB) and enzymes required for the production of its substrate tetradecanal (LuxC, LuxD and LuxE). The biochemical requirements for the production of light by these gene products are available in the cytoplasm of all aerobically growing organisms so cells that express the *luxCDABE* operon emit light spontaneously (1).

Codon usage can limit the applicability of naturally occurring reporter genes to some species. For example, genes of the bacterial genus *Streptomyces* have a GC content that is typically >70% and there is evidence that translation in these organisms favours G and C residues in the wobble position of most codons (2). Furthermore, in *Streptomyces coelicolor* there is only one tRNA for the leucine-encoding codon TTA and it is encoded by the gene *bldA* (3). *bldA* is dispensable for viability as it is only required for the expression of a subset of non-essential genes: several of these are required for antibiotic synthesis and the morphogenetic events that culminate with sporulation in this organism. Of greatest relevance to the work we report here, *bldA* is developmentally regulated such that its product is present at very low abundance during the first 24 hours of colony growth. For this reason, genes bearing TTA codons are poorly expressed

and developmentally regulated. Evidence suggests that this may also be the case in *Streptomyces clavuligerus, Streptomyces halstedii, Streptomyces exfoliatus,* and *Streptomyces griseus* (4-7). The *luxCDABE* operon of *P. luminescens* is relatively rich in codons having either A or T in the wobble position (69%) suggesting that these genes would not be efficiently expressed in *Streptomyces* or many other high-GC organisms. Furthermore, there are 63 TTA codons in this operon: even in the event that these genes were translated efficiently they would be subject to developmental regulation via *bldA* that would largely eliminate their experimental utility.

A number of reporter systems have been developed for use in the streptomycetes; however, to date none has been adopted for routine analysis of gene expression. The most successful of these is probably the *egfp* gene, which encodes the enhanced green fluorescent protein (8). While the green fluorescent protein has revolutionized cell biology in this organism (9-13), it is not a particularly good transcriptional reporter. Its drawbacks are that it tends to photo-bleach rapidly, often compromising the weak signals from developmentally regulated promoters and furthermore, *S. coelicolor* exhibits significant auto-fluorescence when illuminated with blue light (8). Other reporter genes that have been developed for use in *Streptomyces* include *amy*, *xylE* and *melC*, each with its own advantages and disadvantages (14-16). More recently, the *luxAB* genes from the bacterium *Vibrio harveyii* have been employed in *Streptomyces* (17).

One drawback to these genes is that in order to visualize bioluminescence it is necessary to add the *n*-decanal substrate to cells and it is not clear that this substrate can pass through the cell wall of all *S. coelicolor* cell types with equivalent efficiency. Furthermore, if *S. coelicolor* is to be adapted for high throughput research, a simple reporter that does not require an added substrate would be advantageous. At present, the vast majority of molecular analysis of gene expression involves direct assessment of transcript levels using S1 nuclease analysis, primer extension, northern blotting or RT-PCR. Each of these has important virtues but all are excessively labour-intensive for high throughput experiments.

In this work, we have constructed an entirely synthetic 5668 bp *luxCDABE* operon that lacks TTA codons and in which the majority of codons end in a G or C. We show that the genes are functional in *S. coelicolor* and that they accurately report the expression profiles of several genes. We suggest that this new synthetic gene cluster will prove applicable to many gene expression experiments in the streptomycetes, as well as other high-GC bacteria.

2.3 Results

2.3.1 Assembly of a high-GC *luxCDABE* operon

We initiated this work by designing *luxC, luxD, luxA, luxB* and *luxE* sequences *in silico* that were optimized for expression in *S. coelicolor* and flanked by restriction endonuclease recognition sites for assembly into an operon (**Figure 2.1C**). The Shine-Dalgarno site upstream of the *S. coelicolor tuf1* gene was included 8 bp upstream of the methionine-encoding start codon of each gene.

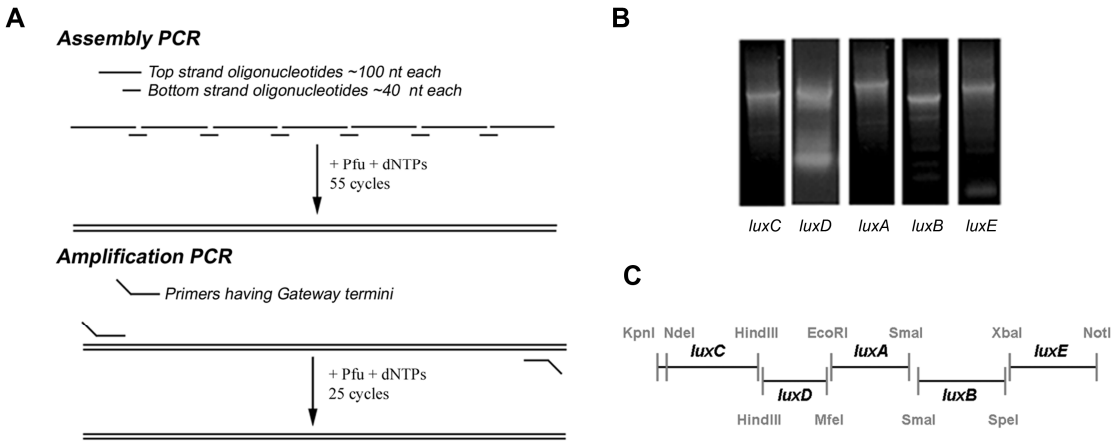

Figure 2.1. Assembly of the high-GC *luxCDABE* operon (A) Schematic diagram of assembly/PCR strategy. In the assembly PCR, the top strand and bottom strand oligonucleotides were mixed in equimolar amounts. A PCR reaction of 55 cycles annealed the primers to create the full length product. In the amplification PCR, the PCR product from the assembly PCR was incubated with primers which introduced sites for entry into the Gateway cloning vector sites for subsequent cloning steps. **(B) Individual amplified *lux* genes assembled using the PCR strategy.** PCR products resulting from the two rounds of PCR are shown. Expected sizes were *luxC* 1.5 kb, *luxD* 1.0 kb, *luxA* 1.1 kb, *luxB* 1.1 kb, and *luxE* 1.1 kb. (**C) Strategy for *luxCDABE* operon assembly.** The high-GC *lux* genes were assembled into an operon through restriction sites corresponding to the pBluescript II SK+ multiple cloning site.

Oligonucleotides of 90-110 nt corresponding to the top strand of each sequence and 40 nt bridging oligonucleotides complementary to the 20 nt at either end of adjacent top strand oligonucleotides were used to amplify the individual *lux* genes (**Figure 2.1A**). Genes were then assembled in a two-step process. First, equimolar mixtures of top strand and bridging DNAs for each *lux* gene were subjected to a 55 cycle assembly reaction with Pfu DNA polymerase (**Figure 2.1A – Assembly PCR**). Second, the products of assembly reactions were gel purified, incubated with PCR primers complementary to the 5' and 3' ends of the genes, each including the sequences necessary for Gateway-mediated recombinational cloning and for subsequent operon assembly, and subjected to a 25 cycle amplification reaction (**Figure 2.1A – Amplification PCR**). The products of amplification reactions for *luxC, luxD, luxA, luxB* and *luxE* are shown in **Figure 2.1B**. Fragments of the appropriate size were then gel purified and introduced into the Gateway entry vector pDONR221 using bacteriophage λ integrase.

Table 2.1. Isolation of individual *lux* genes after assembly by PCR

Gene	Clones Isolated	Clones with Correct Size	Best Candidate	Mutations*
luxA	40	8	A22	3 bp deletion at 109 5 bp deletion at 882
luxB	35	14	B23	point mutation at 907 point mutation at 912
luxC	10	7	C4	1 bp insertion at 551 3 bp deletion at 575
luxD	15	6	D2	None
luxE	25	6	E5	point mutation at 726 1 bp deletions at 970, 988 and 1026

* altered protein sequence

We determined the sequence of a number of alleles for each cloned *lux* gene to establish whether the assembly and amplification reactions had yielded the desired sequences. We identified one perfect allele of *luxD* as well as alleles of the other genes that contained a small enough number of sequence alterations for correction by site-directed mutagenesis (**Table 2.1**). Once we obtained perfect clones for each gene, we assembled a synthetic *luxCDABE* operon (genbank accession EF173694) in the vector pBluescript II SK+ (**Figure 2.1C**). The assembled operon was excised as a 5668 bp NdeI-NotI fragment and ligated to the backbone of the vector pRT801 (30). The transcriptional terminators flanking the multiple cloning site of pIJ8660 (8) were transferred as restriction fragments to flank *luxCDABE*, creating the vector pMU1 (**Figure 2.2A**). This vector can be conjugated directly from *E. coli* to *S. coelicolor* and integrates site-specifically into the *S. coelicolor* chromosome to ensure single copy gene expression (30).

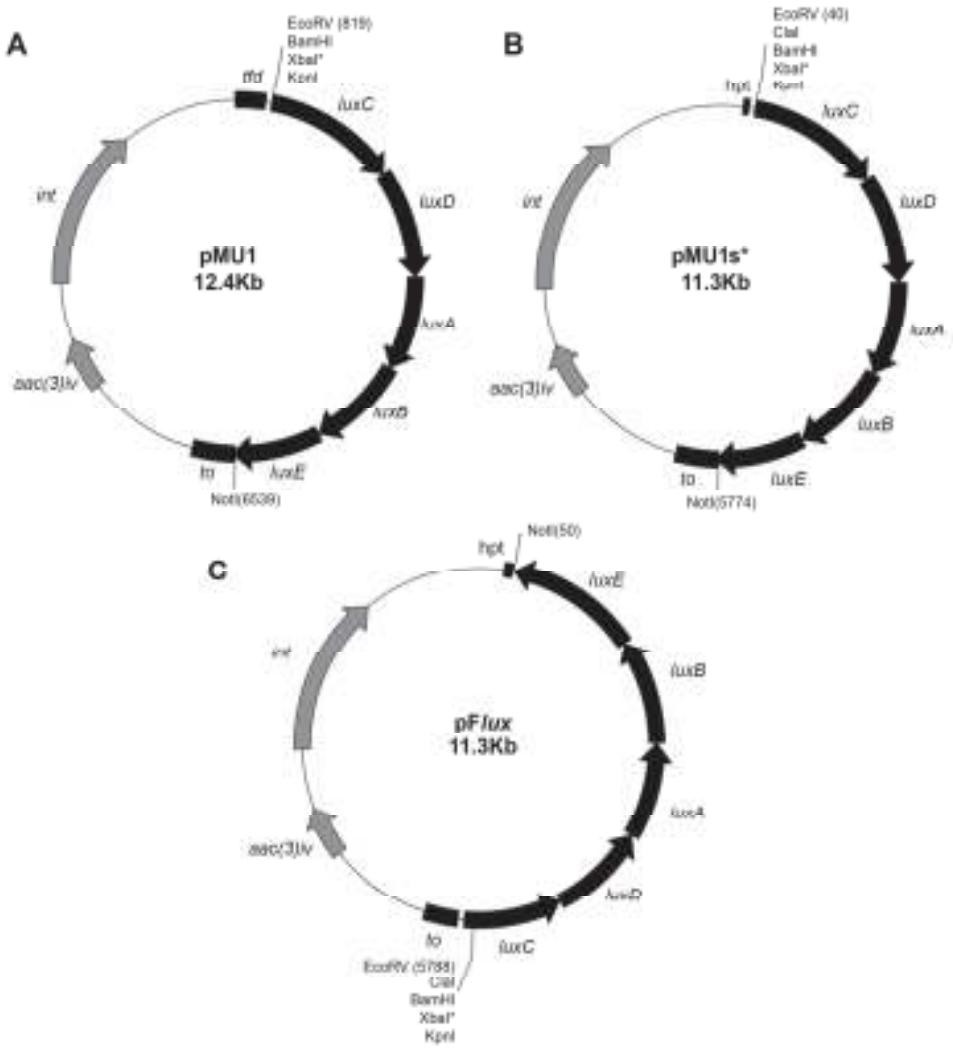

Figure 2.2. The evolution of the pMU1 plasmid. *tfd*, the major transcriptional terminator of phage *fd*; *to*, transcriptional terminator of phage λ; hpt, hairpin terminator from 16S rRNA; *aac(3)IV* apramycin-resistance cassette selectable in *E. coli* and *Streptomyces; int,* the integrase gene and attachment site of the ΦBT1 phage. (non-unique sites are marked with *). **(A) The pMU1 and pMU1* plasmids.** pMU1 is the original published plasmid. pMU1* replaces *luxD* and improves luminescence. pMU1* is not shown as the vector map was unchanged by *luxD* alterations. **(B) The pMU1s* plasmid.** Improved stability of the plasmid during cloning and conjugal transfer to *Streptomyces* by replacing the upstream terminator with a hairpin terminator (this plasmid has significant background luminescence). **(C) pF*lux* plasmid.** The orientation of the *luxCDABE* operon was reversed to remove background luminescence.

2.3.2 Bioluminescence directed by the *ramC, whiEp1* and *hrdB* promoters

To test the expression of our synthetic *lux* operon in *S. coelicolor,* we introduced DNA fragments upstream of *luxCDABE* containing the promoters for the gene *hrdB,* which encodes an *S. coelicolor* housekeeping sigma factor and those of two genes involved in morphogenesis: *ramC* and *whiE*. *ramC,* which is the first gene in the *ramCSAB* operon, encodes an enzyme required for production of the morphogenetic peptide SapB (22, 31, 32). The expression of this operon depends on a number of other developmental genes including *bldB* and *ramR* (13, 33). The *whiE* gene cluster expresses a polyketide synthase that produces a grey pigment that is deposited on the spore surface late in development and depends on the activity of earlier developmental genes. *whiEp1* is the stronger of two divergent promoters within the *whiE* gene cluster (34). We then introduced the resulting plasmids p^{hrdB}lux, p^{ramC}lux and p^{whiEp1}lux into the morphogenetically wild type strains M145 and M600 as well as developmental mutants *bldB* (35), *ramR* and *whiG* (36).

Strains containing pMU1 and the promoter-*lux* fusions were then tested for luminescence on solid media during a three day time course. Luminescence was measured at 8 hour time points using an Envision multilabel plate reader. Luminescence patterns in the wild type strains M145 and M600 were very similar and as a result only the data for M145 are shown.

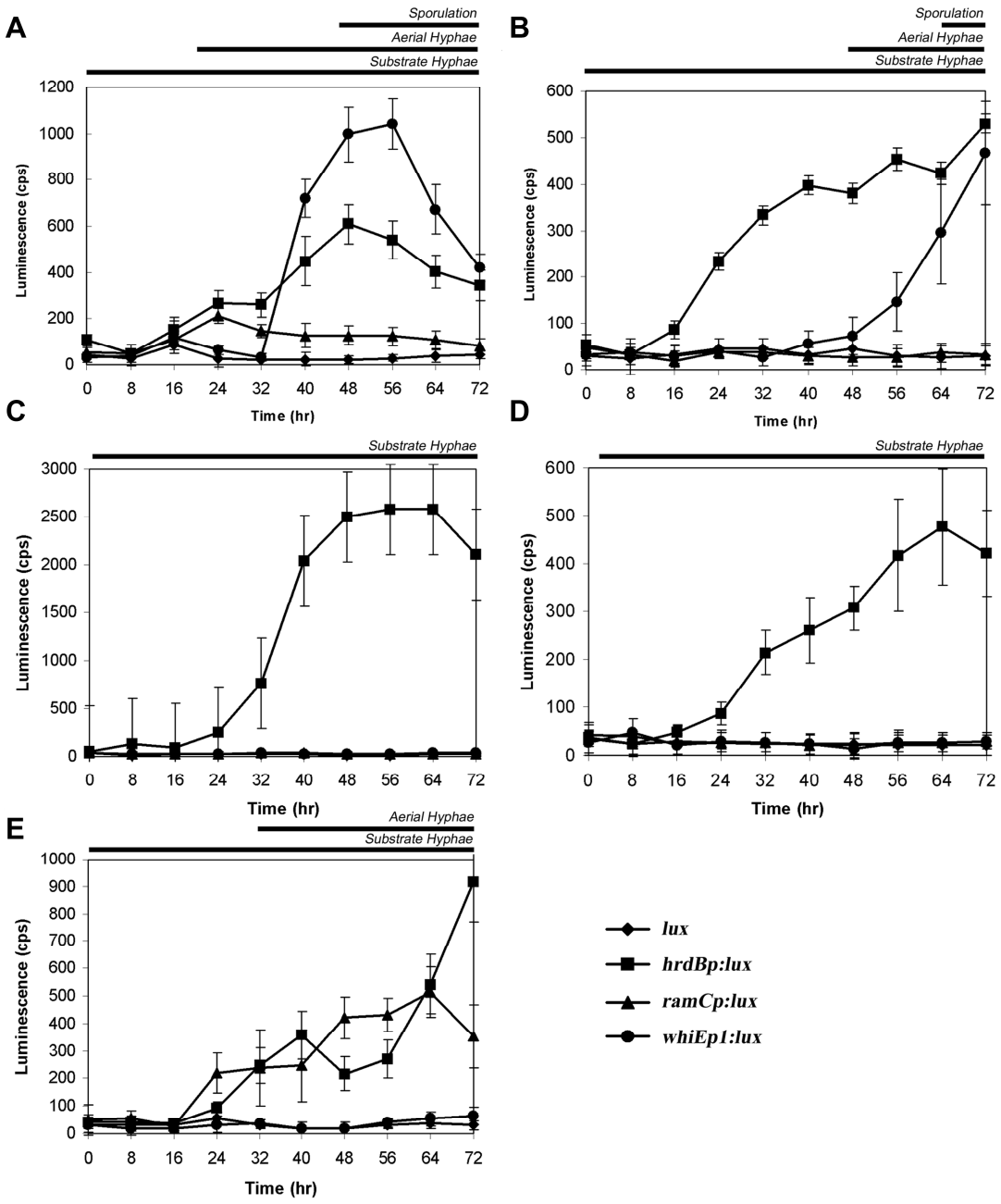

Figure 2.3. Activity of the *lux* operon in *S. coelicolor* when grown on solid media. p^{hrdB}lux, p^{ramC}lux, p^{whiEp1}lux and pMU1 were introduced into wild type M145 and developmental mutants *bldB, ramR* and *whiG.* M145 containing the *lux* promoter fusions grown on R2YE **(A)** and MS agar **(B)**. Developmental mutant strains, *bldB* **(C)** *ramR* **(D)** and *whiG* **(E)** containing the *lux* promoter fusions grown on R2YE.

We observed no background luminescence in strains containing the promoterless *lux* operon in any strains. In contrast, the *hrdB* promoter fusion was clearly active in wild type and the three developmental mutants (**Figure 2.3A-E**). Little to no expression was detected during the first 8 hours of growth. However by 16 hours, *hrdB* promoter-dependent luminescence could be detected that was reproducible and well above that of the promoterless control. Luminescence levels directed by this promoter were similar in M145, the *ramR* and *whiG* mutants, reaching values of ~500 cps. Interestingly, *hrdB* promoter-directed luminescence was reproducibly ~5x higher in the *bldB* mutant than in the other strains, routinely reaching 2500 cps (**Figure 2.3C**).

In wild type cells, detectable luminescence from *whiEp1* was achieved after 40 hours of growth and reached a peak of ~1000 cps at 56 hours (**Figure 2.3A**). In agreement with the established requirements for activation of the *whiE* cluster, no significant luminescence was detected from the *whiEp1* promoter in any of the developmental mutants tested as they are blocked in either aerial hyphae formation (*bldB* and *ramR*) or maturation of the aerial hyphae prior to division (*whiG*) (**Figure 2.3C, D** and **E**).

The formation of aerial hyphae on rich media depends on the *ramCSAB* gene cluster as it is the source of a surfactant molecule, SapB that facilitates upward growth of the spore-forming cells. We measured bioluminescence directed by p^{ramC}lux on the rich medium R2YE and observed significant luminescence at 16

hours, which then peaked at 24 hours and decreased at subsequent time points (**Figure 2.3A**). We detected no *ramC* promoter-dependent luminescence in either the *bldB* and *ramR* mutant strains, consistent with previous results (13, 33) (**Figure 2.3C** and **D**). In the *whiG* null strain, no observable luminescence was detected for the first 16 hours of growth from the *ramC* promoter, after which expression increased throughout the rest of development reaching a peak value of ~500 cps at 64 hours (**Figure 2.3E**). This is consistent with western blot analysis demonstrating that RamC production is up-regulated in many of the *whi* mutants, including *whiG*, such that RamC accumulates at later time points instead of peaking at aerial hyphae formation (Tamara O'Connor, Dachuan Zhang and Justin Nodwell, unpublished results).

The *ram* gene cluster is dispensable for aerial mycelium formation when *S. coelicolor* is grown on some carbon sources such as mannitol and indeed, SapB is not produced under these conditions (37). In agreement with this, when we cultivated wild type M145 cells containing the *lux* fusions on MS agar (a mannitol containing media) we detected luminescence directed by both the *hrdB* and *whiEp1* promoters but none from the *ramC* promoter (**Figure 2.3B**). Luminescence from the *hrdB* promoter reached similar levels of expression as when grown on rich media with values reaching ~500 cps. Expression from *whiEp1* was delayed in comparison to growth on rich media, as was the production of aerial hyphae and grey pigmented spores.

2.3.3 Advancements in the *luxCDABE* transcriptional reporter

The luminescence observed in our synthetic *luxCDABE* gene cluster was much lower than expected. Maximum luminescence was found to be around 2500 cps in our synthetically derived gene cluster; while values from the wild type *lux* cluster can be in excess of 200,000. Originally we believed this lower luminescence to be a side effect of using the gene cluster in *Streptomyces*; however, it was discovered that the *luxD* gene was designed to match a misannotated version of the LuxD (Accession #AAK98553) and actually contained 4 regions with amino acid sequence changes – N32T, T39P, Y236H and a 3 amino acid change starting at 246 GEN to CVF. We suspected the sequences changes in *luxD* most likely caused the lower then expected luminescence. To create the idealized *luxCDABE* gene cluster, a high-GC *luxD* gene was synthesized with the corrected amino acid sequence (purchased from Mr. Gene) and replaced within the gene cluster, creating the second version of our *lux* plasmid, pMU1*. The sequence changes in *luxD* did not result in changes in the overall *lux* plasmid restriction sites (**Figure 2.2A**). To establish the effect of replacing *luxD* on luminescence in *Streptomyces*, *whiEp1* was used for comparison of luminescence for the two plasmids (**Figure 2.4**). The overall pattern of luminescence achieved in pMU1 and pMU1* are identical with peak luminescence around 110 hours. Maximum values differ significantly with the original pMU1 values reaching only 150 cps (**Figure 2.4A**) and the new pMU1* plasmid reaching upwards of 8000 cps (**Figure 2.4B**), approximately 50X higher

luminescence. This increased luminescence also improves the sensitivity of the reporter, reaching detectable luminescence in pMU1* at 70 hours, 15 hours earlier than the original plasmid.

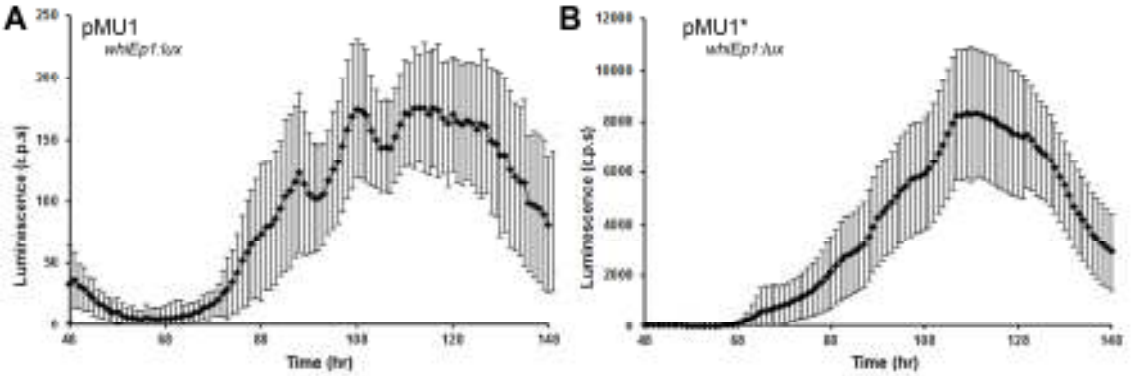

Figure 2.4. Comparison of the luminescence from pMU1 and pMU1* using *whiE*p1. Luminescence readings were taken every hour from 48 hours to 148 hours with growth occurring at room temperature. **(A)** Luminescence from *whiE*p1 in pMU1* achieved maximum detectable luminescence of 150 c.p.s. **(B)** Luminescence from *whiE*p1 in pMU1* reached maximum luminescence of 8000 c.p.s.

One final improvement to the plasmid was to overcome a technical issue. For unknown reasons, cloning promoter fragments upstream of the *lux* gene cluster was exceedingly difficult. We speculated this could be the result of 2 repeated 500 bp regions in the transcriptional terminators that flanks *luxCDABE* in combination with the large size of the plasmid (~12 kb). To improve the ease of use, we have modified the plasmid to remove the repeated sequence by replacing the upstream terminator with a 40 bp hairpin loop which also shortened

the overall plasmid size to ~11kb. In addition, an extra restriction site was added in the multiple cloning site (ClaI), creating the plasmid pMUs* (**Figure 2.2B**). These changes have reduced the instability of the plasmid dramatically and improved the ease of cloning. Unfortunately, an unexpected side-effect of replacing the upstream terminator was strong background luminescence after 24 hours of growth, reaching very high levels of expression of 4000 cps at 68 hours (**Figure 2.5**). This background is most likely due to read-through from the integrase gene downstream from the *lux* operon. To overcome this, the *lux* operon was cloned in the reverse orientation, creating the pFlux plasmid (**Figure 2.2C**). The multiple cloning site orientation and spacing was maintained in relation to the *luxCDABE* operon, as in the previous versions. The change in orientation of the *lux* cluster removed the background demonstrated by the pMU1s* plasmid (**Figure 2.5**), correcting the final issue with this reporter plasmid (summarized in **Table 2.2**).

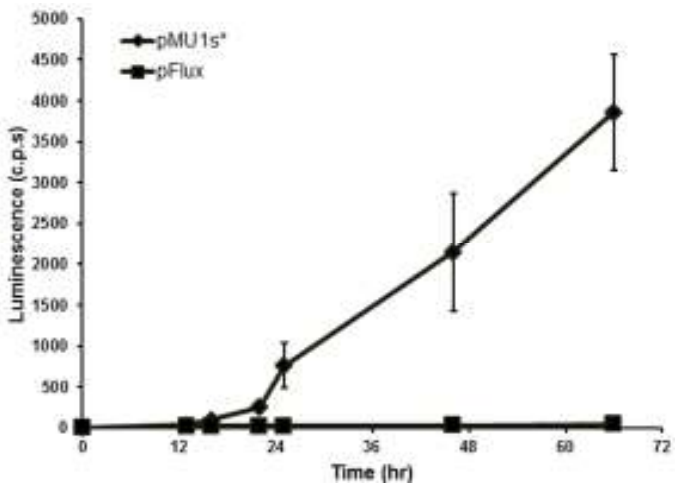

Figure 2.5. Reduced background from the pF*lux* plasmid. *S. coelicolor* strains containing promoterless vectors pMUs* and pF*lux* were inoculated on R5M and grown at 30°C with luminescence readings taken until 68 hours of growth.

Table 2.2. Evolution of the *lux* plasmid for use in *Streptomyces*

Plasmid Name	Change	Result
pMU1	optimized *luxCDABE*	luminescent streptomycetes
pMU1*	replacement of *luxD*	improved luminescence
pMU1s*	change of upstream terminator ClaI added to MCS	improved plasmid stability increased background
pFlux	orientation of *luxCDABE* reversed	reduced background

2.4 Discussion

We have constructed a *luxCDABE* gene cluster optimized for expression in high-GC bacteria and established that the cluster is active in *S. coelicolor.* Fusions of this cluster to promoters for the genes *hrdB, ramC* and *whiE* faithfully report their activities via spontaneous bioluminescence. The activity of each promoter behaved precisely as expected: the *hrdB* promoter fusion was active in wild type cells and developmental mutants, the *ramC* promoter required the genes *bldB* and *ramR* but not *whiG,* and *whiEp1* was dependent on *bldB, ramR* and *whiG.* These fusions also reiterated previously demonstrated medium dependence of the developmental genes: the *ramC* promoter was active on rich medium but not on MS agar while *whiEp1* and *hrdB* were active on both media. We predict therefore that this gene cluster will be widely applicable for routine gene expression studies in *S. coelicolor* and other high-GC micro-organisms. These data suggest that this reporter will greatly facilitate future gene expression studies in *Streptomyces*. Variation in fatty acid pools needed for biosynthesis of the *lux* substrate may vary in some growth conditions and we imagine that this could

compromise some expression data. Should problematic behaviour arise, a solution would be to resort to adding *n*-decanal substrate such that the luminescence is more consistent.

The utility of this operon is evident in the fact that there was no exogenous background luminescence, strains bearing the promoterless control exhibited similar luminescence readings as strains lacking the genes altogether. This is not the case for reporters such as *egfp* (*S. coelicolor* exhibits significant autofluorescence), *xylE* or *amy.* We found that this reporter works exceptionally well in 96-well format, in turn permitting a large number of replicates for each measurement. While S1 nuclease analysis is an excellent tool for mapping 5' ends of transcripts and can be used for comparative analysis of promoter activities, it is simply too labour intensive to be used with this density of replicates.

A simple reporter system for the *Streptomyces* bacteria will be a very useful research tool. These data demonstrate that our synthetic *luxCDABE* operon is active in substrate and aerial hyphae and that it is not significantly toxic to either cell type. It provides a means of accurately monitoring expression of several genes as a function of colony growth and furthermore, can provide a simple means for detecting epistasis relationships between genes. We predict that this construct will be highly useful for studies in *S. coelicolor* and other high-GC bacteria.

2.5 Material & Methods

2.5.1 Bacterial strains, plasmids and culture conditions. Bacterial strains are listed in **Table 2.3**. Plasmids were propagated in the *E. coli* strain XL1 Blue. *E. coli* strains were grown at 37^{o}C in Luria broth media. For introduction of plasmids into *Streptomyces coelicolor,* plasmids were transformed into *E. coli* strain ET12567 containing the pUZ8002 plasmid allowing for conjugal transfer of the plasmid (29). *S. coelicolor* strains were grown on solid R2YE or MS agar at 30^{o}C. Antibiotic concentrations used were 100 µg/ml of ampicillin, 50 µg/ml of kanamycin, 34 µg/ml of chloramphenicol, 50 µg/ml of apramycin and 25 µg/ml naladixic acid.

Table 2.3. Strains used in this work

Strain	Genotype	Reference or Source
Streptomyces coelicolor A3(2)		
M145	Prototrophic, SCP1- SCP2-	(47)
M600	Prototrophic, SCP1- SCP2-	(48)
bldB	M145 *bldB::aphI*	(35)
ramR	M600 *ramR::vph*	Marie Elliot
whiG	M145 *whiG::hyg*	(36)
Escherichia coli		
XL1 Blue	*recA1 endA1 gyrA96 hsdR17 supE44 relA1 lac{F' proAB lacI^{Q}ZΔM15Tn10 (Tet^{r})]*	Stratagene
ET12567/pUZ8002	*dam13::*Tn*9 dcm-6 hsdM hsdR recF143 zjj-201::*Tn*10 galK2 galT22 ara14 lacY1 xyl-5 leuB6 thi-1 tonA31 rpsL136 hisG4 tsx-78 mtlI glnV44,* containing the nontransmissible *oriT* mobilizing plasmid pUZ8002	(29)

2.5.2 Sequence and oligonucleotide design. Coding sequences for each of the *P. luminescens lux* genes were designed *in silico,* substituting codons enriched in GC residues and known to be frequently found in highly expressed *S. coelicolor* genes. The resulting genes have a GC content of 69%, similar to that of the *S. coelicolor* chromosome. The length of flanking sequences were shortened such that, following operon assembly the only non-coding sequence upstream of *luxD, luxA, luxB* and *luxE* was the 12 bp required to accommodate a correctly positioned Shine-Dalgarno sequence. The site upstream of the *tuf1* gene of *S. coelicolor* was chosen for the Shine-Dalgarno and added upstream of each gene as we predicted that this would facilitate efficient initiation of translation.

Synthetic DNAs of ~100 nt were designed based on the top strand of each gene. Each oligonucleotide was analyzed using the ssDNA folding program mfold (see http://www.bioinfo.rpi.edu/applications/mfold/old/dna/). Sequences that exhibited significant intramolecular base pairing in the 20 nucleotides at their 5' or 3' end were modified to reduce the strength of, or eliminate, base pairing. This was accomplished by shifting the end by a few base pairs (and making a corresponding shift in the neighbouring oligonucleotide) and/or by making codon changes consistent with translation in *Streptomyces* that did not change the encoded protein sequence. 40-mer bridging oligonucleotides complementary to the 20 nt at the 5' and 3' ends of adjacent top strand oligonucleotides were then designed. The result of this was a set of 53 ~100-mer oligonucleotides encoding

the top strand of the five *lux* genes and 52 40-mer bridging oligonucleotides. All of these oligonucleotides were run on 10% denaturing polyacrylamide sequencing gels (29:1 acrylamide: bis-acrylamide), visualized by UV-shadow casting, excised, eluted into water and quantified by spectrophotometry. In addition to these assembly sequences, amplification oligonucleotides were designed to amplify each individual *lux* gene, introducing Gateway recombination sequences at each end and, just inside these recombination sequences, recognition sequences for restriction endonucleases that would be used for subsequent assembly of the genes into an operon. Primer sequences are available in the supplemental text of the published version of this chapter.

2.5.3 Assembly and amplification. Each *lux* gene was assembled in reactions containing the top strand and corresponding bridging oligonucleotides (both at 25 μM) in Pfu reaction buffer and 5 μM dNTPs. PCR cycle regimens for each gene were as follows: an initial denaturation step was carried out at 94^{o}C followed by 55 cycles of denaturation (94^{o}C for 1 min), annealing (45^{o}C for 40 sec) and extension (72^{o}C for 1 min). Following this, 15 μM of the appropriate amplifications primers were used to amplify each *lux* gene in reactions primed with 5 μl of assembly reaction – conditions for PCR were identical except that 25 cycles of amplification were used instead of 55 and a terminal extension incubation of 4 min, at 72 ^{o}C was added. All reactions were carried out using a robo-cycler 96 (Stratagene) with zero ramp time between temperatures. Fragments were

visualized by agarose gel electrophoresis, excised and purified using QIAEX II gel extraction.

2.5.4 Cloning, correction and operon assembly. Each amplified fragment was cloned into the Gateway entry vector pDONR221 by *in vitro* recombination as specified by the manufacturer. Transformants containing clones with inserts of the correct length were identified by colony PCR using the amplification primers. Multiple clones were isolated for each *lux* gene (see **Table 2.1**) and subjected to sequencing with the universal primers and, where necessary, primers internal to the cloned insert. Most of the cloned inserts had sequence errors relative to the expected sequence however some of these had no effect on the sequence of the encoded protein. One allele, *luxD2* encoded a perfect LuxD protein and was not subjected to correction mutagenesis. The best alleles of the other four genes: *luxC4, luxA22, luxB23* and *luxE5* were subjected to Quickchange mutagenesis to correct all sequence errors that affected the protein product. Finally, each gene was excised from the pDONR221 vector and assembled into an operon in pBluescript II SK+ (**Figure 2.1C**). Once assembled, the operon was excised as a 5668 bp NdeI-NotI fragment and inserted into the backbone of pRT801 cut with the same enzymes (30). The phage *fd* transcriptional terminator *tfd* and the phage λ transcriptional terminator *to* were cloned from pIJ8660 (8) to flank the *lux* operon. The terminator *tfd,* multiple cloning site and Shine-Dalgarno of pIJ8660 were cloned by adding a HindIII site upstream of *luxCDABE* at the BamHI site in

pRT801+*luxCDABE* with the oligonucleotides 5' GATCGGGAAGCTTCCC 3' and 5' CTAGCCCTTCGAAGGG 3' and digesting both pRT801+*luxCDABE* and pIJ8660 with HindIII and BamHI. The *to* terminator was removed from pIJ8660 by digesting with NotI and EcoRI and cloned into pRT801+(tfd)*luxCDABE* by blunting the EcoRI site with Klenow (New England Biolabs) and inserting this fragment downstream of *luxCDABE* by cutting the pRT801+(tfd)*luxCDABE* plasmid with NotI and EcoRV, creating the plasmid pMU1. The relevant structural details of pMU1 are shown in **Figure 2.2A**.

2.5.5 Cloning of the promoter fragments upstream of *luxCDABE*.

Promoter fragments for *hrdB*, *ramC* and *whiEp1* were amplified from *S. coelicolor* M145 chromosomal DNA using Pfu. These fragments were cloned into pMU1 using BamHI and KpnI in the multiple cloning site to create the plasmids, p^{hrdB}lux, p^{ramC}lux and p^{whiEp1}lux.

2.5.6 Growth Curves. Initial growth curves with pMU1 were performed as follows. 96-well white polystyrene plates (Fisher) were inoculated with strains to a confluent lawn into wells containing 200 µl of R2YE or MS agar with 12 replicates inoculated for each promoter-*lux* fusion. Plates were grown at 30^{o}C with luminescence read every 8 hours using the EnVision Multilabel Reader (PerkinElmer). Growth curves for the comparison of *whiEp1* expression in pMU1 and pMU1* were performed in 96-well white polystyrene plates (Fischer) with 1.0 x 10^4 spores inoculated inoculated per well. Plates were grown for 48 hr at 30^{o}C,

then covered with a plastic film and placed into the Victor Multilabel plate reader with luminescence read every hour until 148 hr of growth at room temperature. For the comparison of pMUs* and pF*lux*, 1.0 x10^4 spores of each promoterless vector was inoculated on R5M media grown at 30^{o}C and read for 68 hours with luminescence readings taken using the Victor Multilabel plate reader.

2.5.7 Correction of the *luxD* sequence. For correction of the *luxD* gene, four regions with amino acid sequence changes – N32T, T39P, Y236H and a three amino acid change starting at 246 GEN to CVF were corrected in the nucleotide sequence and this ideal sequence was purchased (Mr. Gene). The ideal *luxD* gene was then cloned into the *luxCDABE* operon using the HindIII and MfeI sites, creating the pMU1* plasmid.

2.6. References

1. Meighen EA (1993) Bacterial bioluminescence: Organization, regulation, and application of the *lux* genes. *FASEB J* 7: 1016-1022.

2. Sanli G, Blaber SI & Blaber M (2001) Reduction of wobble-position GC bases in *Corynebacteria* genes and enhancement of PCR and heterologous expression. *J Mol Microbiol Biotechnol* 3: 123-126.

3. Leskiw BK, Lawlor EJ, Fernandez-Abalos JM & Chater KF (1991) TTA codons in some genes prevent their expression in a class of developmental, antibiotic-negative, *Streptomyces* mutants. *Proc Natl Acad Sci U S A* 88: 2461-2465.

4. Kwak J, McCue LA & Kendrick KE (1996) Identification of *bldA* mutants of *Streptomyces griseus*. *Gene* 171: 75-78.

5. Garda AL, Fernandez-Abalos JM, Sanchez P, Ruiz-Arribas A & Santamaria RI (1997) Two genes encoding an endoglucanase and a cellulose-binding protein are clustered and co-regulated by a TTA codon in *Streptomyces halstedii* JM8. *Biochem J,* 324: 403-411.

6. Servin-Gonzalez L, Castro C, Perez C, Rubio M & Valdez,F. (1997) *bldA*-dependent expression of the *Streptomyces exfoliatus* M11 lipase gene (*lipA*) is mediated by the product of a contiguous gene, *lipR*, encoding a putative transcriptional activator. *J Bacteriol* 179: 7816-7826.

7. Trepanier NK, Jensen SE, Alexander DC & Leskiw BK (2002) The positive activator of cephamycin C and clavulanic acid production in *Streptomyces clavuligerus* is mistranslated in a *bldA* mutant. *Microbiology* 148: 643-656.

8. Sun J, Kelemen GH, Fernandez-Abalos JM &Bibb MJ (1999) Green fluorescent protein as a reporter for spatial and temporal gene expression in *Streptomyces coelicolor* A3(2). *Microbiology* 145: 2221-2227.

9. Grantcharova N, Lustig U and Flardh K (2005) Dynamics of FtsZ assembly during sporulation in *Streptomyces coelicolor* A3(2). *J Bacteriol* 187: 3227-3237.

10. Flardh K (2003) Essential role of DivIVA in polar growth and morphogenesis in *Streptomyces coelicolor* A3(2). *Mol Microbiol* 49: 1523-1536.

11. Jakimowicz D, Mouz S, Zakrzewska-Czerwinska J & Chater KF (2006) Developmental control of a *parAB* promoter leads to formation of sporulation-associated ParB complexes in *Streptomyces coelicolor*. *J Bacteriol* 188: 1710-1720.

12. Ruban-Osmialowska B, JakimowiczD, Smulczyk-Krawczyszyn A, Chater KF & Zakrzewska-Czerwinska J (2006) Replisome localization in vegetative and aerial hyphae of *Streptomyces coelicolor*. *J Bacteriol* 188: 7311-7316.

13. O'Connor TJ, Kanellis P & Nodwell JR (2002) The *ramC* gene is required for morphogenesis in *Streptomyces coelicolor* and expressed in a cell type-specific manner under the direct control of RamR. *Mol Microbiol* 45: 45-57.

14. Ingram C, Brawner M, Youngman P & Westpheling J (1989) *xylE* functions as an efficient reporter gene in *Streptomyces* spp.: Use for the study of *galP1*, a catabolite-controlled promoter. *J Bacteriol* 171: 6617-6624.

15. Flores FJ, Rincon J & Martin JF (2003) Characterization of the iron-regulated *desA* promoter of *Streptomyces pilosus* as a system for controlled gene expression in actinomycetes. *Microb Cell Fact* 2: 5.

16. Paget MS, Hintermann G & Smith CP (1994) Construction and application of streptomycete promoter probe vectors which employ the *Streptomyces glaucescens* tyrosinase-encoding gene as reporter. *Gene* 146: 105-110.

17. Rodriguez-Garcia A, Combes P, Perez-Redondo R, Smith MC & Smith MC (2005) Natural and synthetic tetracycline-inducible promoters for use in the antibiotic-producing bacteria *Streptomyces*. *Nucleic Acids Res* 33: e87.

18. Berdy J. (2005) Bioactive microbial metabolites. *J Antibiot (Tokyo)* 58: 1-26.

19. Loria R, Kers J and Joshi M (2006) Evolution of plant pathogenicity in *Streptomyces*. *Annu Rev Phytopathol* 44: 469-487.

20. Kelemen GH & Buttner MJ (1998) Initiation of aerial mycelium formation in *Streptomyces*. *Curr Opin Microbiol* 1: 656-662.

21. Chater KF (2001) Regulation of sporulation in *Streptomyces coelicolor* A3(2): A checkpoint multiplex? *Curr Opin Microbiol:* 667-673.

22. Willey JM, Willems A, Kodani S & Nodwell,J.R. (2006) Morphogenetic surfactants and their role in the formation of aerial hyphae in *Streptomyces coelicolor*. *Mol Microbiol* 59: 731-742.

23. Schwedock J, McCormick JR, Angert ER, Nodwell JR & Losick,R. (1997) Assembly of the cell division protein FtsZ into ladder-like structures in the aerial hyphae of *Streptomyces coelicolor*. *Mol Microbiol* 25: 847-858.

24. Flardh K, Leibovitz E, Buttner MJ & Chater KF (2000) Generation of a non-sporulating strain of *Streptomyces coelicolor* A3(2) by the manipulation of a developmentally controlled *ftsZ* promoter. *Mol Microbiol* 38: 737-749.

25. Stahler P, Beier M, Gao X & Hoheisel JD (2006) Another side of genomics: Synthetic biology as a means for the exploitation of whole-genome sequence information. *J Biotechnol* 124: 206-212.

26. Wu G, Zheng Y, Qureshi I, Zin HT, Beck T, Bulka B & Freeland SJ (2006) SGDB: A database of synthetic genes re-designed for optimizing protein over-expression. *Nucleic Acids Res*. 10.1093/nar/gkl648.

27. Au LC, Yang FY, Yang WJ, Lo SH. & Kao CF (1998) Gene synthesis by a LCR-based approach: High-level production of leptin-L54 using synthetic gene in *Escherichia coli*. *Biochem Biophys Res Commun* 248: 200-203.

28. Stemmer WP, Crameri A, Ha KD, Brennan TM & Heyneker HL (1995) Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides. *Gene* 164: 49-53.

29. Flett F, Mersinias V & Smith CP (1997) High efficiency intergeneric conjugal transfer of plasmid DNA from *Escherichia coli* to methyl DNA-restricting streptomycetes. *FEMS Microbiol Lett* 155: 223-229.

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

31. Hudson ME & Nodwell JR (2004) Dimerization of the RamC morphogenetic protein of *Streptomyces coelicolor*. *J Bacteriol* 186: 1330-1336.

32. Kodani S, Hudson ME, Durrant MC, Buttner MJ, Nodwell JR & Willey JM (2004) The SapB morphogen is a lantibiotic-like peptide derived from the product of the developmental gene *ramS* in *Streptomyces coelicolor*. *Proc Natl Acad Sci U S A* 101: 11448-11453.

33. Willey J, Schwedock J & Losick R (1993) Multiple extracellular signals govern the production of a morphogenetic protein involved in aerial mycelium formation by *Streptomyces coelicolor*. *Genes Dev* 7: 895-903.

34. Kelemen GH, Brian P, Flardh K, Chamberlin L, Chater KF & Buttner MJ (1998) Developmental regulation of transcription of *whiE*, a locus specifying the polyketide spore pigment in *Streptomyces coelicolor* A3 (2). *J Bacteriol,* 180: 2515-2521.

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

36. Flardh K, Findlay KC & Chater KF (1999) Association of early sporulation genes with suggested developmental decision points in *Streptomyces coelicolor* A3(2). *Microbiology* 145: 2229-2243.

37. Willey J, Santamaria R, Guijarro J, Geistlich M & Losick R (1991) Extracellular complementation of a developmental mutation implicates a small sporulation protein in aerial mycelium formation by *S. coelicolor*. *Cell* 65: 641-650.

38. Yamazaki H, Ohnishi Y & Horinouchi S (2000) An A-factor-dependent extracytoplasmic function sigma factor (sigma(AdsA)) that is essential for morphological development in *Streptomyces griseus*. *J Bacteriol* 182: 4596-4605.

39. Ryding NJ, Kelemen GH, Whatling CA, Flardh K, Buttner MJ and Chater KF (1998) A developmentally regulated gene encoding a repressor-like protein is essential for sporulation in *Streptomyces coelicolor* A3(2). *Mol Microbiol* 29: 343-357.

40. Lee EJ, *et al* (2005) A master regulator sigmaB governs osmotic and oxidative response as well as differentiation via a network of sigma factors in *Streptomyces coelicolor*. *Mol Microbiol* 57: 1252-1264.

41. Kelemen GH, Brown GL, Kormanec J, Potuckova L, Chater KF & Buttner MJ (1996) The positions of the sigma-factor genes, *whiG* and *sigF*, in the hierarchy controlling the development of spore chains in the aerial hyphae of *Streptomyces coelicolor* A3(2). *Mol Microbiol* 21: 593-603.

42. Bibb MJ, Molle V & Buttner MJ (2000) Sigma(BldN), an extracytoplasmic function RNA polymerase sigma factor required for aerial mycelium formation in *Streptomyces coelicolor* A3(2). *J Bacteriol* 182: 4606-4616.

43. Mendez C & Chater KF (1987) Cloning of *whiG*, a gene critical for sporulation of *Streptomyces coelicolor* A3(2). *J Bacteriol* 169: 5715-5720.

44. Potuckova L, Kelemen GH, Findlay KC, Lonetto MA, Buttner MJ & Kormanec,J. (1995) A new RNA polymerase sigma factor, sigma F, is required for the late stages of morphological differentiation in *Streptomyces* spp. *Mol Microbiol* 17: 37-48.

45. Kormanec J, Sevcikova B, Halgasova N, Knirschova R & Rezuchova B (2000) Identification and transcriptional characterization of the gene encoding the stress-response sigma factor sigma(H) in *Streptomyces coelicolor* A3(2). *FEMS Microbiol Lett* 189: 31-38.

46. Fujii T, Gramajo HC, Takano E and Bibb MJ (1996) *redD* and *actII*-ORF4, pathway-specific regulatory genes for antibiotic production in *Streptomyces coelicolor* A3(2), are transcribed in vitro by an RNA polymerase holoenzyme containing sigma HrdD. *J Bacteriol,* 178: 3402-3405.

47. Kieser T, Bibb M, Buttner MJ, Chater K & Hopwood DA (2000) Practical *Streptomyces* Genetics. The John Innes Foundation, Norwich.

48. Chakraburtty R and Bibb M (1997) The ppGpp synthetase gene (*relA*) of *Streptomyces coelicolor* A3(2) plays a conditional role in antibiotic production and morphological differentiation. *J Bacteriol* 179: 5854-5861.

CHAPTER 3

A SCREEN FOR SMALL MOLECULE MODULATORS OF *S. COELICOLOR* GROWTH, DEVELOPMENT AND ANTIBIOTIC PRODUCTION

A portion of this chapter was published in the supplemental information of:

Arryn Craney, Cory Ozimok, Sheila Marie Pimentel-Elardo, Alfredo Capretta and Justin Nodwell. Chemical perturbation of secondary metabolism demonstrates important links to primary metabolism. *Chem. Bio*. 2012; 19(8):1020-1027.

3.1 Abstract

The phenotypic richness of *Streptomyces coelicolor* has been fundamental for genetic studies in this bacterium as alterations in growth, secondary metabolism and development can all be examined visually. We have used this visual progression to take a parallel method of analysis and identify small molecules probes of the *Streptomyces* lifecycle. This chemical biological approach uses small molecules to act like mutations and alter protein function. Understanding how these small molecules exert their effects will aid in further elucidating the mechanisms involved in *S. coelicolor* secondary metabolism and development. As well, these small molecules can be added to other streptomycetes to understand the conservation of their effects and, in the case of secondary metabolites, to trigger enhanced production of compounds of interest.

3.2 Introduction

The streptomycetes have adopted an unusual reproductive strategy for a prokaryote that is reminiscent of the filamentous fungi and progresses through an ordered set of events involving multiple cell types (**Figure 3.1**) (1). *Streptomyces coelicolor* is an excellent model for this developmental process owing to the visible cues that accompany each stage (**Figure 3.1**) (2, 3).

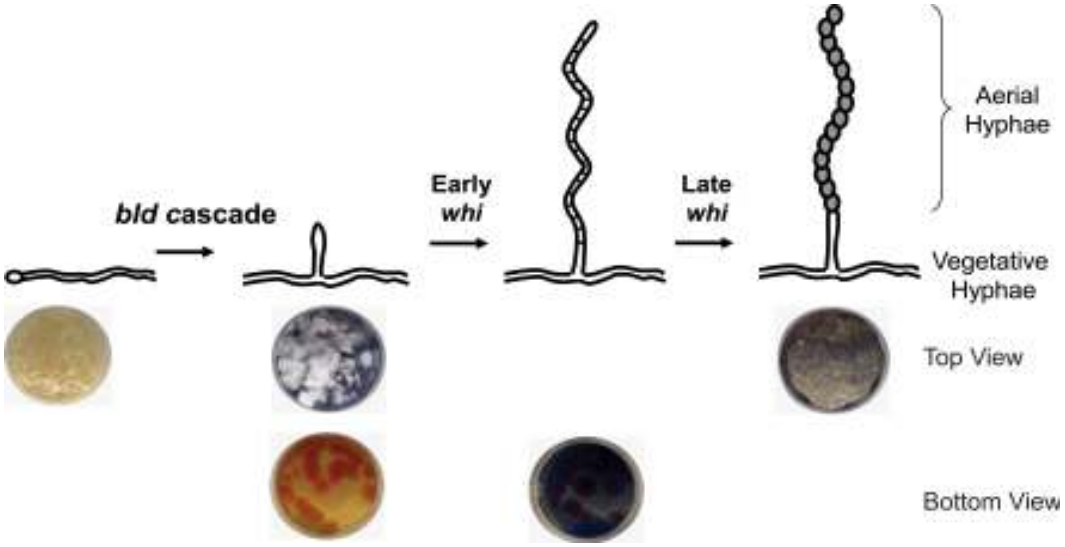

Figure 3.1. The *S. coelicolor* lifecycle. Growth begins with spore germination and the formation of the vegetative hyphae. Surfactant production is essential for aerial hyphae formation and it is believed the role of the extracellular signalling *bld* cascade is to produce and regulate the production of these surfactants (SapB and/or the chaplins). Aerial hyphae rise up from the vegetative hyphae and impart a white phenotype to the growing colony. Around the onset of aerial hyphae formation, the vegetative hyphae produce secondary metabolites. *S. coelicolor* produces two pigments; red pigmented prodiginines are produced first, followed by the production of the blue pigment actinorhodin. A top view of the well permits visualization of growth (presence of colonies), cell differentiation (demonstrated by the white fuzzy layer of aerial hyphae) and sporulation (grey pigmentation of mature spores). A bottom view of the well permits visualization of prodiginine (red) and actinorhodin (blue) production.

Growth begins with spore germination and the formation of the substrate or vegetative mycelium of filamentous cells. The second cell type, the aerial hyphae grow up from the vegetative hyphae into the air approximately 48 hours after vegetative growth begins. These two cell types have very different structures and fates. The vegetative hyphae undergo extensive branching to form a dense mat; cell division is rare in this cell type resulting in long multi-genomic compartments. The aerial hyphae do not branch and undergo a synchronous round of division to produce unigenomic compartments. The function of these cell types is also very different. The aerial hyphae mature into spores; providing a means of dispersion, reproduction and survival for the bacterium. In contrast, secondary metabolites are primarily produced in the substrate hyphae and presumably provide a competitive advantage relative to other environmental organisms.

Every step of the lifecycle is tightly regulated; ensuring both secondary metabolism and development occurs when beneficial to the cell. Regulation of secondary metabolism was discussed in Chapter 1. Many of the genes involved in aerial hyphae formation are referred to as *bld* genes as their deletion prevents aerial hyphae formation, leaving the colony bald in appearance (**Figure 3.1**) (4, 5). Of particular importance to aerial hyphae formation is the lowering of surface tension to allow the aerial hyphae to rise into the air. *S. coelicolor* achieves this by the production of SapB and the chaplins (6). SapB is produced by the *ramCSAB* operon which encodes most of the enzymes for its production and

export (7-10). Secretion of this amphipathic molecule lowers the surface tension (6). The chaplins function in a similar manner. These hydrophobic molecules are produced as a mixture, the short chaplins are secreted into the media and the long chaplins are attached to the cell wall. Their production forms a hydrophobic sheath around the hyphae which also lowers the surface tension and allows the aerial hyphae to rise into the air (11). Once formed, the aerial hyphae go on to develop into mature, grey pigmented spores (12). Genes involved in spore formation are referred to as *whi*, as they leave the colony white in appearance when deleted (**Figure 3.1**) (13). Events in the aerial hyphae can be subdivided into early (prior to cell division) and late processes (post-septation) (14, 15). Interestingly cell division is not required for cell viability. In the absence of the usually essential major cell division protein FtsZ, *S. coelicolor* can grow vegetatively and erect aerial hyphae; however, they cannot undergo septation of the aerial hyphae and sporulation is therefore prevented (16, 17). This represents a unique opportunity to study a generally essential process.

The phenotypic versatility of *Streptomyces coelicolor,* particularly during growth on solid media, has been vital to the understanding of factors involved in growth, development and secondary metabolism. However, new technologies are required to further understand and exploit these processes. A chemical biological strategy offers significant advantages for studying the streptomycetes. Information attained from genetic manipulation can be difficult to apply widely in

the streptomycetes due to the extensive restriction barriers. However, small molecules are readily transferable between organisms. As well, the majority of developmental proteins identified to date have regulatory roles, small molecule manipulation may provide a means to identify proteins involved in enzymatic processes. Finally, small molecules are fast acting and this may result in less compensation by competing pathways and a stronger effect than some genetic mutations. Making use of the visual cues of the *S. coelicolor* lifecycle enables a non-invasive screening method in which all aspects can be examined by eye in a single screen. For example, *S. coelicolor* colony morphology provides cues for growth (the presence or absence of a colony), secondary metabolism (the red and blue pigmented metabolites prodiginine and actinorhodin, respectively), the formation of spore forming cells (the white fuzzy layer of aerial hyphae) and the completion of spore formation (a grey polyketide pigment that coats mature spores) (**Figure 3.1**).

Through a chemical screen against the visual cues of the *S. coelicolor* lifecycle, we have identified small molecules which recapitulate all genetic phenotypes observed. Of particular interest are: small molecules that inhibit *S. coelicolor* growth for their potential as antimicrobial agents; small molecules that interfere with sporulation, as they may also have application as antimicrobial small molecules – in particular, as cell division inhibitors; and those small molecules

that enhance secondary metabolism, as they have the potential application to enhance the secondary metabolites of other streptomycetes.

3.3 Results & Discussion

3.3.1 Screening for bioactive small molecules

To identify small molecules with bioactivity against *S. coelicolor*, we screened the Canadian Compound Collection (CCC) and a small library of kinase inhibitors, the CDK library, (for a total of 30,989) for alterations in the developmental process. Compounds were dissolved in DMSO and screened at 10 μM in duplicate, the final concentration of DMSO per well was 1%. Daily visual inspection in comparison to DMSO controls were recorded over a 10 day period. We were able to identify hits for every aspect of the *S. coelicolor* lifecycle that could be assessed visually. These primary hits are summarized in **Table 3.1**.

Table 3.1. Small molecule induced changes from HTS screen

Phenotype	Number of Hits
Growth	
Complete Growth Inhibition	105
Slow Growth	244
Development	
Inhibition of Aerial Hyphae Formation (bald)	50
Inhibition of Sporulation (white)	196
Precocious Sporulation	137
Secondary Metabolism	
Decreased Actinorhodin and/or Prodiginine	240
Increased Actinorhodin and/or Prodiginine	112

For further characterization of these primary hits, the CCC library was divided into the bioactive subset and the remaining synthetic small molecule subset. The bioactive subset contains small molecules with known bioactive properties such as antibiotics and other drugs. This division was made as the bioactive subset was screened prior to the remainder of the CCC and further progress has been made on their effects. The CDK library was not subjected to follow up as we were unable to purchase these compounds.

3.3.2 Small molecules affecting *S. coelicolor* growth

Small molecules that inhibit S. coelicolor growth may be worthwhile antimicrobial agents

While the streptomycetes are not typical screening organisms for antimicrobial agents, there are many reasons that small molecules capable of inhibiting their growth may be of value. These high–GC Gram positive bacteria are related to many pathogenic bacteria within the actinobacteria, including *Mycobacterium leprae*, *Mycobacterium tuberculosis* and *Corynebacteria diphtheria* (the causative agents of leprosy, tuberculosis and diphtheria, respectively). *S. coelicolor* can be used as a non-pathogenic alternative. There are many actinomycete-specific proteins and *S. coelicolor* can be used to determine their potential as antimicrobial targets. For example, the WhiB-like (Wbl) family is a conserved family of proteins containing a redox sensing iron cluster core, found exclusively within the actinomycetes (18). This family is named after the first identified

protein WhiB in *S. coelicolor* which is involved in sporulation (19). *M. tuberculosis* encodes seven Wbl proteins of which, *whiB1* is essential for viability and *whiB3* is required for virulence, suggesting targeting the Wbl family may have antimicrobial value (20). As well, any small molecule able to kill *S. coelicolor* may have value as an antimicrobial if it possesses a novel mode of action that is well-conserved in bacteria or, if its activity is more potent than currently available antimicrobials.

The majority of small molecules from the bioactive subset that were capable of inhibiting *S. coelicolor* growth were potent antimicrobials, primarily used as antifungals and disinfectants (**Table 3.2**). Of the antifungals identified in the screen, eight were from the azole class of antibiotics. They target an essential cytochrome P450 14α demethylase in fungi and are used clinically for their potent antifungal activity (21). Azoles also have known antibacterial properties, including activity against *Mycobacterium* and *Streptomyces* (22). Actinomycetes possess a large number of cytochrome P450s (*M. tuberculosis* is predicted to encode 20 P450s), including homologues to the fungal P450 demethylase target of the azoles (23). *S. coelicolor* sensitivity to azole antifungals demonstrates its usefulness as an indicator strain for antimicrobials. Disinfectants are generally used to sterilize surfaces from microbial contaminants and are used for their broad spectrum, non-specific growth inhibition. The presence of these

disinfectants suggests that many of the growth inhibitors from the synthetic small molecule subset of the CCC will be toxic non-specific growth inhibitors.

Very few growth inhibitors from the bioactive subset were antibiotics from streptomycete sources, only 5 of the 43 (**Table 3.2**). Of these, *S. coelicolor* is known to be sensitive to apramycin, thiostrepton and viomycin as these are used as selectable markers for genetic manipulation. Our screen also demonstrated sensitivity to tunicamycin B and echinomycin. Tunicamycin B is a nucleoside antibiotic produced by *S. lysosuperficus* which inhibits protein glycosylation (24). Since resistance determinants are known (TmrB in *Bacillus subtilis* (25)), and tunicamycin B is commercially available, this antibiotic may have application as a resistance cassette in *S. coelicolor* for genetic analysis. Echinomycin is a non-ribosomal peptide produced by *S. lasalienis*. Echinomycin is a potent antimicrobial and has been demonstrated to inhibit the growth of methicillin-resistance *Staphylococcus aureus* (26). The ability of five known antibiotics from the bioactive subset to kill *S. coelicolor* further reiterates the effectiveness of using *S. coelicolor* as an indicator for growth inhibition.

Table 3.2. Known Growth Inhibitors in the CCC (Bioactives)

Type	Number
Antibiotics produced by *Streptomyces*	5
Antifungals (azoles)	8
Antifungals (others)	8
Disinfectants	9
Other	13
Total	43

Small molecules resulting in slow growth of S. coelicolor may also be useful antimicrobial agents

Slower growth by *S. coelicolor* was defined by at least a 2 day delay in visible growth in comparison to the DMSO controls. The 244 small molecules in this category are most likely made up molecules with inhibitory concentrations higher than the 10 μM screening concentration and small molecules with antimicrobial properties that *S. coelicolor* has resistance mechanisms to. For example, florfenicol gave a slow growth phenotype and is a synthetic derivative of chloramphenicol. *S. coelicolor* exhibits moderate resistance to chloramphenicol, but high concentrations result in growth inhibition (27). Bacitracin also gives a slow growth phenotype and is known to inhibit *S. coelicolor* at concentrations higher than the primary screening concentrations (28). This suggests higher concentrations of many of these molecules are required for complete growth inhibition. However, the slow growth caused by some small molecules may also be due to resistance mechanism present in *S. coelicolor*. Over half the bioactive compounds that resulted in slow growth are antimicrobials produced by other streptomycetes (**Table 3.3**). As many streptomycetes share the same ecological niche it is plausible that strains share self-protection mechanisms. *S. coelicolor* has known resistance mechanisms to vancomycin (29, 28), tetracycline antibiotics (30), macrolides and lincosamides (31), all of which are present in the slow growth category (**Table 3.3**). While *S. coelicolor* may have resistance to many of these molecules, some are currently used as clinical antibiotics with

potent antimicrobial activity. Vancomycin is used as a last resort drug when other antibiotics fail, for example in treatment of multidrug resistance *S. aureus*. Thus, those small molecules with resistance in *S. coelicolor* may have potent antimicrobial activity on non-resistant bacteria and their activity on other bacteria should be explored.

Table 3.3. Slow growth bioactives produced by the streptomycetes

Compound	Streptomycete/Actinomycete	Reference
Produced by an actinomycete – known resistance mechanism in S. coelicolor vanRS		
Vancomycin	*Amycolatopsis orientalis*	(29)
Tetracyclines		(30)
Chlortetracycline	*S. aureofaciens*	
Minocycline	semi-synthetic derivative of tetracycline	
Doxycycline	semi-synthetic derivative of tetracycline	
Methacycline	semi-synthetic derivative of tetracycline	
Macrolide-Lincosamide-Streptogramin (MLS) resistance		
Lincosamide antibiotics		
Clindamycin	semi-synthetic derivative of clindamycin	(32)
Lincomycin	*S. lincolnensis*	(33)
Macrolides		
Midecamycin	*S. mycarofaciens*	(31)
Erythromycin	*S. erthyrea*	(34)
Roxithromycin	semi-synthetic derivative of erythromycin	
Azithromycin	semi-synthetic derivative of erythromycin	
Tylosin	*S. fradiae*	(35)
Troleandomycin	synthetic derivative of oleandomycin (*S. antibioticus*)	
Produced by a Streptomycete – no known resistance mechanism in S. coelicolor		
Doxorubicin	*S. peucetius*	
Daunorubicin	*S. peucetius*	
Oligomycin	*S. avermitilis*	
Tobramycin	*S. tenebrarius*	
Rifampicin	*Amycolatopsis rifamycinica*	
Nigericin	*S. hygroscopicus*	
Chromomycin A3	*S. griseus*	
Antibiotic A-23187	*S. chartreusensis*	
Lasalocid	*S. lasaliensis*	
Narasin	*S. albus*	
Florfenicol	related to chloramphenicol (*S. venezuelae*)	

Exploring the utility of synthetic small molecules that affect S. coelicolor growth

In general, those small molecules from the bioactive subset that were able to reduce growth in *S. coelicolor* were potent antimicrobials. Based on this, it may be possible to identify potent antimicrobial agents from the remaining synthetic small molecules subset (62 of the 105 growth inhibitors). In support of this, 3 clusters of small molecules were identified from the synthetic growth inhibitors (**Figure 3.2**). Cluster 1 was also identified in a screen against growth in *S. aureus* (36). These small molecules have been demonstrated to inhibit transcription and translation in *S. aureus*. Cluster 2 also has the capacity to kill *M. tuberculosis*; although the mechanism of action is unclear (37). Cluster 3 was not identified in any other chemical screen.

To fully explore the small molecules affecting *S. coelicolor* growth (growth inhibition and slow growth), titrations on a panel of microbes should be performed to identify potent antimicrobial agents for further study. This panel would represent a diverse range of organisms (Gram negative, Gram positive and fungi) as it is important to determine the spectrum and potency of antimicrobial activity and avoid those small molecules that are generally toxic to all cells. Of interest from the current data is identifying the spectrum of activity and mechanism of action of Cluster 2 and Cluster 3.

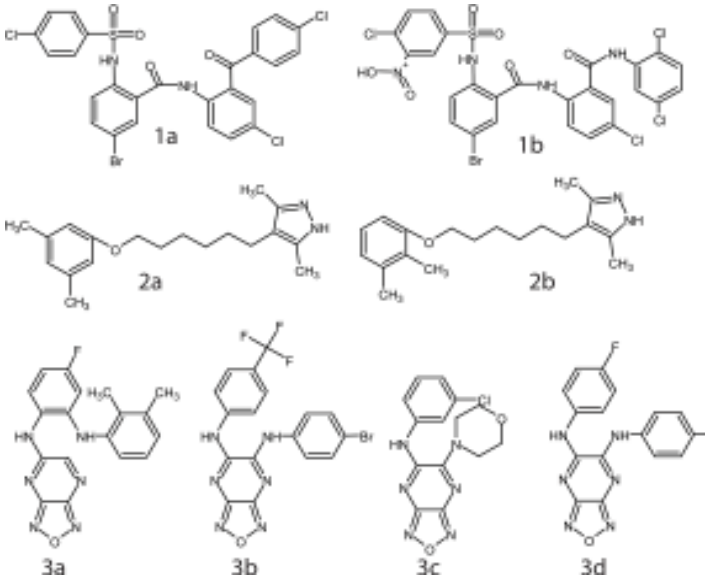

Figure 3.2. Clusters of synthetic small molecules capable of inhibiting *S. coelicolor*. Cluster 1 – also identified as growth inhibitors in *S. aureus* and demonstrated to inhibit transcription and translation. Cluster 2 – also identified as inhibitors of *M. tuberculosis* growth, the mechanism of inhibition is unknown. Cluster 3 – no reported antimicrobial activity.

Depending on the spectrum of activity there are many tools available for target identification. Libraries are available in many bacteria where proteins have been individually over-expressed or knocked down (38, 39). These libraries can be screened for changes in the inhibitory concentration of the small molecule, providing a possible target. For example, platencimycin an antibiotic produced by *S. platensis* was identified as an inhibitor of FabF due to increased sensitivity in the indicator strain *S. aureus* when FabF was knocked down by antisense RNA (40). The development of resistant mutants can also be used as these mutations often map to the target process. Finally, effects in combination with known

antimicrobials can provide the target process and can be performed in any sensitive strain (see (39, 41) for reviews on techniques to identify targets).

3.3.3 Molecules impairing the developmental process (bald, white)

Small molecules preventing aerial hyphae formation

Interestingly, very few compounds prevented aerial hyphae formation (50 compounds resulted in a bald phenotype). Follow up with the bioactive subset (28 compounds), resulted in only a single small molecule which reproduced the bald phenotype, cinchonine (**Figure 3.3**). Cinchonine is an alkaloid produced by fungi within the Cinchona plants (42). The ability of this metabolite to inhibit *S. coelicolor* development may be a strategy of self-defence when growing in the same niche. Aerial hyphae formation is known to be inhibited by surfactin, a secondary metabolite produced by *Bacillus subtilis*, also resulting in a bald phenotype and is thought to be a defense strategy (43).

Figure 3.3. Cinchonine inhibits aerial hyphae formation in *S. coelicolor*. Cinchonine was added at 10 µM to *S. coelicolor* M145 on R5M. Growth after 7 days was compared to the DMSO control in duplicate. The image was scanned from the top of the plate to capture changes in development. Grey pigmentation in the DMSO control signified sporulation. Cinchonine addition results in a bald phenotype.

The remaining synthetic subset has yet to be explored. Confirming the bald phenotype and determining the mechanism in which they prevent aerial hyphae formation will be of importance for this class of small molecule. The majority of proteins identified to date are involved in regulatory aspects and the use of small molecule inhibitors may identify enzymes involved in aerial hyphae formation.

Small molecules interfering with sporulation

196 small molecules were capable of interfering with spore development and were identified by interfering with visible grey pigmentation. The colour of the aerial hyphae in the presence of these compounds ranged from white to pale grey (**Figure 3.4A**). Since the grey pigment is developed late in the lifecycle, small molecules interfering with this pigment could represent alterations in any process during the lead up to mature spore formation.

The lifecycle of *S. coelicolor* can be thought of as a hierarchy of gene expression, as precise expression at each stage of growth in the correct cell type is required to execute the entire process. Based on this hierarchy it should be possible to map the action of these molecules to the stage of development affected (2, 4, 14, 15, 44, 45). For example, it would be expected that all molecules would show reduced expression of *whiE* as this measures grey pigment production – the process used to first identify these molecules. To begin dividing these compounds into possible processes, we created a library of promoters fused to

the *luxCDABE* operon of genes involved in various stages of development (**Figure 3.5**). Although this library is not a complete representation of the genes involved in development, genes from each stage of development have been included and provide an initial set of promoters to begin mapping the developmental blocks.

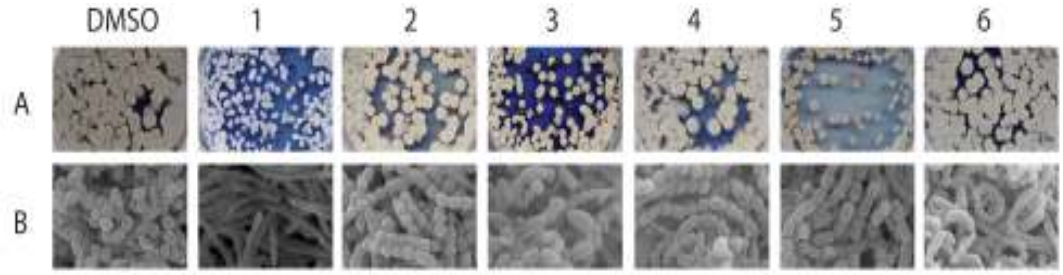

Figure 3.4. Examples of molecules causing a white phenotype. Compounds were added at 10 µM to *S. coelicolor* grown on R5M. Images were taken after 7 days growth. (**A**) Plate images of *S. coelicolor*; small molecule addition results in phenotypes from white (1) to off-white (2 and 3) to pale grey (4-6) compared to the deeply grey DMSO control. (1 - lincomycin, 2 - minocycline, 3 - bleomycin, 4 - carmofur , 5 - roxithromycin, 6 - chlorphensin) (**B**) Scanning electron micrographs in the presence of molecules causing a white phenotype. Abundant septated spores can be seen in the DMSO control. No spores were detected in the presence of compound 1. Compound 2 resulted in the occasional double sized spore. Compound 3 resulted in over-sized spores with a lima bean shape. Compound 4 resulted in misshapen spores. Compound 5 spores did not round up resulting in a box-like appearance. Spores in the presence of compound 6 were generally longer and curled.

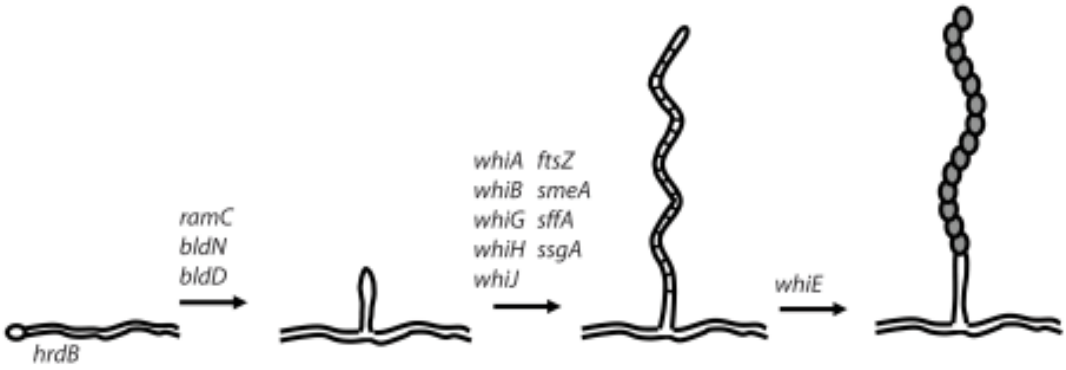

Figure 3.5. Promoter-*lux* fusion library for studying *S. coelicolor* development. 14 developmental promoters were fused upstream of the *lux* operon. *hrdB* is a vegetative sigma factor used as a control for gene expression analysis. *ramC*, *bldN* and *bldD* are required for aerial hyphae formation (9, 46, 47). *whiA*, *whiB*, *whiG*, *whiH* and *whiJ* are required prior to septation in the aerial hyphae (14). *ftsZ*, *smeA*, *sffA* and *ssgA* are involved in septation in the aerial hyphae (2). *whiE* is required for grey pigment production in the mature spore (48).

Three promoter fusions were chosen to broadly map the processes affected: *hrdB*p, *ramC*p and *whiE*p1. *hrdB*p was used as a measure of growth, as this gene encodes the vegetative housekeeping sigma factor and is often used as a control (49). Defects in the expression of *hrdB* suggest these molecules are generally toxic to the cell and the cells therefore have slowed growth, resulting in impaired sporulation. Since these molecules would have non-specific effects on development, they should be removed from subsequent analysis. *ramC*p was used to measure aerial hyphae formation as it is involved in the production of the surfactant SapB which is required at this stage of development (9, 45, 50, 51). Reduced expression of *ramC*p would occur if the molecules interfere prior to aerial hyphae formation and suggest processes in the vegetative hyphae are

most likely being targeted. Finally, *whiE*p1 was used as a measurement for grey pigmentation as genes in this operon encodes for production of the grey pigment (12). Since all molecules were originally identified by defects in grey pigmentation, all molecules would result in reduced *whiE*p1 expression. Those molecules that do not interfere with *whiE* expression are most likely false positives from the primary screen.

Table 3.4. Summary of small molecules impairing sporulation

Process Affected	Change in *lux* fusions	Number of molecules
Growth Inhibition	delay in *hrdB*, *ramC* and *whiE*	35
Aerial Hyphae Formation	decrease/delay in *ramC* only	10
Aerial Hyphae Development	decrease/delay in *ramC* and *whiE*	26
Aerial Hyphae Maturation	delay/decrease in *whiE* only	27
False Positive	no change	74

The 176 small molecules (both bioactive and synthetic subsets from the CCC) were grouped by their effect on these promoters (**Table 3.4**). Three groups stand out, those that interfere early in development only affecting *ramC* promoter activity (implying these molecules block the development process prior to aerial hyphae formation), those that interfere with *ramC* and *whiE* promoter activity (these molecules may block all stages of aerial hyphae formation and maturation) and those that only affect expression of *whiE* (suggesting these small molecules affect processes after aerial hyphae formation). Applying these molecules to the remaining promoter fusions in the library will further map the stage of developmental block.

Once these molecules have been mapped to the stages of growth they effect, over-expression studies can be performed to identify the target pathways. Over-expression of development genes may bypass the ability of the small molecules to interfere with development and imply these inhibitors target that particular pathway. In addition, microscopy can be used to determine the phenotype in the presence of these small molecules and comparisons can be made to known genetic defects (**Figure 3.4B**). Chemically modified versions of the molecules can also be used to link the small molecule to its protein target and identify the target protein. Finally, transcriptional analysis can be performed to assess global changes in the presence of these molecules – possibly identifying new processes involved in *S. coelicolor* development.

Small molecules impairing sporulation may have application as antimicrobial agents

Cell division is only required for spore formation in *S. coelicolor* and is not essential for vegetative growth (17). This presents an interesting avenue as some small molecules with a white phenotype may be interfering with the cell division apparatus. The essential nature of cell division in all other bacteria means that screens that target cell division are generally *in vitro* assays against the GTPase activity of the main cell division protein FtsZ (52). *S. coelicolor* offers a cell-based assay to identify cell division inhibitors and the opportunity to target other aspects of the cell division apparatus. One caveat to this approach is cell

division can also be arrested by DNA damage resulting in the SOS response (53-56). Those small molecules capable of inducing the SOS response were eliminated (only 1 molecule from the synthetic subset, Charul Jani and Justin Nodwell, unpublished data).

In the pursuit of cell division inhibitors, Charul Jani has screened this class of molecules for filamentation of *Bacillus subtilis*, a hallmark of inhibition of cell division, and is currently working to determine the target of three filament-inducing molecules. Of these, 2 are believed to target proteins other than FtsZ (Charul Jani and Justin Nodwell, unpublished data).

3.3.4 Precocious sporulation

One hundred and thirty seven small molecules were able to reduce the time to sporulation, defined by the ability to produce grey pigmentation at least one day earlier than the DMSO controls. The majority of genetic factors involved in sporulation to date have been identified by their inhibitory effects on development when deleted or mutated; using stimulatory small molecules on the developmental process would approach studying sporulation from a different perspective.

These compounds could affect many processes resulting in early sporulation. They could alter the nutritional status (eliciting a stress response resulting in sporulation or improving growth leading to early sporulation), or change events relevant to aerial hyphae formation (for example, acting like a surfactant) and aerial hyphae development (for example, improve production of the grey pigment itself). The developmental promoter-*lux* fusion library created for studying the molecules which impair sporulation can be used to determine the stages of development affected by this class of molecules. **Figure 3.6A** shows two examples from the bioactive subset which improve production of *whiE*p1 expression.

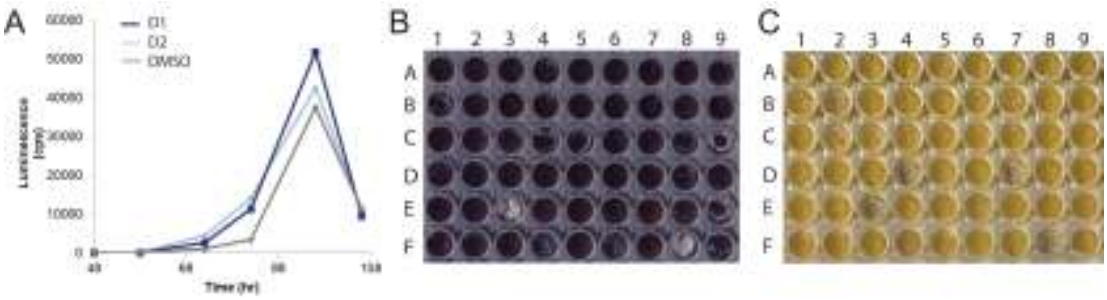

Figure 3.6. Examples of follow up assays with the molecules causing precocious sporulation. The bioactive subset of the CCC causing faster sporulation were analysed at 10 µM. (**A**) Improved *whiE*p1-*lux* expression in the presence of D1 and D2 (D1 - MK-886 and D2 - Oleic Acid) on R5M media. (**B**) 2 faster molecules were able to complement aerial hyphae formation in the *ramC* null mutant on R5M (E3 - theaflavin digallate and F4 - rutilantinone). (**C**) Ability of the faster molecules to improve sporulation of *S. lividans* after 7 days growth on R5M (B2 - Geranylgeranoic acid, D4 - N-oleoyldopamine, D7 - U-73122, E3 - theaflavin digallate and F4 - rutilantinone).

Mutants deficient in aerial hyphae formation (bald) or sporulation (white) can be used to assess the ability of these molecules to bypass genetic defects (**Figure 3.6B**). **Figure 3.6B** demonstrates the ability of two molecules to bypass the defect in development created by the absence of *ramC*. This suggests these two molecules act on processes prior to aerial hyphae formation; for example, perhaps acting as surfactants to restore aerial hyphae formation. The developmental promoter-*lux* fusion library can also be introduced into these mutant strains to assess the level of complementation. For example, *ramC*p-*lux* can be used to determine the expression of the surfactant SapB and *ftsZ*p-*lux* can be used to detect cell division in the aerial hyphae; while *whiE*p1-*lux* would assess the production of the grey pigment. Titrations of these molecules should also be performed to identify the most effective working concentration for each. As well, global transcriptional analysis can be used to determine the effects of these small molecules outside the known regulatory cascade. Finally, addition of these molecules to other streptomycetes can be used to determine if the effect on sporulation is conserved amongst the streptomycetes (**Figure 3.6C**).

3.3.5 Changes in pigment production

Decreased Pigment Production

Two hundred and forty small molecules were capable of interfering with *S. coelicolor*'s pigment production. These small molecules could be divided into those reducing both visible pigments (actinorhodin and prodiginine), only

actinorhodin, or only prodiginine (**Figure 3.7**). Examples shown in **Figure 3.7** are those molecules that did not have a significant effect on sporulation; however, these molecules can be further divided by their effect on development as some had strong defects in the formation of aerial hyphae and the development of grey pigmentation. This class of molecules offers an opportunity to study small molecules that have a negative effect on secondary metabolism. Many regulators in *S. coelicolor* have been identified genetically by reduced pigment production when deleted (see **Table 1.3** Chapter 1). Chemically inactivating proteins involved in positive regulation of secondary metabolism would result in decreases in secondary metabolism, mimicking the effects of genetic deletions.

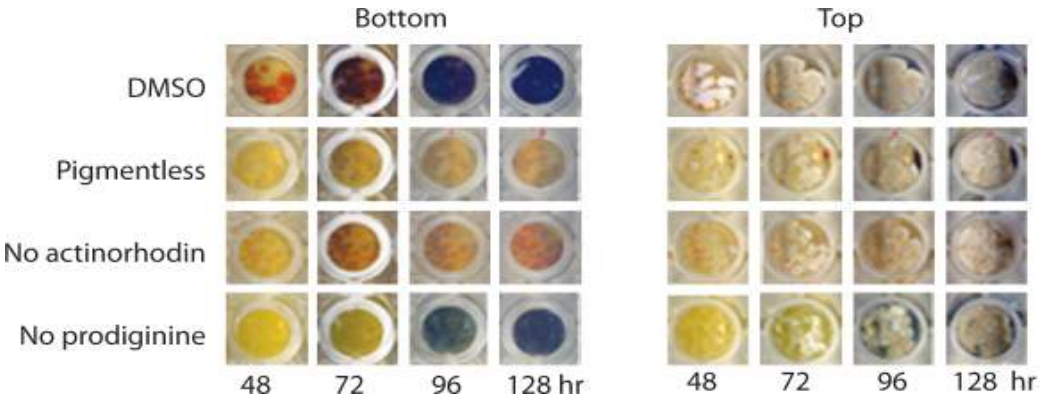

Figure 3.7. Examples of small molecules which reduced pigment production in *S. coelicolor*. Compounds were used at a concentration of 10 μM and compared daily to the DMSO control plated on R5M. Pigment production was assessed from the bottom view. Prodiginine production is denoted by red pigmentation in the colony; while actinorhodin production is denoted by blue pigmentation. Development was assessed from the top view. White colonies denote aerial hyphae formation and grey signifies sporulation. The examples depicted are from the bioactive subset of the CCC (N,N-hexamethyleneamiloride addition resulted in pigmentless colonies, reduced prodiginine production resulted from acriflavinium and reduced actinorhodin production is from 3,7-dimethoxyflavone addition).

Compounds which reduce pigment production may be acting in two general ways: altering the regulatory network which prevents expression of the metabolites or inhibiting key steps during biosynthesis of the metabolites, preventing accumulation. To distinguish between these two possibilities, strains over-expressing the pathway specific activator for actinorhodin (ActII-4) or prodiginine (RedD) can be used, as these would bypass the natural regulation within the cell (57); thus, small molecules still able to interfere with pigment production would most likely alter biosynthesis. Those molecules no longer able to interfere with pigment production likely function by altering the regulatory network leading to the expression of the metabolites, as their effects have been bypassed. Once grouped into regulatory inhibitors and biosynthetic inhibitors, follow up experiments could include transcriptional analyses to identify the pathways affected by the regulatory inhibitors, and over-expressing the biosynthetic genes to determine the targets of the biosynthetic inhibitors. Determining the processes affected by these molecules will lead to a greater understanding of secondary metabolism regulation and biosynthesis of *S. coelicolor*`s pigmented antibiotics.

Increased Pigments

Small molecules which improved pigment production were mainly identified by increased or early production of actinorhodin. The 112 compounds from the primary screen (synthetic and bioactive subsets of the CCC) were compared for their ability to stimulate actinorhodin production. Molecules with strong antimicrobial activity and those with low actinorhodin stimulating activity were removed; this left 19 small molecules capable of strongly activating actinorhodin production (named ARC, for antibiotic remodeling compounds **Figure 3.8D**). Interestingly, these molecules had varying effects on prodiginine production and were grouped by their effects on yields of this red pigment (6 molecules increased prodiginine production, 6 decreased prodiginine production and 7 had no effect **Figure 3.8A-C**). While all 19 molecules warrant further analysis, of particular interest were: ARC2-5 as they are a set of related structures; ARC6 as it is the most potent actinorhodin stimulating molecule; ARC11 as it is the MreB inhibitor A22 (its addition results in strong prodiginine production); and ARC15/18 as they are a second set of structurally related molecules.

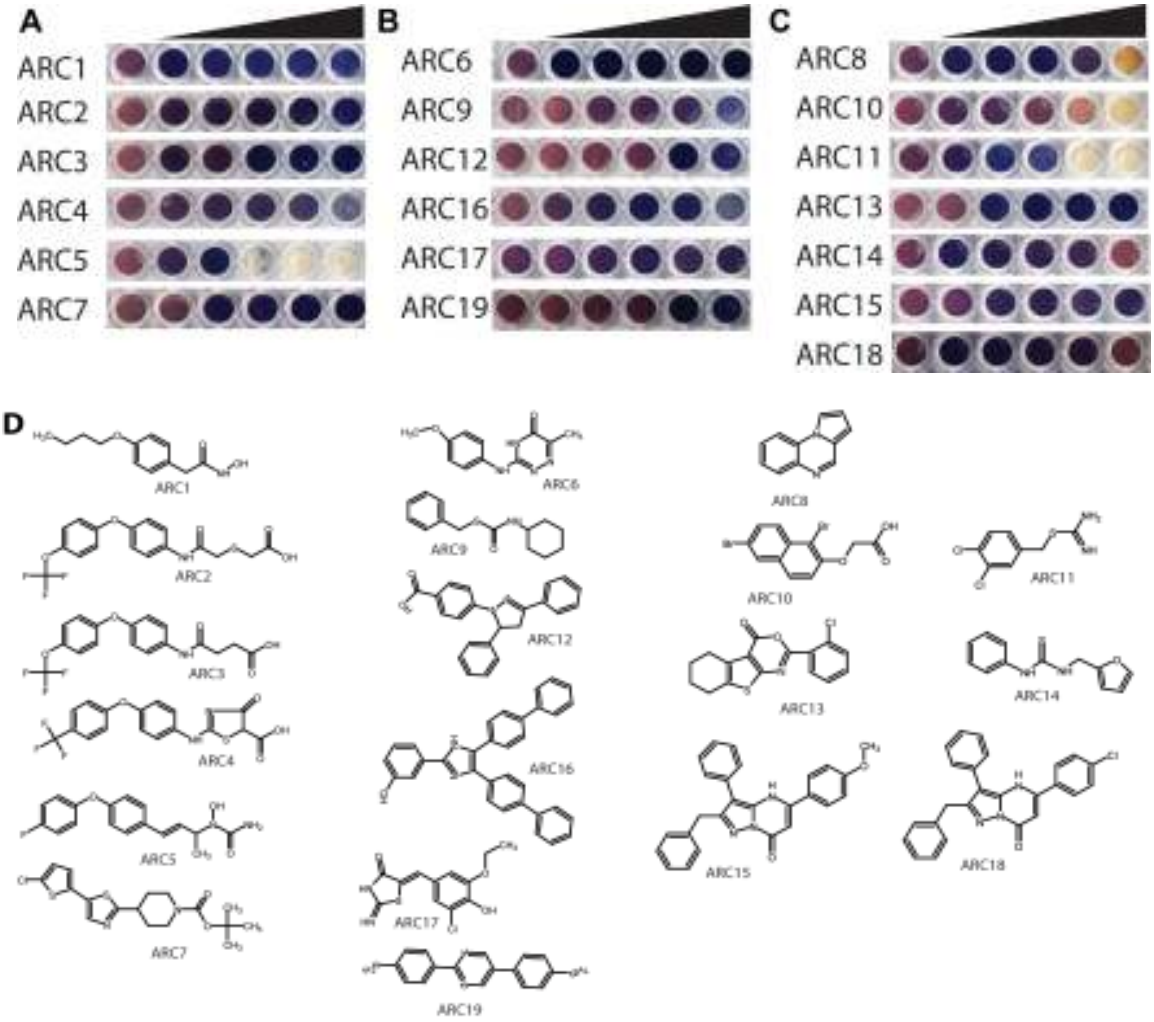

Figure 3.8. 19 actinorhodin stimulating molecules. The ARCs (antibiotic remodeling compounds) were assessed for their ability to induce actinorhodin production on R5M media. Actinorhodin production was determined by the development of a blue pigment. Compounds were added as a titration from 0, 7.8, 15.6, 31.3, 62.5, 125, 250 μM in the presence of *S. coelicolor* J1501 and plates scanned after 5 days growth. (**A**) ARCs that reduces yield of prodiginines (**B**) ARCs that increase prodiginine yields (**C**) ARCs that did not effect prodiginine production (**D**) Structures of the 19 ARCs.

Determining the mechanism by which these molecules improve secondary metabolism is of prime importance. Transcriptional analysis (microarray or RNA sequencing) in the presence and absence of these ARC molecules can be used to determine any resulting changes in gene expression and provide insights into the pathways leading to increased actinorhodin production. As well, the ability of the ARCs to exert their effect can be assessed in deletion and over-expression strains of genes known to be involved in regulating secondary metabolism. The inability of an ARC to exert its effect on one of these strains would suggest a target pathway or protein. For target identification, versions of the ARCs can be created that contain affinity tags allowing for covalent binding of the molecule to its protein target and provide a tag for purification of the protein-small molecule pair. Once purified, the identity of the protein target can be determined by mass spectrometry and confirmed by deletion and over-expression of the putative target.

Once the pathways and targets for the 19 ARC molecules have been elucidated, their application to improving secondary metabolism in other streptomycetes can be evaluated. Only small molecules that target well conserved pathways would warrant application to these other strains. Identifying small molecule effectors of secondary metabolism will greatly benefit drug discovery in the streptomycetes as they can be easily added to multiple strains to alter secondary metabolite

production, providing a non-invasive high throughput method to alter producer strains.

3.4 Conclusion

Genetic strategies have been used to alter the visual aspects of the *S. coelicolor* lifecycle and this has been vital to understanding growth, development and secondary metabolism within the streptomycetes. Through our screen for chemical modulators of these visual aspects of *S. coelicolor* colonies, we have recapitulated the phenotypes identified by genetic manipulation. In the future, characterizing the effects of these small molecules will lead to a greater understanding of *Streptomyces* biology.

3.5 Material and Methods

3.5.1 High throughput screen for small molecules altering *S. coelicolor* growth. *S. coelicolor* M145 spores were dispensed at 20 cfu/well into 96 well screening plates containing 200 µl R5M agar using a Biomex FX liquid handler (Beckman Coulter). Two microlitres of compounds were dispensed in duplicate from 1 mM master plates into the screening plate to give a final screening concentration of 10 µM of the Canadian Compound Collection of small molecules (McMaster HTS facility). Outer columns were reserved for control molecules with DMSO used as normal growth control and thiostrepton (10 µM) used as a growth inhibition control. Plates were shaken to distribute the compounds evenly and then dried in a laminar flow hood for 2 hours. Phenotypes were recorded over 10 days growth at 30^{o}C.

One challenge was establishing growth conditions for screening in 96 well plates. *S. coelicolor* exhibits its full range of developmental phenotypes on solid media, not in liquid media and the appearance of individual colonies responds to humidity, density of inoculum, time, medium constituents, pH and other environmental variables. It was necessary therefore to establish growth conditions where we could observe reproducible colony behaviour in >60,000 wells on >700 96-well plates over 4 months of screening. We therefore prepared a batch of 1 x 10^{15} spores of the strain *S. coelicolor* M145 at a concentration of 2 x 10^{13} cfu/ml. These spores were stored in frozen aliquots and diluted by 1 x 10^{-9}

prior to inoculation so that each well received ~20 cfu. R2YE and MS agar were not suitable for screening because colony morphology was not sufficiently reproducible in this format. We found R5M medium gave robust growth, development and was compatible with the solvent in which the compound library was dissolved in dimethyl sulfoxide (DMSO). Concentrations of DMSO greater than 3% impaired aspects of the *S. coelicolor* lifecycle and given the storage concentration of the compound library this limited the small molecule concentration for screening to 10 μM and 1% DMSO.

Our source of small molecules was the Canadian Compound Collection, supplemented by the 1,000 CDK library of kinase inhibitors. The Canadian Compound Collection library consists of 25,989 synthetic molecules and 3,580 natural products and known bioactives, sourced from Chembridge, Maybridge, Prestwick, BIOMOL, Sigma and Microsource.

We screened 1000-3000 molecules, in duplicate, at a time. Plates were prepared and monitored daily over 10 days of growth at 30^{o}C. Changes in colony appearance were recorded as a function of time and well. Results are summarized in **Table 3.1**.

3.5.2. Compounds for follow up analysis of primary hits. With the exception of the white molecules and increased pigment molecules, only the bioactive subset of the CCC was subjected to follow up analysis. Compounds for initial

follow up were obtained from the McMaster HTS facility. Follow up assays were performed on R5M at 10 μM unless otherwise stated.

3.5.3 Initial confirmation of small molecules blocking aerial hyphae formation. Twenty eight compounds from the bioactive subset were applied at 10 μM to *S. coelicolor* M145. Growth, aerial hyphae production and spore formation were monitored visually over 7 days.

3.5.4. Analysis of small molecules interfering with sporulation. One hundred and ninety six small molecules (CCC synthetic and bioactive) were applied at 10 μM to *S. coelicolor* M145 strains harbouring luminescent reporters (empty vector, *hrdB*p, *ramC*p and *whiE*p1). Luminescence was read over 120 hrs of growth at 30^{o}C using the Victor Multilabel Plate Reader.

3.5.5. Analysis of small molecules causing precocious sporulation. The 48 compounds from the bioactive subset were applied at 10 μM to *S. coelicolor* M145 harbouring the luminescent reporter monitoring *whiE*p1 expression. Luminescence was read over 108 hr of growth at 30^{o}C using the Victor Multiplabel Plate Reader. These molecules were also applied to *S. lividans* and *S. coelicolor* Δ*ramC* and monitored visually for aerial hyphae production and sporulation.

3.5.6. Small molecules resulting in decreased pigment production. The 35 compounds from the bioactive subset were applied to *S. coelicolor* M145 at 10

μM and monitored visually for alterations in the production of the red pigment (prodiginine) and blue pigment (actinorhodin) as well as sporulation.

3.5.7 Confirmation of small molecules which increased pigment production.

The 112 molecules which increased pigment production (both CCC synthetic and bioactive) were applied to *S. coelicolor* M145 at 10 μM and monitored for increased pigment production. Thirty compounds were purchased for further follow up (Sigma, Maybridge, Chembridge). This was narrowed down to 19 small molecules by removing those molecules with strong antimicrobial activity. These 19 molecules were visually assessed in *S. coelicolor* J1501 as a titration (0, 7.8, 15.6, 31.3, 62.5, 125, 250 μM).

3.6 References

1. Flardh K (2003) Essential role of DivIVA in polar growth and morphogenesis in *Streptomyces coelicolor* A3(2). *Mol Microbiol* 49: 1523-1536.

2. McCormick JR & Flardh K (2012) Signals and regulators that govern *Streptomyces* development. *FEMS Microbiol Rev* 36: 206-231.

3. van Wezel GP & McDowall KJ (2011) The regulation of the secondary metabolism of *Streptomyces*: New links and experimental advances. *Nat Prod Rep* 28: 1311-1333.

4. Kelemen GH & Buttner MJ (1998) Initiation of aerial mycelium formation in *Streptomyces*. *Curr Opin Microbiol* 1: 656-662.

5. Pope MK, Green BD & Westpheling J (1996) The *bld* mutants of *Streptomyces coelicolor* are defective in the regulation of carbon utilization, morphogenesis and cell--cell signalling. *Mol Microbiol* 19: 747-756.

6. Capstick DS, Willey JM, Buttner MJ & Elliot MA (2007) SapB and the chaplins: Connections between morphogenetic proteins in *Streptomyces coelicolor. Mol Microbiol* 64: 602-613.

7. Keijser BJ, van Wezel GP, Canters GW & Vijgenboom E (2002) Developmental regulation of the *Streptomyces lividans* ram genes: Involvement of RamR in regulation of the *ramCSAB* operon. *J Bacteriol* 184: 4420-4429.

8. Kodani S*, et al* (2004) The SapB morphogen is a lantibiotic-like peptide derived from the product of the developmental gene *ramS* in *Streptomyces coelicolor. Proc Natl Acad Sci U S A* 101: 11448-11453.

9. O'Connor TJ, Kanellis P & Nodwell JR (2002) The *ramC* gene is required for morphogenesis in *Streptomyces coelicolor* and expressed in a cell type-specific manner under the direct control of RamR. *Mol Microbiol* 45: 45-57.

10. Willey JM, Willems A, Kodani S & Nodwell JR (2006) Morphogenetic surfactants and their role in the formation of aerial hyphae in *Streptomyces coelicolor. Mol Microbiol* 59: 731-742.

11. Elliot MA*, et al* (2003) The chaplins: A family of hydrophobic cell-surface proteins involved in aerial mycelium formation in *Streptomyces coelicolor. Genes Dev* 17: 1727-1740.

12. Kelemen GH*, et al* (1998) Developmental regulation of transcription of *whiE*, a locus specifying the polyketide spore pigment in *Streptomyces coelicolor* A3 (2). *J Bacteriol* 180: 2515-2521.

13. Chater KF (2001) Regulation of sporulation in *Streptomyces coelicolor* A3(2): A checkpoint multiplex?. *Curr Opin Microbiol* 4: 667-673.

14. Flardh K, Findlay KC & Chater KF (1999) Association of early sporulation genes with suggested developmental decision points in *Streptomyces coelicolor* A3(2). *Microbiology* 145 (Pt 9): 2229-2243.

15. Kelemen GH*, et al* (1996) The positions of the sigma-factor genes, *whiG* and *sigF*, in the hierarchy controlling the development of spore chains in the aerial hyphae of *Streptomyces coelicolor* A3(2). *Mol Microbiol* 21: 593-603.

16. Flardh K, Leibovitz E, Buttner MJ & Chater KF (2000) Generation of a non-sporulating strain of *Streptomyces coelicolor* A3(2) by the manipulation of a developmentally controlled *ftsZ* promoter. *Mol Microbiol* 38: 737-749.

17. McCormick JR & Losick R (1996) Cell division gene *ftsQ* is required for efficient sporulation but not growth and viability in *Streptomyces coelicolor* A3(2). *J Bacteriol* 178: 5295-5301.

18. Fowler-Goldsworthy K*, et al* (2011) The actinobacteria-specific gene *wblA* controls major developmental transitions in *Streptomyces coelicolor* A3(2). *Microbiology* 157: 1312-1328.

19. Soliveri J, Brown KL, Buttner MJ & Chater KF (1992) Two promoters for the *whiB* sporulation gene of *Streptomyces coelicolor* A3(2) and their activities in relation to development. *J Bacteriol* 174: 6215-6220.

20. Smith LJ*, et al* (2010) *Mycobacterium tuberculosis* WhiB1 is an essential DNA-binding protein with a nitric oxide-sensitive iron-sulfur cluster. *Biochem J* 432: 417-427.

21. Lamb D, Kelly D & Kelly S (1999) Molecular aspects of azole antifungal action and resistance. *Drug Resist Updat* 2: 390-402.

22. McLean KJ*, et al* (2002) Azole antifungals are potent inhibitors of cytochrome P450 mono-oxygenases and bacterial growth in Mycobacteria and streptomycetes. *Microbiology* 148: 2937-2949.

23. McLean KJ*, et al* (2010) The *Mycobacterium tuberculosis* cytochromes P450: Physiology, biochemistry & molecular intervention. *Future Med Chem* 2: 1339-1353.

24. Alex L & Sharma G (1992) Tunicamycin resistance in *Saccharomyces cerevisiae*: Genetic studies. *FEMS Microbiol Lett* 70: 43-47.

25. Noda Y, Yoda K, Takatsuki A & Yamasaki M (1992) TmrB protein, responsible for tunicamycin resistance of *Bacillus subtilis*, is a novel ATP-binding membrane protein. *J Bacteriol* 174: 4302-4307.

26. Park YS, Shin WS & Kim SK (2008) In vitro and in vivo activities of echinomycin against clinical isolates of *Staphylococcus aureus. J Antimicrob Chemother* 61: 163-168.

27. Vecchione JJ, Alexander B,Jr & Sello JK (2009) Two distinct major facilitator superfamily drug efflux pumps mediate chloramphenicol resistance in *Streptomyces coelicolor. Antimicrob Agents Chemother* 53: 4673-4677.

28. Hesketh A*, et al* (2011) Genome-wide dynamics of a bacterial response to antibiotics that target the cell envelope*. BMC Genomics* 12: 226.

29. Hong HJ, Paget MS & Buttner MJ (2002) A signal transduction system in *Streptomyces coelicolor* that activates the expression of a putative cell wall glycan operon in response to vancomycin and other cell wall-specific antibiotics*. Mol Microbiol* 44: 1199-1211.

30. Rodriguez-Garcia A, Combes P, Perez-Redondo R, Smith MC & Smith MC (2005) Natural and synthetic tetracycline-inducible promoters for use in the antibiotic-producing bacteria *Streptomyces. Nucleic Acids Res* 33: e87.

31. Hara O & Hutchinson CR (1990) Cloning of midecamycin(MLS)-resistance genes from *Streptomyces mycarofaciens*, *Streptomyces lividans* and *Streptomyces coelicolor* A3(2)*. J Antibiot (Tokyo)* 43: 977-991.

32. Coats JH & Argoudelis AD (1971) Microbial transformation of antibiotics: Phosphorylation of clindamycin by *Streptomyces coelicolor* muller*. J Bacteriol* 108: 459-464.

33. Marshall VP*, et al* (1991) Purification of lincosaminide O-nucleotidyltransferase from *Streptomyces coelicolor* muller*. J Antibiot (Tokyo)* 44: 895-900.

34. Puzynina GG, Danilenko VN, Vasil'chenko LG, Mkrtumian NM & Lomovskaia ND (1979) Determination of *Streptomyces coelicolor* A3(2) resistance to erythromycin*. Genetika* 15: 1730-1738.

35. Wiley PF, Baczynskyj L, Dolak LA, Cialdella JI & Marshall VP (1987) Enzymatic phosphorylation of macrolide antibiotics*. J Antibiot (Tokyo)* 40: 195-201.

36. Ruble JC*, et al* (2007) Structure-activity relationships of bioisosteres of a carboxylic acid in a novel class of bacterial translation inhibitors*. Bioorg Med Chem Lett* 17: 4040-4043.

37. Ananthan S*, et al* (2009) High-throughput screening for inhibitors of *Mycobacterium tuberculosis* H37Rv*. Tuberculosis (Edinb)* 89: 334-353.

38. Xu HH, *et al* (2010) *Staphylococcus aureus* TargetArray: Comprehensive differential essential gene expression as a mechanistic tool to profile antibacterials. *Antimicrob Agents Chemother* 54: 3659-3670.

39. Barker CA, Farha MA & Brown ED (2010) Chemical genomic approaches to study model microbes. *Chem Biol* 17: 624-632.

40. Wang J, *et al* (2006) Platensimycin is a selective FabF inhibitor with potent antibiotic properties. *Nature* 441: 358-361.

41. Azad MA & Wright GD (2012) Determining the mode of action of bioactive compounds. *Bioorg Med Chem* 20: 1929-1939.

42. Maehara S, Simanjuntak P, Kitamura C, Ohashi K & Shibuya H (2011) Cinchona alkaloids are also produced by an endophytic filamentous fungus living in cinchona plant. *Chem Pharm Bull (Tokyo)* 59: 1073-1074.

43. Straight PD, Willey JM & Kolter R (2006) Interactions between *Streptomyces coelicolor* and *Bacillus subtilis*: Role of surfactants in raising aerial structures. *J Bacteriol* 188: 4918-4925.

44. Claessen D, de Jong W, Dijkhuizen L & Wosten HA (2006) Regulation of *Streptomyces* development: Reach for the sky!. *Trends Microbiol* 14: 313-319.

45. Willey J, Schwedock J & Losick R (1993) Multiple extracellular signals govern the production of a morphogenetic protein involved in aerial mycelium formation by *Streptomyces coelicolor*. *Genes Dev* 7: 895-903.

46. Bibb MJ, Molle V & Buttner MJ (2000) Sigma(BldN), an extracytoplasmic function RNA polymerase sigma factor required for aerial mycelium formation in *Streptomyces coelicolor* A3(2). *J Bacteriol* 182: 4606-4616.

47. Elliot MA, Bibb MJ, Buttner MJ & Leskiw BK (2001) BldD is a direct regulator of key developmental genes in *Streptomyces coelicolor* A3(2). *Mol Microbiol* 40: 257-269.

48. Kelemen GH, *et al* (1998) Developmental regulation of transcription of *whiE*, a locus specifying the polyketide spore pigment in streptomyces coelicolor A3 (2). *J Bacteriol* 180: 2515-2521.

49. Shiina T, Tanaka K & Takahashi H (1991) Sequence of *hrdB*, an essential gene encoding sigma-like transcription factor of *Streptomyces coelicolor* A3(2): Homology to principal sigma factors. *Gene* 107: 145-148.

50. Willey J, Santamaria R, Guijarro J, Geistlich M & Losick R (1991) Extracellular complementation of a developmental mutation implicates a small sporulation protein in aerial mycelium formation by *S. coelicolor. Cell* 65: 641-650.

51. Nguyen KT*, et al* (2002) A central regulator of morphological differentiation in the multicellular bacterium *Streptomyces coelicolor. Mol Microbiol* 46: 1223-1238.

52. Ma S & Ma S (2012) The development of FtsZ inhibitors as potential antibacterial agents. *ChemMedChem* 7: 1161-1172.

53. Modell JW, Hopkins AC & Laub MT (2011) A DNA damage checkpoint in *Caulobacter crescentus* inhibits cell division through a direct interaction with FtsW. *Genes Dev* 25: 1328-1343.

54. Mo AH & Burkholder WF (2010) YneA, an SOS-induced inhibitor of cell division in *Bacillus subtilis*, is regulated posttranslationally and requires the transmembrane region for activity. *J Bacteriol* 192: 3159-3173.

55. Ogino H, Teramoto H, Inui M & Yukawa H (2008) DivS, a novel SOS-inducible cell-division suppressor in *Corynebacterium glutamicum. Mol Microbiol* 67: 597-608.

56. Cordell SC, Robinson EJ & Lowe J (2003) Crystal structure of the SOS cell division inhibitor SulA and in complex with FtsZ. *Proc Natl Acad Sci U S A* 100: 7889-7894.

57. McKenzie NL & Nodwell JR (2007) Phosphorylated AbsA2 negatively regulates antibiotic production in *Streptomyces coelicolor* through interactions with pathway-specific regulatory gene promoters. *J Bacteriol* 189: 5284-5292.

CHAPTER 4

THE ARC2 SERIES REMODELS SECONDARY METABOLISM THROUGH INHIBITION OF FATTY ACID BIOSYNTHESIS

Adapted from:

Arryn Craney, Cory Ozimok, Sheila Marie Pimentel-Elardo, Alfredo Capretta and Justin Nodwell. Chemical perturbation of secondary metabolism demonstrates important links to primary metabolism. *Chem. Bio*. 2012; 19(8):1020-1027.

4.1 Abstract

The mechanisms that limit the production of secondary metabolites in the laboratory are not well understood and this has impeded the discovery of many important compounds. We have identified a set of small molecules, the ARC2 series, that remodel the yields of secondary metabolites in many prokaryotes and do so by inhibiting fatty acid biosynthesis. This demonstrates a particularly intimate relationship between this primary metabolic pathway and secondary metabolism and suggests a new approach to enhancing the yields of metabolites for discovery and biochemical characterization.

4.2 Introduction

The secondary metabolites produced by bacteria exhibit potent biological activities and have been developed extensively for antimicrobial, anticancer and other vital therapeutic applications. Most of these molecules are polyketides and non-ribosomal peptides and are produced by biochemical pathways that are encoded in discrete islands within the genome. The genes encoded in these islands normally include all the metabolite-specific biochemical steps; however, certain precursor molecules and other resources are drawn from primary metabolism (1, 2).

An outstanding source of natural products and drug leads are the actinobacteria. One of these, *Streptomyces coelicolor*, is a superb model system for secondary metabolism due to its production of pigmented secondary metabolites; including the blue polyketide actinorhodin (3), and the red prodiginines, undecylprodiginine and streptrorubin B (4). These molecules can be readily detected visually and quantified spectrophotometrically.

Understanding the metabolic and regulatory mechanisms that determine the yields of secondary metabolites from producer organisms is important for the discovery and exploitation of these molecules. A major outcome of genome sequencing has been the discovery that each streptomycete chromosome encodes many more secondary metabolites than can be detected from routine laboratory culture (5, 6). Efforts to enhance output have focused on the

engineering of strains to overexpress pathway specific (7) and pleiotropic regulators (8) of secondary metabolism and on moving biosynthetic gene clusters into engineered overproduction strains (9, 10). These approaches have been successful but are technically cumbersome and not amenable to high throughput.

In this work, we show that one set of the actinorhodin stimulating molecules (ARCs) identified in Chapter 3, ARC2-5, target fatty acid biosynthesis - demonstrating that the relationship between primary and secondary metabolism is one limit on secondary metabolite yields.

4.3 Results

4.3.1 The ARC2 series remodels secondary metabolism in the actinomycetes

Four of the 19 molecules actinorhodin-stimulating molecules from our HTS screen, ARC2, 3, 4 and 5 (referred to as the 'ARC2 series' for antibiotic remodeling compounds, **Figure 4.1A**), had similar structures and stimulated blue pigmentation during growth on solid medium. Spectrophotometric analysis of liquid cultures demonstrated that the ARC2 series reduced yields of the prodiginines (**Figure 4.1B**). Using LC-MS analysis we found that ARC2 induced changes in the production of 2 other *S. coelicolor* secondary metabolites. Yields of the germidicins (11) (germicidin B/C 183.107 $[M+H]^+$ and germicidin A 197.125 $[M+H]^+$) were enhanced ~3 fold by ARC2 (**Figure 4.1C**) while yields of the daptomycin-like calcium-dependent antibiotic (CDA) (1495.485 $[M+H]^+$) (12, 13) were reduced ~2-fold (**Figure 4.1D**). The ARC2 series can therefore pleiotropically remodel secondary metabolism in *S. coelicolor*.

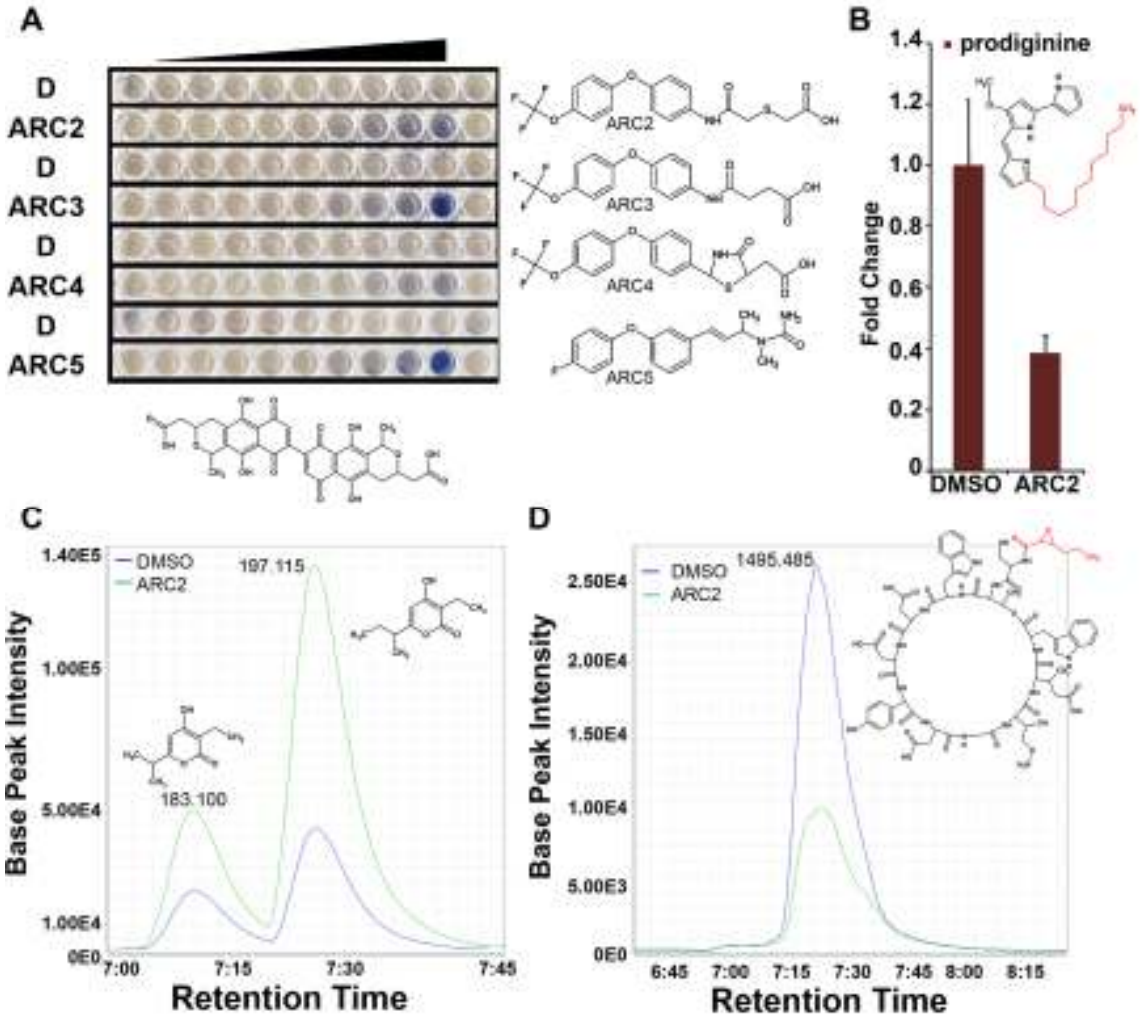

Figure 4.1. The ARC2 series of actinorhodin-inducing molecules. (A) The structures of the ARC2 series and their effect on blue pigmentation of *S. coelicolor* colonies. Molecules were added at 0.78, 1.56, 3.13, 6.25 12.5, 25, 50, 100 μM to solid medium. The DMSO solvent only controls are indicated (D) and in wells 1 and 10 of compound rows. The plate was photographed following three days growth. **(B)** Spectrophotometric analysis of prodiginine yields in *S. coelicolor* M145 grown in the presence of 25 μM ARC2 or with a solvent only control (DMSO) after 3 days growth in liquid R5M. **(C)** The effects of 10 μM ARC2 on germicidin production (germicidin B/C 183.107 $[M+H]^+$ and germicidin A 197.125 $[M+H]^+$) were assessed using LC-MS analysis after 7 days growth on solid R5M. **(D)** The production of calcium-dependent antibiotic (CDA) (1495.485 $[M+H]^+$ for CDA 4a) was assessed in the presence of DMSO and 10 μM ARC2 using LC-MS after 7 days growth on solid R5M. FabI-dependent fatty acid moieties of prodiginine (*B*) and calcium-dependent antibiotic (*D*) are shown in red.

These effects led us to investigate the action of ARC2 on secondary metabolism in other bacteria. We added ARC2 to cultures of *Kutzneria* sp. 744, *S. pristinaespiralis* ATCC 25486 and *Streptomyces peucetius* 27952 and observed their small molecule output using LC-MS (**Figure 4.2**). ARC2 enhanced yields of desferrioxamine B and E (14) (561.362 and 601.360 $[M+H]^+$ respectively) in *S. pristinaespiralis* (**Figure 4.2A**); doxorubicin (544.196 $[M+H]^+$), baumycin (674.258 $[M+H]^+$) and three unknown molecules (417.103 $[M+H]^+$, 433.097 $[M+H]^+$ and 615.356 $[M+H]^+$) in *S. peucetius* (15) (**Figure 4.2B**) and an unknown metabolite (252.175 $[M+H]^+$) in *Kutzneria* (**Figure 4.2C**). Yields of some secondary metabolites were unchanged: for example, ARC2 had no effect on yields of pristinamycin IIA (526.259 $[M+H]^+$) in *S. pristinaespiralis* (16). These data suggest that the molecular target and mechanism of ARC2 is conserved in these diverse actinomycetes.

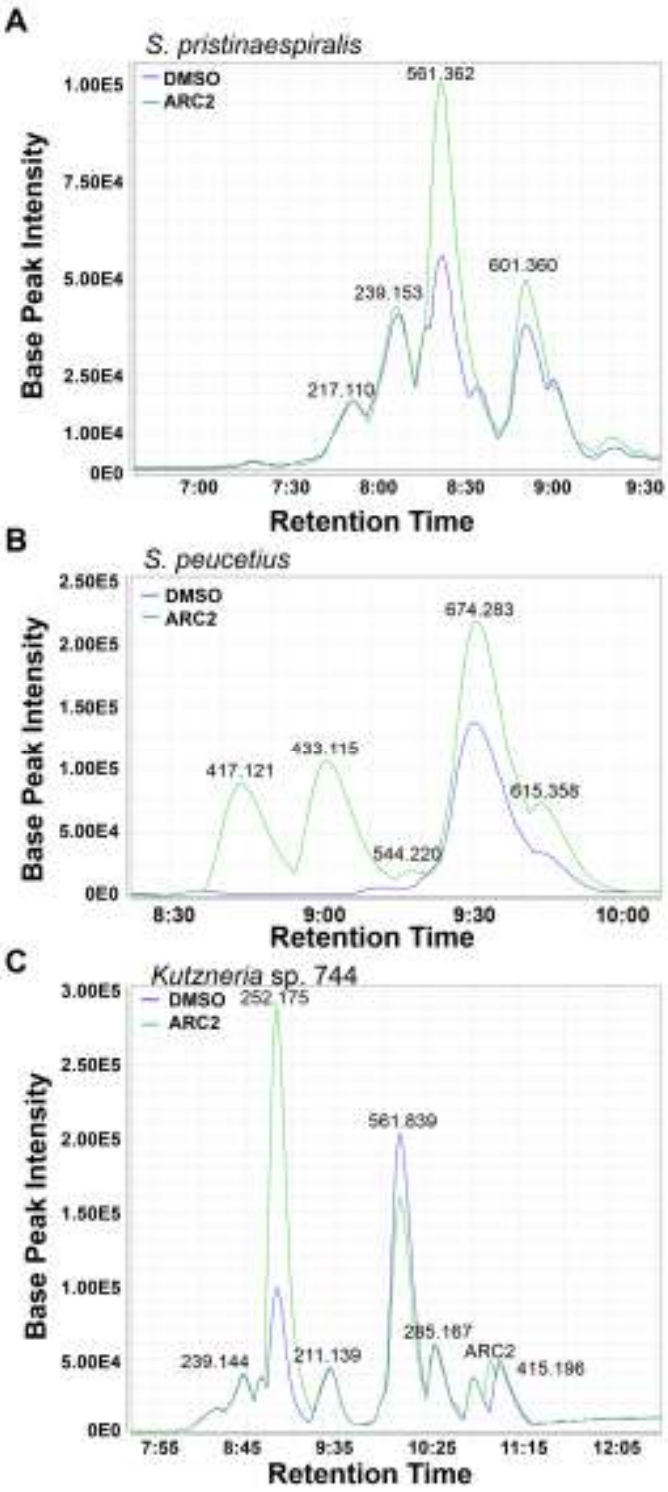

Figure 4.2. ARC2-enhanced yields of secondary metabolites. The effect of 10 µM ARC2 on secondary metabolite output was assessed in **(A)** *Streptomyces pristinaespiralis,* **(B)** *Streptomyces peucetius* and **(C)** *Kutzneria* using LC-MS analysis. All strains were grown for 7 days in the presence of either DMSO or 10 µM ARC2. Enhanced yields of desferrioxamine B and E (561.362 and 601.360 $[M+H]^+$ respectively) were observed in *Streptomyces pristinaespiralis* **(A)**, doxorubin and baumycin (544.220 $[M+H]^+$ and 674.283 $[M+H]^+$ respectively) and three unknown molecules (417.121, 433.115, 615.356 $[M+H]^+$ in *S. peucetius* **(B)** and an unknown metabolite (252.175 $[M+H]^+$) in *Kutzneria* **(C)**.

4.3.2 Structural determinants of ARC2 activity

To gain further insight into the structural features of the ARC2 series, we generated a structure-activity library by sourcing related commercial molecules (compounds 1 – 8) and by chemical synthesis (compounds 9 – 32) and exploring the capacity of these molecules to alter secondary metabolism in *S. coelicolor* (**Figure 4.3A**).

Changes to the core structure (bottom left panel in **Figure 4.3A**) such as the removal of ring A (compound 6) or replacement of the ether linkage between rings A and B with a ketone (compound 5) eliminated the enhancement of blue pigmentation. Insertion of an extra carbon between rings A and B similarly abolished activity (compound 7). These data suggested that the core shape of the molecules and rotational freedom around the ether linkage were important for biological activity.

The tails on ring B were the most variable feature (R_4 and R_5) (top right panel in **Figure 4.3A**); alterations introduced in the structure-activity analysis had similarly variable effects on biological activity. We investigated 13 substitutions at R_4 and R_5 in the context of a functional ring A (a Cl at R_1 – see below) and found that most variants exhibited relatively weak actinorhodin-stimulating activity or were inactive. One variant that proved functional was an acetamide group at R_4 (**Figure 4.3A**, compound 16 top left panel).

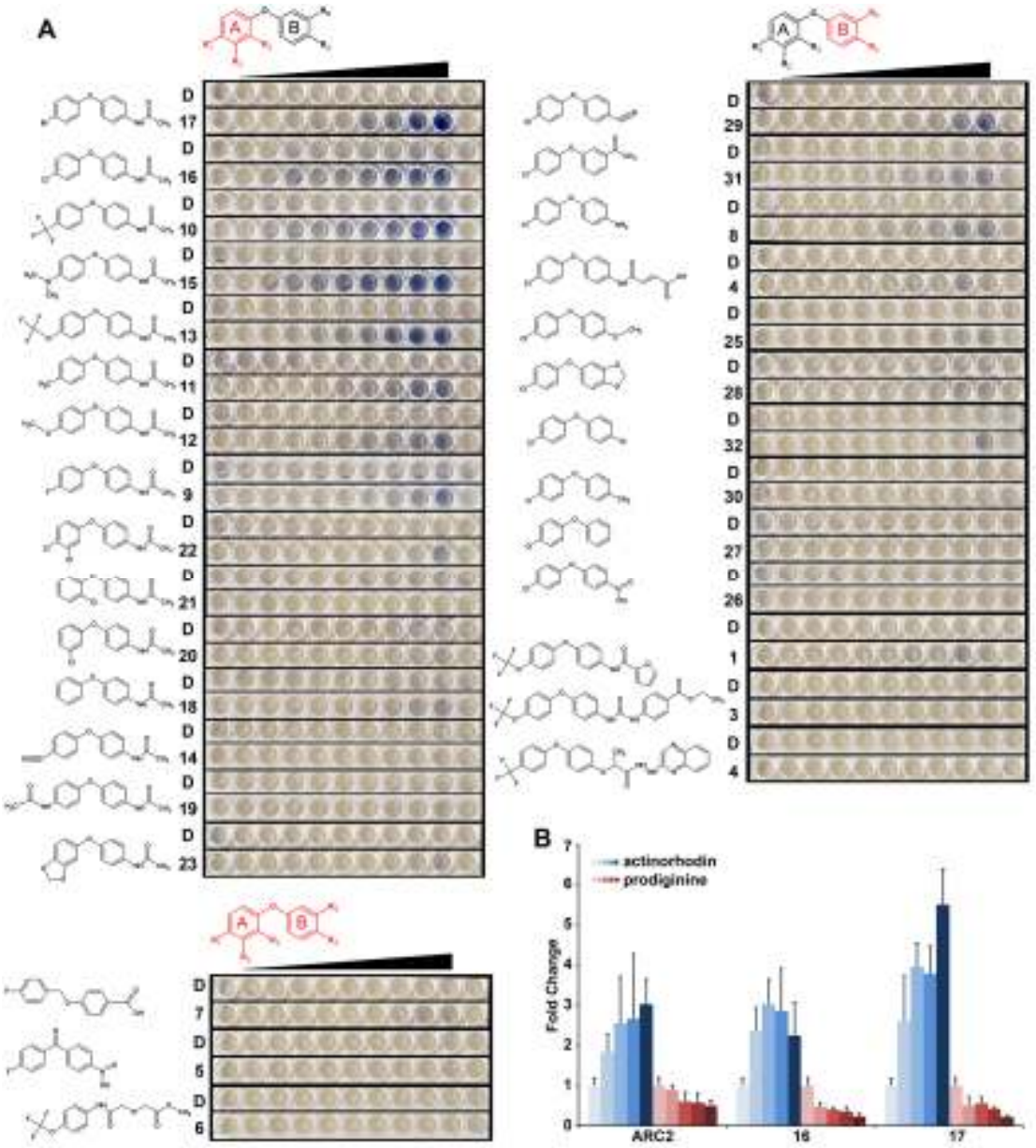

Figure 4.3. Effects of ARC2 and ARC2 analogues on actinorhodin and prodiginine biosynthesis. (A) Compounds 1 – 32 based on the ARC2 series were plated as a titration (0.78, 1.56, 3.13, 6.25 12.5, 25, 50, 100 μM) on solid MS agar in the presence of *S. coelicolor* M145 spores. The DMSO solvent controls are indicated (D) and wells 1 and 10 of compound rows. **(B)** Titrations of increasing concentration of ARC2 and compounds 16 and 17 (1.56, 6.25, 12.5, 25 μM) were assessed for effects on actinorhodin and prodiginine production and compared to a DMSO. Actinorhodin production is shown in blue and prodiginine in red.

We similarly generated compounds having varied substitutions on ring A in the context of an acetamide substitution at position R_4 of ring B (top left panel in **Figure 4.3A**) – variation on this ring had a greater influence on function. For example, we found that the absence of an electron-withdrawing group at R_1 compromised biological activity (compound 18). The electron-withdrawing group could not be moved to R_2 or R_3 without reducing or eliminating activity (compounds 20, 21 and 22). The nature of this electron-withdrawing group was also important. Of halogen substitutions, Br>Cl>F (compare compounds 9, 10, 16, and 17 in **Figure 4.3A**). While halogens were most active, other electron withdrawing substitutions at R_1 were also tolerated including $N(CH_3)_2$ and OCH_3 (compare compounds 11,12 and 15 in top left panel of **Figure 4.3A**). The only substitution that did not agree with this rule was CH_3 which proved functional (compound 11).

This structure-activity analysis generated several molecules that were active at lower concentrations than ARC2 (compounds 10, 13, 15, 16 and 17; see left panel of **Figure 4.3A**) and confirmed that the important features for activity were two benzene rings joined by an ether linkage with, in most cases, an electron-withdrawing group at the R_1 position.

As was the case with the initial hits of the ARC2 series, we observed a correlation between enhanced yields of γ-actinorhodin and reduced yields of the prodiginines during liquid culture (**Figure 4.3B**). The conserved structural features of the

active molecules and related phenotypic effects are consistent with a shared molecular mechanism of action for the ARC2 series.

4.3.3 The ARC2 series target fatty acid biosynthesis

Structural analysis of the ARC2 series suggested a possible molecular target for the series. Several of the most active molecules (e.g. compounds 16 and 17) were structurally similar to the known antimicrobial agent triclosan, an inhibitor of fatty acid synthesis (**Figure 4.4A**). Triclosan is a covalent inhibitor of the FabI enoyl reductase which catalyzes the final and rate-limiting step in fatty acid biosynthesis (17). We were intrigued by this similarity because fatty acid biosynthesis shares the precursors acetyl-CoA and malonyl-CoA with polyketide biosynthesis. The ARC2 series might therefore shunt resources from primary metabolism thereby enhancing the yields of some polyketides. Indeed, of the five metabolites that we were able to identify that were enhanced in yield by ARC2, four were polyketides that depend on the cellular acyl-CoA pool: actinorhodin, doxorubicin, baumycin and germicidin. Desferrioxamine belongs to the nonpeptide hydroxymate siderophores and at least one step in its biosynthesis is dependent on acetyl-CoA as well (18). Furthermore, the two metabolites that were reduced in yields by ARC2, the prodiginines and calcium-dependent antibiotic, both contain fatty acid moieties (highlighted with red on the structures in **Figure 4.1B** and **4.1D**) that are generated via primary metabolism and require the enoyl reductase activity of FabI (13, 19). FabI exhibits significant end-to-end

sequence similarity and is highly conserved in both its catalytic center and triclosan-binding residues (20, 21). This could be consistent with the broad spectrum of these molecules to influence secondary metabolism in diverse organisms. We have addressed this hypothesis in several ways.

First, we explored triclosan itself, finding that it exhibited a potent antimicrobial effect on *S. coelicolor* growth in the 1 to 50 μM range where the ARC2 series influenced pigmentation (the minimum inhibitory concentration was 3.13 μM). However, when we reduced the triclosan concentration below the MIC, we observed a clear stimulation of blue pigmentation (**Figure 4.4B**). This led us to synthesize 3 additional compounds containing the 2-phenoxyphenol core of triclosan (compounds 33, 34 and 35) (**Figure 4.4A**), each of which has been shown previously to inhibit FabI (22). Compound 33 inhibited *S. coelicolor* growth with an MIC of 25 μM, compounds 34 and 35 were less potent antimicrobials. Importantly, all three triclosan/ARC2 chimeras stimulated blue pigmentation of *S. coelicolor* colonies at concentrations that were sub-inhibitory for growth (**Figure 4.4B**). The ability of triclosan and these known FabI inhibitors to mimic the effects of ARC2 is consistent with a shared mechanism involving the inhibition of FabI.

Table 4.1. Antibiotic Resistance of ErmE and FabV in M145

Antibiotic	ErmE*(μM)	FabV(μM)
Triclosan	3.13	100
Kanamycin	6.25	6.25
Thiostrepton	0.78	0.78
Rifamycin	12.5	12.5
Tetracycline	50	50
Novobiocin	12.5	12.5

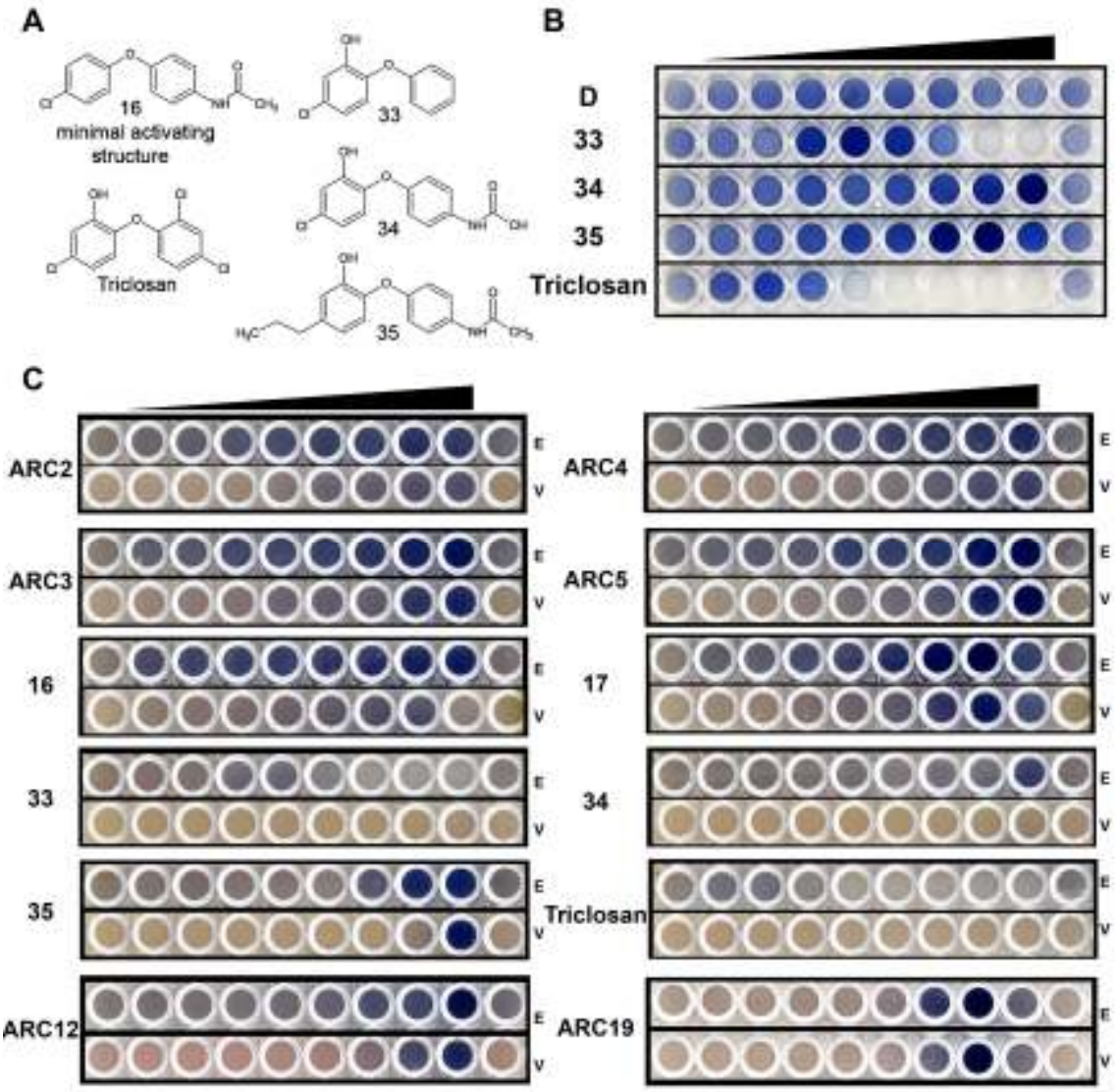

Figure 4.4. The ARC2 series acts through partial inhibition of fatty acid biosynthesis. (A) Comparison of the minimal activating structure (compound 16) with triclosan and three triclosan/ARC2 series chimeras (compounds 33-35). **(B)** The effects of these compounds on actinorhodin yields on solid MS agar. **(C)** ARC2-5, compounds 16 and 17, the triclosan/ARC2 chimeras (compounds 33, 34 and 35) and triclosan were added to cultures of *S. coelicolor* expressing the triclosan resistance allele *fabV* (rows denoted V) or a control vector (E). Similar analysis was applied to ARC12 and ARC19 which are unrelated to the putative or known FabI inhibitors. The effects of these compounds on actinorhodin yields were assessed at 0.78, 1.56, 3.13, 6.25 12.5, 25, 50, 100 µM. Solvent controls are indicated (D) and in wells 1 and 10 of compound rows. Actinorhodin yield was assessed by blue pigmentation after 7 days growth on solid MS agar.

To test this hypothesis genetically, we created a *S. coelicolor* strain that expressed the *fabV* gene from *Pseudomonas aeruginosa* which encodes a triclosan-resistant FabI paralogue (23). The expression of this gene increased the MIC of triclosan against *S. coelicolor* from 3.13 μM to 100 μM but had no effect on the MICs of antibiotics that target translation or DNA replication (**Table 4.1**). We compared the effects of ARC2, ARC3, ARC4, ARC5 and compounds 16, 17, 33, 34 and 35 on actinorhodin yields in the resistant strain (V) and an isogenic control (E) (**Figure 4.4C**). Consistent with the targeting of fatty acid biosynthesis by these molecules, we found that the expression of FabV impaired the effects of all of the molecules on blue pigmentation. We observed residual actinorhodin-stimulating activity at the highest concentrations of some compounds (50 and 100 μM); however, there was a clear bypass effect of *fabV* at lower concentrations. We then tested two other unrelated actinorhodin-inducing molecules from our primary screen, ARC12 and 19. Neither of these is structurally related to ARC2. The capacity of these molecules to stimulate actinorhodin yields was not influenced by the *fabV* gene (**Figure 4.4C**). We also explored the effect of *fabV* on the reduction of prodiginine yields by the ARC2 series (**Figure 4.5**) and found that indeed, this triclosan-resistant enoyl reductase restored undecylprodiginine (394.287 $[M+H]^+$) yields in the presence of ARC2. These data further support a mechanism of action for the ARC2 series that involves the inhibition of FabI.

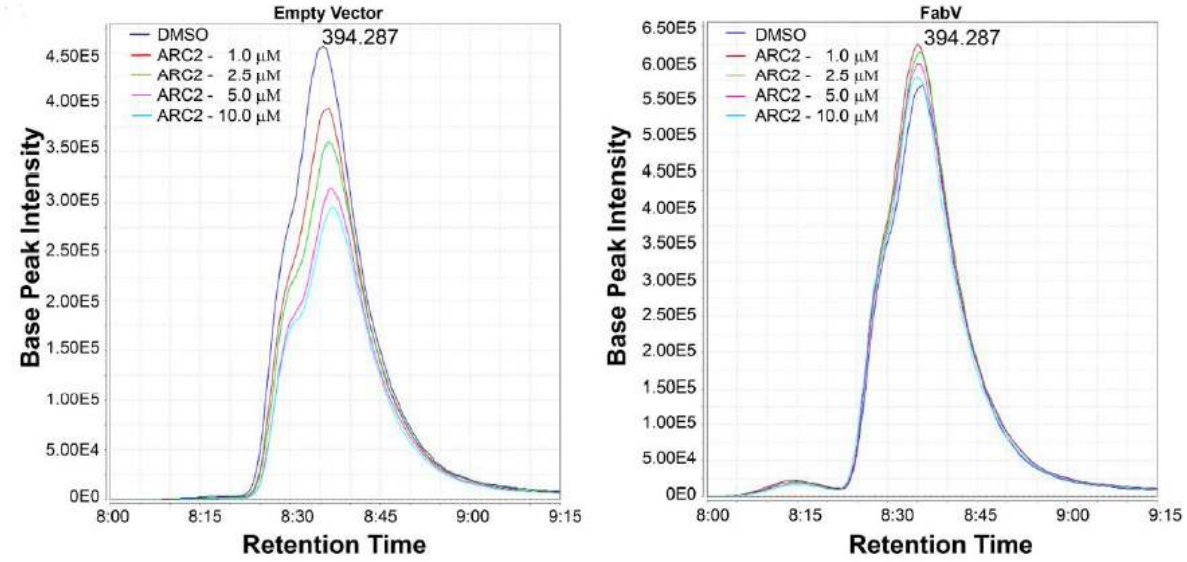

Figure 4.5. Complementation of the decreased production of prodiginine by ARC2 through FabV expression. The effect of 1.0, 2.5, 5.0 and 10.0 µM ARC2 on prodiginine production in a *fabV* expressing or control vector strain of *S. coelicolor* in comparison to DMSO. LC-MS analysis was used to detect prodiginine production (394.286 [M+H]) after 7 days growth on MS agar.

Finally, we assessed the ability of the ARC2 series to interfere with the biochemical activity of purified FabI *in vitro.* We established an enzyme assay using *E. coli* FabI where the consumption of NADH during the reduction of the substrate crotonyl-CoA was measured (24). As expected, we observed 100% inhibition of FabI with 0.01 µM triclosan. Compound 33 inhibited FabI activity to completion at 1 µM; compounds 34 and 35 inhibited ~80% of the activity at 10 µM, consistent with their less potent antimicrobial activities. Confirming a capacity to impair FabI activity, 50 µM ARC2 and compound 17 were able to inhibit FabI activity by 55% to 56%, respectively (**Table 4.2**). This work confirmed that the

ARC2 series inhibit the FabI enoyl reductase and that inhibition of this step is important for their effects on secondary metabolism.

Table 4.2. FabI Inhibition

Compound	Concentration (μM)	% Inhibition
Triclosan	0.005	86
Triclosan	0.01	100
33	1	100
34	10	75
35	10	81
ARC2	50	55
17	50	56

4.4 Discussion

We show that the ARC2 series is related to the antibiotic triclosan structurally and through the capacity to interfere with FabI activity *in vitro* and *in vivo.* The fact that expression of a triclosan resistant enoyl reductase, FabV, confers triclosan resistance and interferes with the effect of these molecules on actinorhodin and prodiginine biosynthesis demonstrates that impairing fatty acid biosynthesis remodels secondary metabolite yields. The reduction in *S. coelicolor*'s two lipid-containing secondary metabolites, prodiginine and calcium-dependent antibiotic, also correlates well with inhibition of fatty acid biosynthesis as there would be less lipids available for incorporation into these two metabolites.

The relationship between polyketide and fatty acid biosynthesis appears to be a very intimate one. For example most polyketide biosynthetic clusters, including those of actinorhodin, doxorubicin, baumycin and germicidin do not encode

designated malonyl-CoA transferases for loading a starter unit onto the polyketide acyl carrier protein. Instead, they recruit the malonyl-CoA:ACP transacylase enzyme from the fatty acid biosynthetic machinery for this purpose (25). The simplest interpretation of our data therefore, is that there is a balance between fatty acid and polyketide biosynthesis in cells that helps set the upper limit on polyketide yields. This balance could be manifested at the level of precursor supply, malonyl-CoA:ACP transacylase availability or both. Inhibiting this primary metabolic pathway therefore favours the polyketides and certain other metabolites having shared precursors (*e.g.* desferrioxamine). Polyketide biosynthesis is related to fatty acid biosynthesis biochemically; however, most polyketide biosynthetic pathways do not possess or require a designated enoyl reductase. This includes the pathways for the ARC2-stimulated polyketides that we have identified: actinorhodin, doxorubicin, baumycin and germicidin.

Partial inhibition, rather than complete inhibition of fatty acid biosynthesis is a critical feature of the most useful molecules we have identified. Fatty acid biosynthesis is essential for viability as demonstrated by the antimicrobial effect of triclosan and at least two of the ARC2 analogues. Indeed, the most potent inhibitors where characterized by a narrow concentration range, just below their MIC for growth (*e.g.* 0.78 μM to 3.13 μM for triclosan and 3.13 μM to 12.5 μM for compound 33), at which they influenced secondary metabolism. In contrast, the weaker inhibitors exhibited a broader concentration range at which they could

influence secondary metabolism and this range was not limited by a strong antimicrobial effect at higher concentrations. For example, for 16 and 17 we observed a stimulation of actinorhodin biosynthesis over the two orders of magnitude between 0.78 μM and 100 μM.

Many strategies have been applied to the improvement in yields of polyketides and other secondary metabolites including the over-expression of pathway specific or pleiotropic regulators (2). There have also been efforts to genetically enhance the availability of precursors (1). While all of these approaches have met with success, they require the genetic engineering of starting strains. This is difficult to accomplish in high throughput and is often prevented by the existence of restriction barriers to genetic engineering. In contrast, a strategy for secondary metabolite screening that depends on small molecule perturbation is, in principle, readily scalable to large numbers of strains using existing high throughput technologies. We note that while we observed relatively modest 2-5 fold enhancement in the yields of actinorhodin, which is normally produced during routine laboratory culture, the effect of ARC2 on some other metabolites was more dramatic. Indeed, we observed the induction of two unknown compounds in *Streptomyces peucetius* (417.103 m/z and 433.097 m/z) that could not be detected in the absence of ARC2. This suggests that this strategy can be applied to strains in the absence of genome sequence information and that this could facilitate the discovery of new metabolites.

4.5 Material and Methods

4.5.1 Bacterial strains, plasmids and culture conditions. Plasmids propagation and cloning were carried out in the *Escherichia coli* strain DH5α. *E. coli* strains were grown at 37^{o}C in Luria broth media. Antibiotic concentrations were 100 μg/ml of ampicillin, 50 μg/ml of kanamycin, 34 μg/ml of chloramphenicol, 25 μg/ml naladixic acid and 50 μg/ml of apramycin. All antibiotics were purchased from Sigma. For introduction of plasmids into *Streptomyces coelicolor,* plasmids were transformed into *E. coli* strain ET12567 containing the pUZ8002 plasmid allowing for conjugal transfer of the plasmid (26). *S. coelicolor* M145 was used for the HTS screen and subsequent follow up, unless stated elsewhere. All actinomycetes strains were grown on solid R5M (100g maltose, 10.12g $MgCl_2$*$6H_20$, 0.5g K_2SO_4, 0.2g Difco casaminoacids, 10g yeast extract, 11.46g TES, 4 ml trace elements, 10 ml (0.5%) K_2PO_4, 4 ml (5M) $CaCl_2$*$2H_20$, 15 ml (20%) L-proline, 7ml 1M NaOH per litre) or MS agar at 30^{o}C (27). *Kutzneria* sp. 744, *S. pristinaespiralis* ATCC 25486 and *Streptomyces peucetius* 27952 were used for LC-MS analysis.

4.5.2 Actinorhodin and prodiginine production. Liquid R5M cultures containing 2 x 10^7 cfu/ml of *S. coelicolor* were grown for 16 hours at 30^{o}C and sub-cultured into 5 ml aliquots containing either DMSO or compounds of interest. DMSO concentration was kept constant at 0.1% for liquid analysis. Actinorhodin and prodiginine production were assessed after 72 hours in the presence of 25

μM compounds or DMSO, unless otherwise stated. Both metabolites where measured spectrophotometrically as previously described (27). Briefly, 1 ml aliquots were centrifuged and supernatants extracted with 1:1 chloroform:acidic methanol. The chloroform phase was read at 542 nm corresponding to γ-actinorhodin. The pellet was weighed and 1 ml of acidic methanol added, the supernatant was then read at 530 nm, corresponding to prodiginines. Analysis of actinorhodin production was also assessed visually on MS agar with DMSO concentrations kept constant at 1%.

4.5.3 LC-MS analysis. R5M was inoculated with *S. coelicolor, Kutzneria* sp. 744, *S. pristinaespiralis* ATCC 25486 and MS agar was inoculated with *Streptomyces peucetius* 27952 in the presence of either DMSO or 10 μM ARC2 and grown for 7 days at 30^{o}C. Strains were extracted in an equal volume of butanol, sonicated and left overnight. Extracts were filtered through Whatman paper and evaporated. Extracts were resuspended in 1:1 CH_3CN:H_2O and subjected to LC-MS analysis. CDA was extracted after 7 days growth on R5M as previously described (28). LC-MS analysis was performed on an Agilent 1200 “RR” series LC system coupled to a Bruker micrOTOF II time-of-flight MS equipped with an ESI ionization source. LC was carried out using a Phenomenex Kinetex C18 column (50x2.1mm, 2.6μm, 100 Å), H_2O + 0.1% formic acid (A) and CH_3CN + 0.1% formic acid (B) as solvents, 40˚C column block and the following gradient: flow 0.2 ml/min, 0.5 min 5% B, 5-9 min 95% B, 10-15 min 5% B. The MS

conditions were set to a capillary voltage of 4.5 kV for positive mode, nebulising gas pressure (N2) of 3 barr, dry gas flow rate (N2) of 6 L/min, temperature at 200˚C and a scan rate of 1 Hz.

4.5.4 Structure activity library construction and analysis. Compounds 1 to 8 were purchased from Maybridge. Compounds 9 to 35 were synthesized as described in the supplemental information of the published version of this chapter. Activity was assessed visually by the production of actinorhodin when spotted as a titration (0.78, 1.56, 3.13, 6.25 12.5, 25, 50, 100 μM) on MS agar. Outer columns contain DMSO only and DMSO was kept constant at 1%.

4.5.5 FabV complementation. FabV was PCR amplified from *Pseudomonas aeruginosa* PAO1 chromosomal DNA and introduced into pSET152 containing the *ermE** promoter through the use of the EcoRV and BamHI restriction sites. pSET152*ermE**p and pSET152 *ermE**p*fabV* were introduced into *S. coelicolor* by conjugal transfer. The ability of FabV to compliment the antimicrobial activity of triclosan and other antibiotics was assessed as a titration of compounds (0.78, 1.56, 3.13, 6.25 12.5, 25, 50, 100 μM) on solid MS agar with outer wells containing DMSO only. Prodiginine complementation was assessed through LC-MS analysis. *S. coelicolor* M145 containing empty vector or FabV were grown for 7 days on solid MS agar in the presence of DMSO or ARC2 (1, 2.5, 5, 10 μM), extracted with butanol, dried and resuspended in 1:1 CH_3CN:H_2O.

4.5.6 FabI Enzyme Assay. The assay was conducted as previously described (24). Briefly, 10 μM of *E. coli* FabI (Sino Biological) in the presence of PIPES (30 mM), NaCl (150 mM), NADH (250 μM) and crotonyl-CoA (200 μM) was assessed for its ability to convert the crotonyl-CoA substrate. The rate of reaction was determined by the decrease in absorbance at 340 nm corresponding NADH consumption during crotonyl-CoA conversion by FabI. Crotonyl-CoA was added last to initiate the reaction.

4.6 References

1. Olano C, Lombo F, Mendez C & Salas JA (2008) Improving production of bioactive secondary metabolites in actinomycetes by metabolic engineering. *Metab Eng* 10: 281-292.
2. Martin JF & Liras P (2010) Engineering of regulatory cascades and networks controlling antibiotic biosynthesis in *Streptomyces*. *Curr Opin Microbiol* 13: 263-273.
3. Okamoto S, Taguchi T, Ochi K & Ichinose K (2009) Biosynthesis of actinorhodin and related antibiotics: Discovery of alternative routes for quinone formation encoded in the act gene cluster. *Chem Biol* 16: 226-236.
4. Mo S, *et al* (2008) Elucidation of the *Streptomyces coelicolor* pathway to 2-undecylpyrrole, a key intermediate in undecylprodiginine and streptorubin B biosynthesis. *Chem Biol* 15: 137-148.

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

6. Bentley SD, *et al* (2002) Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* 417: 141-147.

7. Laureti L, *et al* (2011) Identification of a bioactive 51-membered macrolide complex by activation of a silent polyketide synthase in *Streptomyces ambofaciens. Proc Natl Acad Sci USA* 108: 6258-6263.

8. McKenzie NL, *et al* (2010) Induction of antimicrobial activities in heterologous streptomycetes using alleles of the *Streptomyces coelicolor* gene *absA1*. *J Antibiot* (Tokyo) 63: 177-182.

9. Gomez-Escribano JP & Bibb MJ (2011) Engineering *Streptomyces coelicolor* for heterologous expression of secondary metabolite gene clusters. *Microb Biotechnol* 4: 207-215.

10. Komatsu M, Uchiyama T, Omura S, Cane DE & Ikeda H (2010) Genome-minimized *Streptomyces* host for the heterologous expression of secondary metabolism. *Proc Natl Acad Sci USA* 107: 2646-2651.

11. Aoki Y, Matsumoto D, Kawaide H & Natsume M (2011) Physiological role of germicidins in spore germination and hyphal elongation in *Streptomyces coelicolor* A3(2). *J Antibiot* (Tokyo) 64: 607-611.

12. Lewis RA, *et al* (2011) Active site modification of the beta-ketoacyl-ACP synthase FabF3 of *Streptomyces coelicolor* affects the fatty acid chain length of the CDA lipopeptides. *Chem Commun* 47: 1860-1862.

13. Hojati Z, *et al* (2002) Structure, biosynthetic origin, and engineered biosynthesis of calcium-dependent antibiotics from *Streptomyces coelicolor*. *Chem Biol* 9: 1175-1187.

14. Barona-Gomez F, *et al* (2006) Multiple biosynthetic and uptake systems mediate siderophore-dependent iron acquisition in *Streptomyces coelicolor* A3(2) and *Streptomyces ambofaciens* ATCC 23877. *Microbiology* 152: 3355-3366.

15. Lomovskaya N, *et al* (1998) The *Streptomyces peucetius dpsY* and *dnrX* genes govern early and late steps of daunorubicin and doxorubicin biosynthesis. *J Bacteriol* 180: 2379-2386.

16. Mast Y, *et al* (2011) Characterization of the 'pristinamycin supercluster' of *Streptomyces pristinaespiralis*. *Microb Biotechnol* 4: 192-206.

17. Bergler H, Fuchsbichler S, Hogenauer G & Turnowsky F (1996) The enoyl-[acyl-carrier-protein] reductase (FabI) of *Escherichia coli*, which catalyzes a key regulatory step in fatty acid biosynthesis, accepts NADH and NADPH as cofactors and is inhibited by palmitoyl-CoA. *Eur J Biochem* 242: 689-694.

18. Barona-Gomez F, Wong U, Giannakopulos AE, Derrick PJ & Challis GL (2004) Identification of a cluster of genes that directs desferrioxamine biosynthesis in *Streptomyces coelicolor* M145. *J Am Chem Soc* 126: 16282-16283.

19. Williamson NR, Fineran PC, Leeper FJ & Salmond GP (2006) The biosynthesis and regulation of bacterial prodiginines. *Nat Rev Microbiol* 4: 887-899.

20. Lu H & Tonge PJ (2008) Inhibitors of FabI, an enzyme drug target in the bacterial fatty acid biosynthesis pathway. *Acc Chem Res* 41: 11-20.

21. Gago G, Diacovich L, Arabolaza A, Tsai SC & Gramajo H (2011) Fatty acid biosynthesis in actinomycetes. *FEMS Microbiol Rev* 35: 475-497.

22. Tipparaju SK, *et al* (2008) Design and synthesis of aryl ether inhibitors of the *Bacillus anthracis* enoyl-ACP reductase. *ChemMedChem* 3: 1250-1268.

23. Zhu L, Lin J, Ma J, Cronan JE & Wang H (2010) Triclosan resistance of *Pseudomonas aeruginosa* PAO1 is due to FabV, a triclosan-resistant enoyl-acyl carrier protein reductase. *Antimicrob Agents Chemother* 54: 689-698.

24. Liu N, Cummings JE, England K, Slayden RA & Tonge PJ (2011) Mechanism and inhibition of the FabI enoyl-ACP reductase from *Burkholderia pseudomallei*. *J Antimicrob Chemother* 66: 564-573.

25. Koppisch AT & Khosla C (2003) Structure-based mutagenesis of the malonyl-CoA:Acyl carrier protein transacylase from *Streptomyces coelicolor*. *Biochemistry* 42: 11057-11064.

26. Flett F, Mersinias V & Smith CP (1997) High efficiency intergeneric conjugal transfer of plasmid DNA from *Escherichia coli* to methyl DNA-restricting streptomycetes. *FEMS Microbiol Lett* 155: 223-229.

27. Kieser T, Bibb M, Buttner MJ, Chater KF & Hopwood DA (2000) Practical *Streptomyces* Genetics, (The John Innes Foundation, Norwich), pp 613.

28. Powell A, *et al* (2007) Engineered biosynthesis of nonribosomal lipopeptides with modified fatty acid side chains. *J Am Chem Soc* 129: 15182-15191.

SUMMARY AND CONCLUDING REMARKS

We have performed the first chemical screen against the lifecycle of *S. coelicolor* and this work provides the foundation for studying the streptomycetes by chemical manipulation. The visual progression of both development and secondary metabolism was instrumental in our search for small molecules with activity against the *S. coelicolor* lifecycle, enabling us to recapitulate all phenotypes chemically which have been achieved by genetic manipulation. We believe exploring these visual cues through a chemical biological approach will further the understanding of this important class of bacteria and provide chemical tools for widespread use within the streptomycetes.

While the screen identified numerous worthwhile processes for study by chemical means, there are a few of particular interest. *S. coelicolor* can be used as an organism to identify potential antimicrobial agents. *S. coelicolor* is related to many pathogenic bacteria within the actinobacteria, including *M. tuberculosis* and can provide a non-pathogenic alternative for screening purposes. In support of this, we have identified small molecules which inhibit *S. coelicolor* growth that were previously identified to kill *M. tuberculosis*. The streptomycetes also possess many protein families in common with eukaryotic fungi, including the target of the azole class of antifungals – suggesting we may also be able to identify growth inhibitors with antifungal properties. As well, the dispensability of the usually essential process of cell division in other bacteria provides a cell-based assay against this process. Small molecules interfering with cell division in

S. coelicolor will result in a white phenotype as cell division is only required during aerial hyphae maturation. This presents a unique opportunity to detect cell division inhibitors and evaluate their potential as novel antimicrobials. Those compounds whose addition resulted in precocious sporulation represent an interesting class of molecules for further study as this phenotype is under reported by genetic analysis. Since factors stimulating development are relatively unexplored, the use of molecules causing precocious sporulation may lead to a greater understanding of this complex process. Finally, altering secondary metabolism chemically will be of great value as these molecules can be used to further understand secondary metabolism and exploit the secondary metabolite potential of the actinomycetes to identify metabolites with novel bioactive properties. The compounds identified in our analysis are capable of affecting *S. coelicolor*'s pigmented secondary metabolites in many different ways – some have similar effects on both pigments (either increasing or decreasing yields), others have opposing effects, while the remainder alter the production of only a single pigment. These phenotypes are further confounded by the effect of these compounds on development, as their addition can also result in improved or impaired sporulation. This implies that these compounds are affecting *S. coelicolor* through a variety of different mechanisms and understanding their effects will provide insight into the many factors governing secondary metabolism and offer a number of chemical tools for application in the actinomycetes.

The creation of our high-GC luminescent transcriptional reporter, based on the *luxCDABE* cluster from *P. luminescens*, will enable the analysis of changes in gene expression by these small molecules and aid in understanding how they exert their effects. Prior to the development of this reporter, transcriptional analysis of chemical manipulation in *S. coelicolor* would have been impossible as thorough analysis requires multiple concentrations and multiple time points throughout its 5 days lifecycle – too tedious to perform by RNA methods. This reporter has also been used in other research projects in the Nodwell lab, including the study of a conserved family of transcriptional regulators, the TetR-like proteins. It has aided in identifying the ligands sensed by these regulators and in understanding their expression patterns during growth. Our synthetic *lux* genes is also in use in research laboratories throughout the world and has even been reported to give luminescence in *Mycobacterium*. This high-GC luminescent transcriptional reporter will be invaluable to *Streptomyces* research, enabling routine genetic analysis to be performed with ease.

Many strategies have been adopted to improve secondary metabolite production in the streptomycetes. These strategies can be cell-based such as mutagenesis, ribosomal engineering and the over-expression of pleiotropic regulators, or cluster-based, where the individual biosynthetic cluster is altered to improve production. As well, scanning genomes can identify the classes of secondary metabolites present and reveal a species bioactive potential. These methods are vital for the continual supply of new secondary metabolites that can be used to

benefit human health. Our strategy in this endeavour has been to identify small molecules capable of altering secondary metabolism. Small molecule manipulation has significant advantages for the discovery of novel secondary metabolites. There is no requirement for genetic manipulation which can be tedious in widespread application. As well, the ease of addition allows titratable modification to the producer strain to fine-tune the level of response, a feat difficult to achieve genetically. Most importantly, it is amenable to high throughput analysis which is required in order to identify novel bioactive metabolites. For example, those libraries of streptomycetes that have been previously screened could be rescreened in the presence of small molecules as changes in their secondary metabolic profiles may lead to the production of previously undetected metabolites.

By making use of the pigmented secondary metabolites of *S. coelicolor*, we were able to identify small molecules with bioactivity against this process. Hits from the primary screen were narrowed down to 19 ARCs (antibiotic remodeling compounds) characterized by a strong ability to stimulate actinorhodin production. Work presented in this thesis delves into a set of 4 structurally related ARCs, the ARC2 series. The ARC2 series was chosen as we found it significant that of over 30,000 compounds this structure was identified repeatedly. This, in combination with the opposing effect the ARC2 series exhibits on blue and red pigment yields suggested this would be an interesting set of compounds to begin studying secondary metabolism by chemical perturbation. The target

process of the ARC2 series was revealed to be inhibition of fatty acid biosynthesis, specifically the enoyl reductase FabI which is involved in the last step of fatty acid chain elongation. Since fatty acid biosynthesis is essential for growth, weak inhibition is key to the ARC2 series effect on secondary metabolism. While strong inhibitors of FabI, such as triclosan, were also capable of altering secondary metabolism, their potent growth inhibition limits their utility to a very narrow working concentration just below their inhibitory concentration. The ARC2 series remodels the secondary metabolic profile of *S. coelicolor* with improved production of two acetyl-CoA dependent metabolites, actinorhodin and the germicidins and reduced production of two lipid dependent metabolites, calcium-dependent antibiotic and the prodiginines. The simplest explanation of this is that weak inhibition of fatty acid biosynthesis leads to a buildup of the fatty acid precursor acetyl-CoA, resulting in improved production of acetyl-CoA dependent metabolites. As well, the inhibition of lipid production would extend to those lipid containing secondary metabolites, resulting in their decreased production. While the influence of ARC2 on the acetyl-CoA pool has yet to be measured to confirm this hypothesis, the changes in *S. coelicolor* secondary metabolism correlate with this hypothesis.

The ARC2 series targets a well conserved process within the cell, fatty acid biosynthesis. The conservation of this process ensures the target is always present and widespread use of the ARC2 series to modify secondary metabolism should be achievable. We believe this unexploited bacterial response to fatty

acid inhibition provides a unique opportunity to identify previously uncharacterized secondary metabolites. Currently, we are using the ARC2 series to identify novel bioactive metabolites within the actinomycetes. Consistent with the alterations observed by the ARC2 series in *S. coelicolor*, we observe both increased and decreased production of secondary metabolites. We speculate based on the changes in *S. coelicolor*, decreased metabolites are most likely lipid containing or lipid-dependent metabolites; while increases are likely to be those metabolites that rely on the primary metabolic pool of acetyl-CoA precursors such as the polyketides. Characterization of these metabolites in the actinomycetes is underway to determine if this paradigm holds true in general for the observed secondary metabolic changes.

Further understanding of the ARC2 series is necessary to truly exploit its effect on secondary metabolism. Preliminary transcriptional analysis in the presence of ARC2 (see **Appendix 2**) suggests the pleiotropic regulator AfsS may be involved in the ARC2 effect. AfsS expression results in improved secondary metabolism in *S. coelicolor* and the induction of this regulator would heighten the stimulatory effect achieved by the ARC2 series, possibly adding a second level of response in addition to fatty acid inhibition. While AfsS involvement remains to be experimentally determined, this is an interesting avenue to pursue.

One final area to explore is the effect of the ARC2 series on the enoyl reductase enzymes encoded within the biosynthetic clusters of some secondary metabolites. While enoyl reductase activity is not required for the majority of

secondary metabolites, there are many examples where this activity is required to produce the final product. It would be interesting to establish whether the ARC2 series inhibits this step of synthesis and if so, will this inhibition result in the production of a new metabolite. Preliminary evidence suggests that the ARC2 series may result in changing the proportion of enoyl reductase dependent metabolites; however, this needs to be further explored.

The ARC2 series exemplifies the utility of chemical probes for use in the streptomycetes. These probes have demonstrated an important link between primary and secondary metabolism that can be exploited to the benefit of secondary metabolism. As well, the involvement of AfsS in the response to fatty acid inhibition would add new information as to how *S. coelicolor* regulates secondary metabolism. Not only has the use of the ARC2 series lead to a greater understanding on *S. coelicolor* secondary metabolism, these chemical probes can be directly applied to all actinomycetes to alter their secondary metabolic profiles; achieving the ultimate goal of this endeavour, to create a simple and effective strategy for improved secondary metabolism in the actinomycetes.

APPENDIX 1

A CLASS OF BACTERIAL MEMBRANE PROTEINS LACKING N-TERMINAL SIGNAL SEQUENCES

Adapted from:

Arryn Craney, Kapil Tahlan, David Andrews, and Justin Nodwell. 2011. Bacterial transmembrane proteins that lack N-terminal signal sequences. PLoS One.6(5):e19421.

A1.1 Abstract

Tail-anchored membrane proteins, which have no signal sequence and are targeted to the membrane independently of the Sec translocon, play critical roles in apoptosis, vesicle trafficking and other vital processes in eukaryotic organisms. Until recently, this class of membrane proteins has been unknown in bacteria. Here we present the results of bioinformatic analysis revealing proteins that are superficially similar to eukaryotic TAMPs in the bacterium *Streptomyces coelicolor.* We demonstrate that at least five of these proteins are bona fide, membrane-spanning proteins, that none has an N-terminal signal sequence and that the C-terminal membrane-spanning domain is sufficient for membrane targeting. Several of these proteins, including a serine/threonine kinase and the SecE component of the Sec translocon, are widely conserved in bacteria.

A1.2 Introduction

Protein translocation into and across the lipid bilayer is an essential process in all kingdoms of life. Most proteins are inserted into the membrane by the well-conserved Sec pathway, consisting of a membrane-spanning translocase SecYEG in bacteria and Sec61 in eukaryotes. Many accessory proteins aid in protein targeting and insertion, including in particular the signal recognition particle (SRP), and its cognate membrane receptor (1, 2). To be targeted to the membrane via the Sec system, a protein must have an N-terminal signal sequence for recognition by the SRP. Signal sequences are typically 20-30 amino acids long and consist of an N-terminal domain with one or more positively charged amino acids, followed by an H-domain of 8-12 hydrophobic residues and, for proteins that are secreted, a C-domain recognition site for peptide cleavage (3). During co-translational targeting the signal sequence is recognized and bound by the SRP as the N-terminus of the nascent polypeptide emerges from the ribosome. The ribosome/nascent peptide is then brought to the membrane for insertion through an interaction with the SRP receptor FtsY (4) and transferred to the SecYEG translocon for insertion (5). [For reviews of Sec translocation see (3, 6 - 9 and 10)].

While the majority of membrane proteins are targeted to the membrane via Signal sequence/Sec translocon-dependent mechanisms, another system has been identified in eukaryotes for targeting tail anchor membrane proteins (TAMPs) (11,

12). Eukaryote TAMPs carry out a wide range of biological functions, many of which involve membranes. Examples include the Bcl-2 protein, a major player in the apoptosis pathway, the SNARE proteins which are involved in vesicle targeting and fusion, and the Sec61β protein, which is a component of the eukaryotic Sec translocon (13, 14). Bcl-2, the SNAREs and Sec61β all lack the N-terminal signal sequence required for SRP-targeting and are instead targeted to the membrane via a single C-terminal transmembrane domain, the tail anchor.

All of the TAMPs that have been investigated biochemically to date are found in eukaryotes. Recently however, a bioinformatic approach was used to demonstrate the existence of TAMP-like proteins in the Gram-negative bacteria *Escherichia coli* and *Rickettsia prowazekii* as well as the archeon *Methanococcus maripaludis* (15). This work suggests that in fact, tail-anchored membrane proteins are universal and that they make up similar proportions of all proteomes (15). Our work adds to this, providing experimental evidence of these bacterial tail anchor membrane proteins. We have taken advantage of a newly developed algorithm, TAMP finder to identify membrane-proteins encoded in the Gram-positive bacterium *Streptomyces coelicolor*. Similarly, we find a large number of proteins that are superficially similar to the eukaryotic TAMPs in that they lack signal sequences and contain single C-terminally located transmembrane domains. We have used several biochemical approaches to test these predictions and find that indeed, many of these proteins are transmembrane

proteins and that the tail sequences are sufficient for membrane targeting. These include important proteins including the SecE component of the translocon and members of the bacterial serine/threonine (Ser/Tthr) protein kinase family.

A1.3 Results

A1.3.1 Putative membrane proteins lacking signal sequences and exhibiting broad conservation in prokaryotes

A genome-wide search using the "TAMP finder" program identified 73 putative tail-anchor membrane proteins (TAMPs) in *S. coelicolor.* This program was designed to identify TAMPs encoded in eukaryotic genomes by seeking polypeptide sequences having the known TAMP properties. These include a putative C-terminally located transmembrane domain, the tail anchor, and the absence of an obvious N-terminal signal sequence. To further test these candidates, we analyzed each of them individually using the transmembrane prediction program TMHMM (16). We restricted subsequent analysis to those proteins having one or, in a few cases two, strongly-predicted transmembrane domains near the C-terminus. We then used SignalP, a program that predicts SRP signal sequences, and visual inspection to further confirm that these proteins lacked candidate N-terminal signal sequences (17). Twenty of the 73 predicted polypeptides identified by TAMP finder met both criteria (**Table A1.1** and **Figure A1.1**). During this analysis, careful consideration was taken in scanning the upstream regions of the predicted translational start site to ensure

proteins were not mis-annotated. Those with mis-annotated start sites that contained N-terminal signal sequences were removed from the analysis.

Table A1.1. Highest Confidence *S. coelicolor* Tail Anchor Membrane Proteins

Protein	Size(aa)	Function	N-terminus	Tail Anchor	Homologues
SCO1166	110	hypothetical	out	AAGLILLIWLPWWAALLIVLGVPAAAYLTLDPSQRRRLRRVSRKEIG	streptomycete
SCO1431	80	hypothetical	out	PKILEHVLGWTLVVVVAMLVVQLGLL	streptomycete
SCO2124*	205	hypothetical	in	WLTTLSIGGFLGGFATLVVRMRTGDEDDDDPGRGAVV	actinomycete
SCO2199	89	hypothetical	out	VGSRRRSSWVSTVVVLGCVAAVIVLLGYLNFRAPY	streptomycete
SCO2900	110	hypothetical	out	TGAPRMERVVPVALVVAGVVGLLALGGTRRRKR	actinomycete
SCO2973	417	Ser/Thr Kinase	in	RRRRIAVGAGAVALVAAIGVGTWLATGGDEDGGGPQDTRNSAPAAP	actinomycete
SCO3544*	132	hypothetical	in	PVALGVSPVASATVASVAAVVALGLGAWCLTQV	actinomycete
SCO3860	576	Ser/Thr Kinase	in	RRRRRPGPPARVALPVLLLALACYAVGFWALTRI	streptomycete
SCO4008	192	TFR	in	APDLLFLLVAMANWAVVVPQMKRILVGGGDAGTDGLRDSIKKAARRI	actinomycete
SCO4033	96	hypothetical	in	AASSGPRVGLIVGIVAAVIVVAAVAWLALG	streptomycete
SCO4174	83	hypothetical	in	HKARSRRRAGLDGATVSGLLTVLCVATLLVTITFAV	*S. lividans*
SCO4646	94	SecE	in	SRNQLTTYTTVVIIFVVIMIGLVTLIDYGFSHAAKYVFG	Most bacteria
SCO4959	85	hypothetical	in	TAARRLMWLLLGAAAVAFTVWALTVQPWVEPPSETTPPVTGWEGWS	streptomycete
SCO5157*	317	CorA	in	DYMPETHWKFGYPLVLSVTVCICLGIHRTLKRNGWL	Most baceteria
SCO5344*	107	SpdD2	in	GGGTAVVLVVGAVLVSMLLAVAITAASVAVCAVVLRSLLASDAKRR	streptomycete
SCO6904	336	hypothetical	out	GADATLWLIGGAAVLIAAGGGALAVARRSRTDSHTQDNTGS	streptomycete
SCO7096	114	hypothetical	in	RRYARLRRMSRVALAVLAATVMVLLVALVLVAAG	streptomycete
SCO7133	113	hypothetical	in	RGTMIAMTAIGLTIFVCTAVVVGSMT	streptomycete
SCO7199	131	hypothetical	in	RRLGRILAGAAALAVLLGLFTCLPEEPPGLPTGPEDTSPPRTSSAVV	streptomycete
SCO7330	78	hypothetical	out	GWAKGPMALILAVVVIFAVGLLGYALALIY	actinomycete

* denotes 2 predicted transmembrane domains at the C-terminus

Putative signal sequences at the N-termini of the *S. coelicolor* FtsY and four other proteins annotated as membrane proteins are shown in **Figure A1.1A**. All have stretches of 8-10 hydrophobic residues: these are the predicted binding sites of the SRP (3). In contrast, the known cytoplasmic protein ActR has a hydrophilic N-terminus (**Figure A1.1B**). Similarly, the 20 candidates listed below ActR, with the exception of SCO6904, also have largely hydrophilic N-termini (**Figure A1.1C**). These proteins therefore lack obvious N-terminal signal sequences. The 'twin-arginine repeat' or TAT pathway is involved in the secretion of folded proteins and has not been implicated in membrane insertion (18). We note however that these candidates also lack the characteristic Z-R-R-φ-X-X (where Z is a polar

residues, X-X are hydrophobic residues and φ is any residue) although SCO4033 has two arginines embedded in the N-terminal sequence A**RR**PRTWAALA. It is unlikely that this could serve as a target for TAT-mediated secretion.

A

SRP Signal Sequence		Average Hydrophobicity	% Hydrophilicity
FstY	MEIVILAVVIAVVVIGALGGLVVGS	-1.0	8
SCO0041	MTPAVAALLMLAGLSEALGRVTPVA	-0.5	12
SCO1101	MILPAALLLLGALTAVLAPRLLARA	-0.7	8
SCO1374	M**RRR**TGAAGIAVAIAAIVPLADPAP	-0.2	16
SCO2176	MEPGKLPLLLALAAFVLTFAVTRVI	-0.7	12

B

Cytoplasmic Control			
ActR	VSRSEEGRPMPEEIPVPPW**RR**PKKA	0.8	48

C

Putative TAMPs			
SCO1166	MPSRTQSPTGSQGRVGVGPQGVLRV	0	36
SCO1431	MTKLSAAKAAHGGKVDTVSGGTPPS	0.1	28
SCO2124	MAEHDSDREDREERDLEREHREGRE	2.0	76
SCO2199	MPEMPDRDDEEREYDLRWAEGAEHK	1.2	56
SCO2900	MADTSDIRTPAQIEADIK**RR**REVLA	0.7	48
SCO2973	MARKIGSRYTAHQILGRGSAGTVWL	-0.2	28
SCO3544	MGARTHFFSRVRP**RR**PPHPVYGGRP	0.2	28
SCO3860	MGEVFAGRYELVDPIGRGGVGAVWR	-0.1	24
SCO4008	MAARDPEATKARIFEAAVAEFARHG	0.4	32
SCO4033	M**RR**WYGVEWVKRA**RR**PRTWAALADS	0.3	40
SCO4174	MTYDGGLPYEEC**RR**AQRLAP**RR**DRA	0.7	44
SCO4646	MTDAVGSIDMPDAQDEAPDSKKSRK	0.9	56
SCO4959	MNRP**RR**EAAA**RR**IMEQSPPRVPADL	0.7	48
SCO5157	MSMIRDLRAAVRPSRVSLRKDGGAY	0.4	44
SCO5344	MLRPKLPTMPQPTSTVTPPAVIEPT	-0.2	20
SCO6904	MHAHLVRSAAATAVAVAGTLTWVPA	-0.7	8
SCO7096	MDAKDEAERGLTVIQDYLYWDSH**RR**	0.5	48
SCO7133	MDTSKQAARGLADIQAFLYREAHLS	0.1	40
SCO7199	MPNPTPESVWTRFLTDSEEAIARSA	0.2	40
SCO7330	MSYVNPDPDPRLSTGLEPGGGVPPG	0.1	28

Figure A1.1. Putative bacterial membrane proteins lacking N-terminal signal sequences. (A) The N-terminal sequences of five strongly predicted *S. coelicolor* transmembrane proteins (FtsY, SCO0041, SCO1101, SCO1374 and SCO2176) are shown, illustrating their high hydrophobicity and correspondingly low hydrophilicity. Hydrophobic residues are shaded grey. Double arginine residues are bolded. **(B)** The N-terminal sequences of the *S. coelicolor* cytoplasmic protein (ActR) and **(C)** 20 predicted tail anchor membrane proteins lacking obvious signal sequences are shown to illustrate their highly hydrophilic N-terminus.

The 20 candidates in **Table A1.1** represent a wide range of important biological functions. These include a conspicuous number of hypothetical membrane proteins of less than 100 amino acids (SCO1431, SCO2199, SCO4033, SCO4174, SCO4959 and SCO7330), two serine/threonine protein kinases (SCO2973 and SCO3860), the SecE component of the Sec translocon (SCO4646) proposed to be a tail-anchored membrane protein in many organisms, including *Archea* (15), a CorA-like Mg^{2+} transporter (SCO5157) (19) and the SpdD2 protein believed to be involved in transfer of plasmid DNA in streptomycetes (SCO5344) (20,21). Many of these proteins are highly conserved in the actinomycetes and two are conserved generally in prokaryotes (15,19) (**Table A1.1**). While, the majority of these proteins are predicted to have a topology with the N-terminus facing into the cell, several are predicted to have their N-termini projecting out of the cell (**Table A1.1**).

While a large number of these proteins are small hypothetical proteins, we are confident that these represent expressed genes rather than artefacts of genome annotation. Only membrane proteins conserved in multiple streptomycetes and possible having orthologues in other actinomycetes were included in our analysis. For example, SCO2900 is predicted to encode a 110 residue polypeptide that is conserved within the streptomycetes and related actinomycetes (**Figure A1.2A**). Conserved features of this protein include three absolutely conserved residues (P29 R34 P37 with respect to the *S. coelicolor* protein sequence) and the C-

terminal transmembrane domain followed by a small stretch of 4 positively charged residues, suggesting an N-terminus "out" orientation. Sequences from related actinomycetes were found to contain an approximately 30 residue deletion upstream of the predicted tail anchor and there is a small C-terminal extension in *Corynebacterium* proteins. SCO7133-like proteins were found in some streptomycetes, and in no other genera (**Figure A1.2B**). Although a C-terminal transmembrane domain is consistently predicted among SCO7133 paralogues, the amino acid identity in this domain is low. Four positively charged residues are located directly upstream of the transmembrane domain, suggesting the N-terminus of this protein is facing into the cell. This predicted topology was shared among the SCO7133-like paralogues. The large N-terminal extension predicted in *S. lividans* TK24 is most likely mis-identification of the start site; regardless, this extended region does not contain an N-terminal signal sequence.

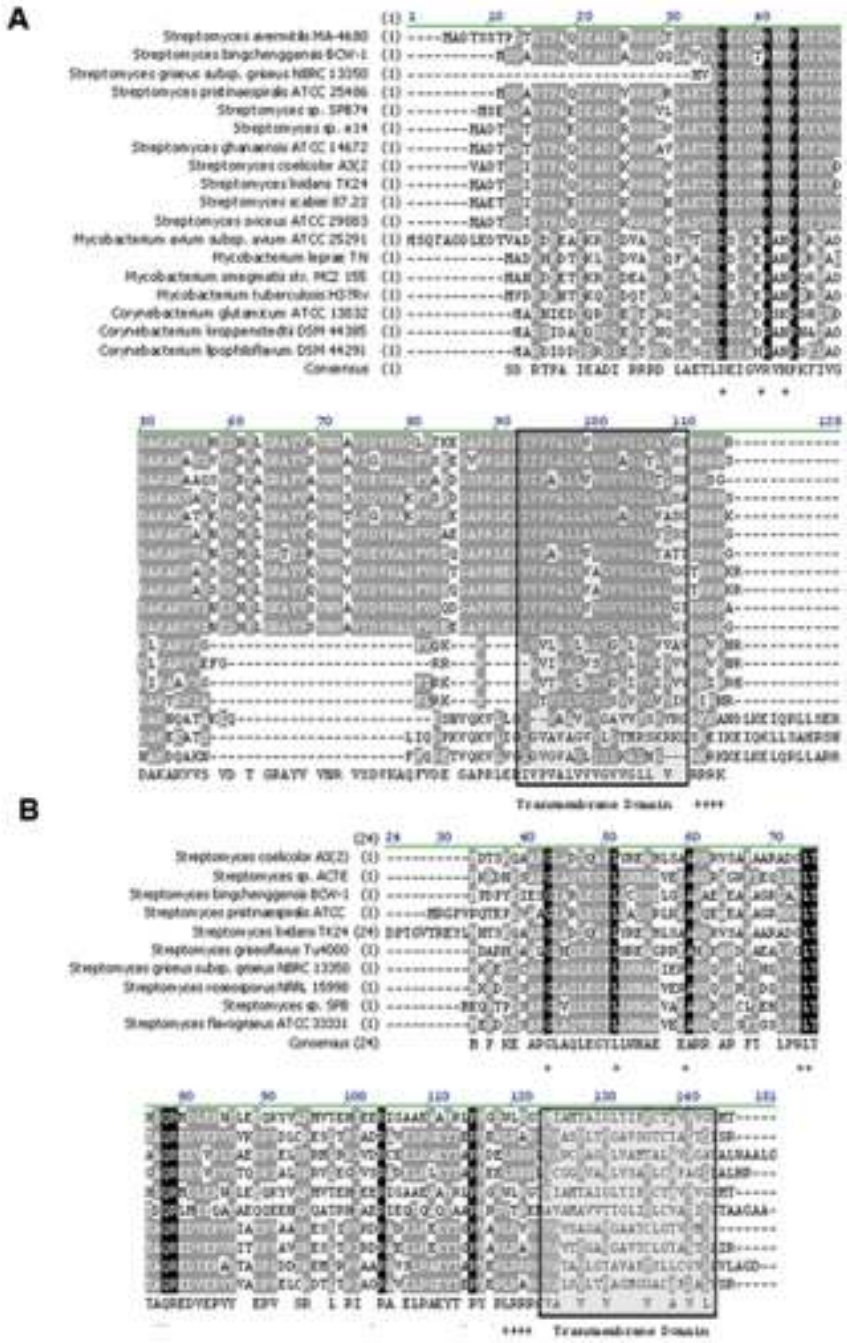

Figure A1.2. Alignments of predicted tail-anchor membrane proteins. (A) SCO2900 from *S. coelicolor* aligned with various orthologues from other streptomycetes and actinomycetes **(B)** SCO7133 from *S. coelicolor* aligned with various orthologues from other streptomycetes. The *S. lividans* extended leader sequence is MGRHRPREDRRPTGTAPTAAPRH. Absolutely conserved residues are shaded black and marked with *, similar residues are shaded grey. Possible topology predicting residues are marked with + and the C-terminal transmembrane domains are boxed and shaded grey.

A1.3.2 Four integral membrane proteins

We chose five of the candidates in **Table A1.1** to test the prediction that they are integral membrane proteins: two small hypothetical proteins (SCO2900 and SCO7133), the Ser/Thr kinase PkaB (SCO2973), SecE (SCO4646) and a predicted TetR-like transcription factor (SCO4008). The known cytoplasmic protein ActR served as a control. All six proteins were expressed in *S. coelicolor* under the thiostrepton-inducible promoter *tipA* such that they had an N-terminal FLAG-tag for visualization by Western analysis.

Protoplasts of cells expressing these proteins were isolated after lysozyme-treatment. The protoplasts were subsequently lysed, fractionated by ultracentrifugation and the pellets and supernatants analyzed by Western analysis with anti-FLAG antibodies. As expected, ActR was found exclusively in the supernatant (**Figure A1.3A**). Similarly, in spite of having a predicted transmembrane domain, SCO4008 was found exclusively in the supernatant, consistent with its probable role as a DNA binding transcription factor. The other four proteins were contained exclusively in the pellets.

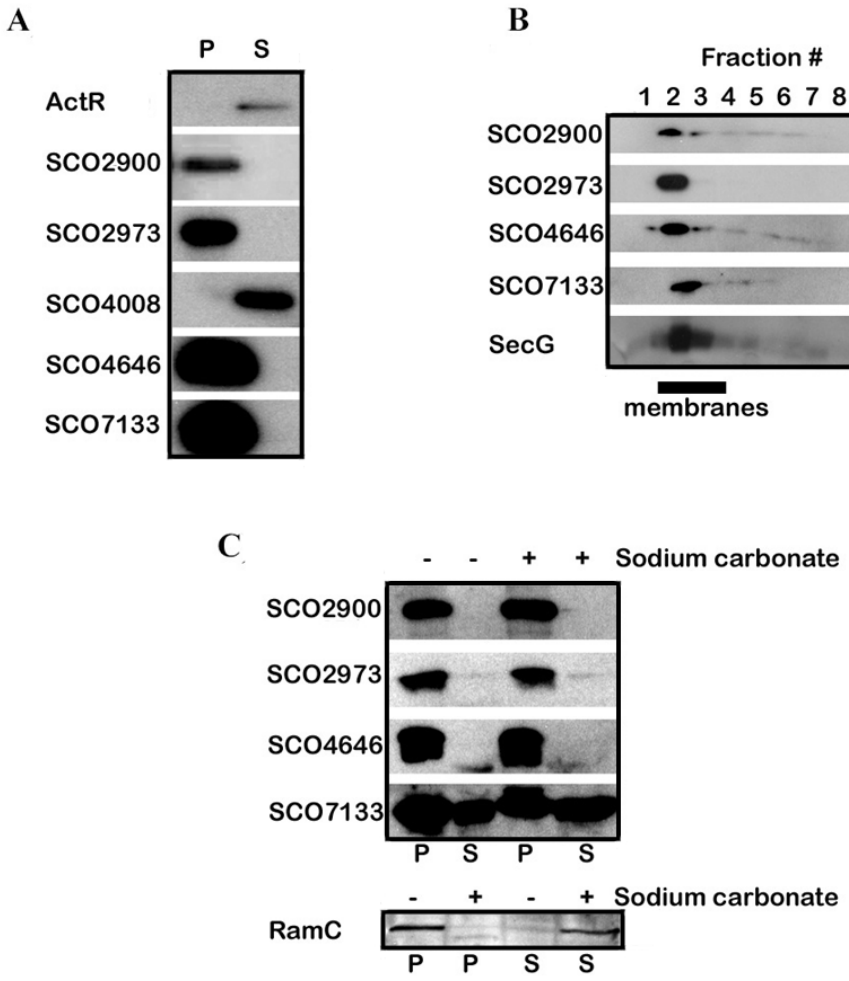

Figure A1.3. Membrane-association of five candidates. **(A)** Cells were fractionated into pellet (P) and supernatant (S) fractions and western blot analysis directed against the FLAG epitope was used to determine the localization of the putative membrane proteins SCO2900, SCO2973, SCO4008, SCO4646 and SCO7133. ActR was used as a cytoplasmic control. **(B)** The pellets from (A) were subjected to sucrose gradient ultracentrifugation and 1 ml fractions were collected with fraction 1 corresponding to the highest density and fraction 10 the lowest. Fractions 2 to 4 (underlined) correspond to sedimentation profiles of known membrane proteins. **(C)** Carbonate extraction of TAMP proteins. Cell lysate was mixed with either sucrose (-) or carbonate (+) and separated into pellet (P) and supernatant (S) fractions. Fractions were subjected to Western blot analysis. The peripheral membrane protein RamC was used as a control.

To determine whether the pellet-associated proteins SCO2900, PkaB, SecE or SCO7133 were membrane-associated, the pellets from this centrifugation step were subjected to sucrose gradient ultracentrifugation. To locate the membrane fractions we used antibodies against the known Sec-dependent transmembrane protein, SecG for western analysis. Consistent with previous analysis of membrane-proteins using this procedure SecG was found primarily in the 2nd and 3rd fractions (**Figure A1.3B**) (22). Consistent with membrane association, SCO2900, SCO2973 (PkaB), SCO4646 (SecE), and SCO7133 were also found predominantly in fraction 2 and 3. None were found in the pellet (**Figure A1.3B**) as would be the case if these proteins were simply insoluble hydrophobic inclusions.

Extraction of the membranes at pH 11.4 using sodium carbonate was then used to distinguish between proteins that were peripherally or integrally associated with the membrane (23). Cells expressing the four candidates demonstrated above to be membrane-associated (**Figure A1.3A** and **A1.3B**) were converted to protoplasts, lysed, subjected to carbonate extraction and then fractionated into membrane-containing (P) and cytosolic (S) fractions. As shown in **Figure A1.3C**, SCO2900, SCO2973 (PkaB) and SCO4646 (SecE) remained entirely in the membrane-containing fraction. Some SCO7133, possibly 30% of the total, was found in the supernatant fractions in this particular experiment. We suspect that this is a result of prolonged induction of the *tipA* promoter that drives expression

of the fusion. Importantly, only a very modest amount of protein was moved from the pellet to the supernatant after carbonate extraction. In contrast, the protein RamC, which we have shown previously to be membrane-associated via interactions with other proteins (24), was almost completely separated from the membranes by treatment with sodium carbonate. This is striking because RamC is an extremely hydrophobic protein and yet could still be rendered soluble in this way. This strongly suggests that the other four proteins remained in the pellet fractions because they are integral membrane proteins.

A1.3.3 Tail sequences are sufficient for membrane targeting

To investigate whether the tail sequences of these proteins are sufficient for membrane targeting, the C-terminal sequences of three candidates, SCO2973 (PkaB), SCO4646 (SecE) and SCO7133 were fused to the cytoplasmic protein eGFP generating eGFP-2973, eGFP-4646 and eGFP-7133. SCO2900 was not included for analysis as its N-terminus is predicted to face out of the cell. Again, these fusions were expressed in *S. coelicolor* using thiostrepton; protoplasts prepared and lysed then fractionated using ultracentrifugation. Fractions containing the fusions were then identified using Western analysis with anti-GFP antibodies. Cross-reactive bands to the eGFP antibody are visible with the eGFP-7133 fusion protein, these bands are also present in the other samples; however, they are not contained in the field of the image. As expected, the

unfused eGFP protein was contained entirely in the supernatant (**Figure A1.4**). In contrast, all fusions to eGFP were found completely in the pellet fractions (**Figure A1.4**).

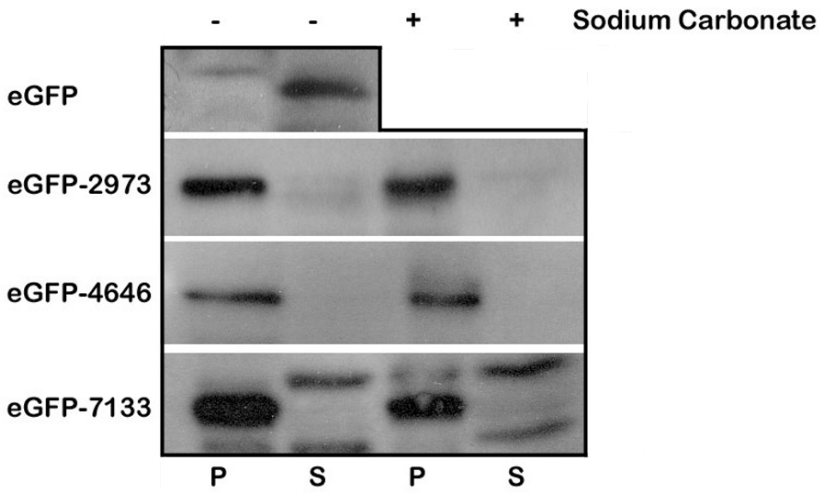

Figure A1.4. Localization of eGFP tail anchor fusions. Putative tail anchor transmembrane domains from SCO2973, SCO4646 and SCO7133 were fused to the C-terminus of the cytoplasmic protein eGFP and localization to the pellet (P) and supernatant (S) fractions was determined in the presence of either sucrose (-) or carbonate (+). Unfused eGFP is shown for comparison. Localization was detected by western blot analysis against eGFP.

The three eGFP-tail sequence fusions were subjected to carbonate extraction to determine whether they behaved as integral membrane proteins. All three proteins remained in the pellet fraction regardless of the treatment with sodium carbonate, suggesting that they were integral membrane proteins (**Figure A1.4**). The ability of the transmembrane domain from the three tail-anchor proteins to relocate eGFP to the pellet and resist carbonate extraction strongly suggests that all information required for targeting to the membrane is found in the C-termini of these proteins.

A1.3.4 Bacterial tail anchor membrane proteins are capable of facing into and out of the cell

During the topology prediction, we noted that while the majority of our putative TAMPs were predicted on the basis of the 'positive charge in' rule (25) to have their N-termini face into the cell, 6 of the 20 were predicted to have their N-termini exterior to the cell, in contrast to the eukaryote paradigm. In order to test this we subjected the 4 candidates (SCO2900, SCO2973, SCO4646 and SCO7133) to Proteinase K digestion with ActR, a cytoplasmic protein, serving as a control for cell lysis (**Figure A1.5**). We found that with high doses of proteinase, all of the fusions were rapidly degraded to the point where they were undetectable by anti-FLAG tag Western analysis (data not shown). At lower proteinase concentrations however, including those shown in **Figure A1.5**, SCO2900 was consistently more sensitive to proteinase digestion than SCO2973, SCO4646 or SCO7133, suggesting this proteins FLAG-tag is external to the cell, along with the bulk of the protein, and that it is therefore susceptible to proteolytic removal. We take this as evidence that while the N-termini of SCO2973, SCO4646 and SCO7133 are intracellular, SCO2900 may project it's N-terminus out of the cell, as predicted by the 'positive charge in' rule (25).

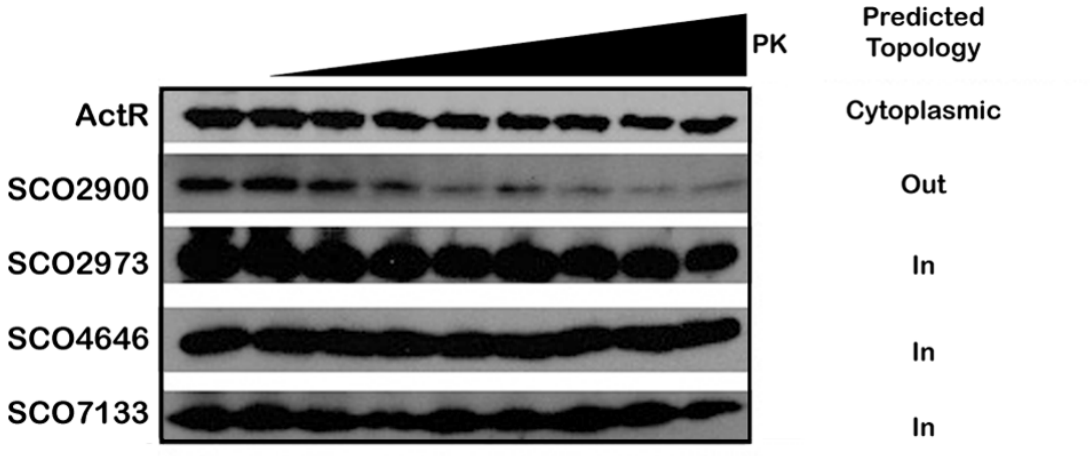

Figure A1.5. Protease protection assay to assess TAMP orientation at the membrane. Protoplasts expressing ActR, SCO2900, SCO2973, SCO4646 and SCO7133 were subjected to increasing concentration of Proteinase K (PK). Exterior facing N-termini were expected to be susceptible to Proteinase K digestion; while inward facing N-termini were expected to be protected. Visualization of the extent of degradation was detected by Western blot analysis against the FLAG epitope.

A1.4 Discussion

We have identified a previously uncharacterized class of bacterial membrane proteins in *S. coelicolor* that lack the N-terminal signal sequences and, rather, depend on C-terminal transmembrane domains for membrane targeting. This is the first time such an observation has been biochemically demonstrated in a prokaryote. Aside from their C-terminal sequences, these proteins do not appear to contain any additional sequence motif for membrane targeting as the C-termini alone from three of these proteins can render eGFP entirely membrane-associated (**Figure A1.4**). Furthermore, the remarkable diversity of the N-terminal domains of these proteins, which exhibit no universally conserved sequence

characteristics, strongly argues for a membrane targeting mechanism that depends primarily, if not entirely on, the C-terminal domains.

Among the candidates that we have worked with here are at least two known proteins of considerable interest, PkaB (SCO2973), one of the so-called "eukaryotic" Ser/Thr protein kinases found in streptomycetes and other prokaryotes, and SecE from the Sec translocon. Orthologues of these proteins have been investigated in several bacteria previously; however, to our knowledge the possible tail-anchoring is a new observation (26, 27, 28).

While little is known about PkaB in *S. coelicolor*, it is closely related to the *Mycobacterium tuberculosis* protein kinase PknA. The *pknA* $_{TB}$ gene is adjacent to a second Ser/Thr kinase gene *pknB*$_{TB}$, (unfortunately referred to as *pkaA* in *S. coelicolor)*. These two mycobacterial kinases have been implicated in cell division and the maintenance of cell shape and it has been suggested that they may phosphorylate components (FtsZ and FipA) of the division apparatus (29). The single C-terminal transmembrane domain has been previously noted; however, the absence of N-terminal signal sequences was not. The importance of PknA in *M. tuberculosis* suggests, that understanding the mechanism with which this kinase targets to the membrane could lead to new drug targets for combating this pathogenic bacteria.

SecE is similarly a highly studied and important protein. Its role in secretion is to aid in forming the protein conducting channel, the SecYEG translocase, by stabilizing SecY and by contributing residues to the active centre in the translocase (5, 30). The *E. coli* SecE protein, arguably the best studied prokaryotic example, is a 127 amino acid, Sec-dependent polytopic transmembrane protein having three transmembrane sequences. In contrast, the *S. coelicolor* orthologue that we have investigated, SCO4646, is a 79 amino acid protein having a single transmembrane domain at its C-terminus: we confirmed that this characterization is not due to a mis-identification of the open reading frame's 5' end. These results are also in agreement with recently published bioinformatic data from the SecE of *M. maripaludis* (15).

The eukaryotic orthologues of SecE, Sec61β, are also well-known tail-anchored transmembrane proteins (12). Intriguingly, our sequence searches suggest that many other prokaryotic SecE orthologues are similar to that of *S. coelicolor* in that they appear to lack signal sequences and have a single, C-terminal transmembrane domain. For example, the SecE orthologues in all the sequences streptomycetes are all predicted to be shorter proteins, similar in length to that of SCO4646, and to have a single predicted transmembrane domain at their C-terminus. Remarkably, the SecE orthologue in the very well-studied model organism *Bacillus subtilis* (NCBI locus tag NP_387981) is also a shorter protein of 59 amino acid residues with a single, C-terminal

transmembrane domain and no obvious N-terminal signal sequence. This appears also to be the case in the important pathogens *Staphylococcus aureus* (NCBI locus tag AAB54017) and *Enterococcus faecalis* (NCBI locus tag EEN75976), both of which are smaller proteins with one predicted C-terminal membrane spanning domain like that of *S. coelicolor*. The *M. tuberculosis* SecE protein is a longer protein of 161 amino acid residues however it too appears to lack a signal sequence and has a single, C-terminal transmembrane domain, unlike that of *E. coli.*

While we have identified this class of bacterial membrane proteins, the targeting apparatus and mechanism remains unknown. We have demonstrated that the C-terminal transmembrane domain is sufficient for localization; suggesting a targeting pathway that is independent of the SRP. Recent bioinformatics suggests that *Archea* and eukaryote TAMPs target via a similar set of machinery, the archeon ArsA and eukaryote equivalent the Get3 complex; however, the bacterial equivalent lacks key residues for membrane protein targeting, suggesting bacterial ArsA is not the TAMP targeting machinery (15). The eukaryotic TAMPs are all predicted to insert their transmembrane domains into the membrane and sit with their N-termini facing the cytosol from either the mitochondrial outer membrane or the endoplasmic reticulum (N-terminus “in”) (31). An interesting development in our analysis is the variation in predicted topology of the TAMPs from *S. coelicolor* (**Table A1.1**) with some facing the

cytosol (N-terminus "in") and some exterior to the cell (N-terminus "out"). Preliminary biochemical evidence has confirmed these 2 bacterial orientations (**Figure A1.5**). This requirement to cross the lipid bilayer may be the reason for a differing targeting mechanism for bacterial TAMPs in contrast to archeon and eukaryotes as translocation machinery may be required for proper translocation across the membrane. Based on known membrane targeting machinery, bacterial possibilities could include YidC as YidC is capable of targeting membrane proteins independent of the Sec translocon (8). It has been previously reported that YidC alone is capable of inserting *E. coli* SecE, a SecE with multiple TMs, into the membrane (32). Despite this possibility, YidC targeting of bacterial TAMPs has yet to be explored. The identification of a new targeting pathway could pose as an important target for an antimicrobial agent, especially in light of a potentially differing targeting pathway from eukaryotes.

A1.5 Material & Methods

A1.5.1 Bioinformatics. The TAMP finder program was used as previously reported (Brito et al, Manuscript in Preparation). Transmembrane domains were detected using the TMHMM software available at http://www.cbs.dtu.dk/services/TMHMM/ and signal sequences were assessed using the SignalP software available online at http://www.cbs.dtu.dk/services/SignalP/. Hydrophilicity and hydrophobicity of the

N-terminal regions was calculated using the online program http://www.innovagen.se/custom-peptide-synthesis/peptide-property-calculator/peptide-property-calculator.asp.

A1.5.2 Strains, plasmids and general growth conditions. *E. coli* strains were grown at 37°C in Luria broth medium. Plasmid construction was performed in *E. coli* strain XL1 blue (Stratagene); while *E. coli* strain ET12567 containing the pUZ8002 plasmid was used for conjugal transfer of plasmids into *S. coelicolor* (33). *S. coelicolor* M145 was used to test the membrane protein predictions. *Streptomyces* strains were grown at 30°C on MS agar for matings and R2YE for general restreaking. Liquid cultures of *S. coelicolor* were grown in R5 medium supplemented with 7% PEG-8000 (34). Antibiotic concentrations were 50 µg/ml kanamycin, 50 µg/ml apramycin, 35 µg/ml chloramphenicol, 30 µg/ml thiostrepton and 25 µg/ml nalidixic acid.

A1.5.3 Construction of TAMP over-expression vectors. Putative membrane proteins SCO2900, SCO2973, SCO4008, SCO4646 and SCO7133 were amplified from *S. coelicolor* chromosomal DNA via PCR introducing a FLAG epitope (DYKDDDDK) at their N-termini for western blot analysis. NdeI and BamHI restriction sites were introduced upstream and downstream of the genes, respectively, to allow for introduction into the *Streptomyces* over-expression vectors pIJ6902 and pIJ8600 (35, 36). The cytoplasmic protein ActR was amplified in the same manner and introduced into pIJ6902.

A1.5.4 Construction of eGFP-tail anchor fusions. The eGFP gene was amplified from the plasmid pIJ8668, removing the stop codon and introducing an XbaI restriction site downstream for introduction into pIJ6902. The eGFP gene was also cloned in a similar manner but containing the stop codon for use as a cytoplasmic control. The putative tail anchor transmembrane domains from SCO2973, SCO4646 and SCO7133 were amplified including 11 residues upstream from the predicted transmembrane domain via PCR, introducing XbaI and BamHI for introduction downstream of the eGFP gene.

A1.5.5 Separation of membrane and cytoplasmic fractions. *S. coelicolor* strains containing the TAMP over-expression vectors and the eGFP-tail anchor fusions were grown in liquid culture for 16 hours prior to induction. Cultures were induced for 45 min with 30 µg/ml thiostrepton. Cells were washed once with 10.3% sucrose and resupended in P buffer containing 2 mg/ml lysozyme (34). Protoplasts were created by incubation at 30^{o}C for 1 hour and harvested by filtering through cotton and centrifugation at 7,000 xg for 10 min (34). The pellet was resuspended in lysis buffer (150 mM Hepes pH 7.3, 150 mM NaCl, 3 mM DTT, 30% glycerol) with protease inhibitor cocktail. Subsequent steps were all performed at 4^{o}C. Protoplasts were sonicated for 2 min at 5 sec intervals following 10 sec rest. The lysate was centrifuged at 7,000 xg for 10 min and the supernatant was centrifuged at 100,000 xg for 1 hr.

A1.5.6 Sucrose gradient ultracentrifugation. Fifty micrograms of total protein from the membrane fractions were loaded to the top of sucrose step gradients containing 60% sucrose (4 ml Tris pH 8), 40% sucrose (4 ml Tris pH 8) and 20% sucrose (3 ml Tris pH 8). Gradients were centrifuged at 100,000 xg for 16 hr at 4^{o}C and 1 ml fractions were collected by piercing a needle in the bottom of the centrifuge tube and collecting the flow through.

A1.5.7 Sodium carbonate extraction. Cell lysate was prepared as described above. The lysate was mixed on ice with an equal volume of 0.2 M sodium carbonate (pH 11.4) or 0.2 M sucrose (pH 7.8) and centrifuged at 4^{o}C for 1 hr. Following centrifugation at 100,000 xg, the supernatant was neutralized with glacial acetic acid and the pellet was resuspended in lysis buffer. The peripheral membrane protein RamC was used as a control for extraction by sodium carbonate. Preparation of lysate for this analysis was performed as previously described (24).

A1.5.8 Visualization of subcellular localization. Western blot analysis was used to determine the localization patterns of the TAMPs and eGFP-tail anchor fusions. For detection of the TAMPs anti-FLAG (Sigma) was used at a concentration of 1:10,000 and anti-eGFP (Invitrogen) was used at a concentration of 1:2,500 for eGFP fusion proteins. Antibodies against RamC were used at a concentration of 1:1,000.

A1.5.9 Proteinase K digestion. A concentration range of 0, 1, 10, 15, 20, 25, 50, 75, 100 µg/µl Proteinase K (Sigma) was added to protoplasts and incubated on ice for 10 minutes. Proteolysis was stopped by the addition of protease inhibitors (Roche), followed by the addition of 3x SDS loading buffer and heating to 95^{o}C for 10 minutes. The degree of Proteinase K digestion was visualized by western blot analysis using the anti-FLAG antibody (Sigma).

A1.6 References

1. Dong H, Jiang J & Li Y (2009) The distinct anchoring mechanism of FtsY from different microbes. *Curr Microbiol* 59: 336-340.
2. Valent QA (2001) Signal recognition particle mediated protein targeting in *Escherichia coli*. *Antonie van Leeuwenhoek* 79: 17-31.
3. Stroud RM & Walter P (1999) Signal sequence recognition and protein targeting. *Curr Opin Struct Biol* 9: 754-759.
4. Shen K & Shan S (2010) Transient tether between the SRP RNA and SRP receptor ensures efficient cargo delivery during cotranslational protein targeting. *Proc Natl Acad Sci* 107: 7698-7703.
5. Papanikou E, Karamanou S & Economou A (2007) Bacterial protein secretion through the translocase nanomachine. *Nat Rev Microbiol* 5: 839-851.
6. du Plessis DJ, Nouwen N & Driessen AJ (2011) The Sec translocase. *Biochim Biophys Acta* 1808: 851-65.
7. Eichler J & Duong F (2010) Break on through to the other side – the Sec translocon. *Trends Biochem Sci* 29: 3-5.
8. Kol S, Nouwen N & Driessen AJ (2008) Mechanisms of YidC-mediated insertion and assembly of multimeric membrane protein complexes. *J Biol Chem* 283: 31269-31273.
9. Gold VA, Duong F & Collinson I (2007) Structure and function of the bacterial Sec translocon. *Mol Membr Biol* 24: 387-394.
10. Rusch SL & Kendall DA (2007) Oligomeric states of the SecA and SecYEG core components of the bacterial Sec translocon. *Biochim Biophysic Acta* 1768: 5 – 12.

11. High S & Abell BM (2004) Tail-anchored protein biosynthesis at the endoplasmic reticulum: the same but different. *Biochem Soc Trans* 32: 659-662.
12. Favaloro V, Spasic M, Schappach B & Dobberstein B (2008) Distinct targeting pathways for the membrane insertion of tail-anchored (TA) proteins. *J Cell Sci* 121: 1832-1840.
13. Borgese N, Brambillasca S & Colombo S (2007) How tails guide tail-anchored proteins to their destinations. *Curr Opin Cell Biol* 19: 368-375.
14. Wattenberg BW, Clark D & Brock S (2007) An artificial mitochondrial tail signal/anchor sequence confirms a requirement for moderate hydrophobicity for targeting. *Biosci Rep* 27: 385-401.
15. Borgese N & Righi M (2010) Remote origins of tail-anchored proteins. *Traffic* 11: 877-885.
16. Krogh A, Larsson B, von Heijne G & Sonnhammer EL (2001) Predicting transmembrane protein topology with a Hidden Markov Model : Application to complete genomes. *J Mol Biol* 305: 567-580.
17. Bendtsen JD, Nielsen H, von Heijne G & Brunak S (2004) Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* 340: 783-795.
18. Natale P, Brüser T & Driessen AJ (2008) Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane — Distinct translocases and mechanisms. *Biochim Biophysic Acta* 1778: 1735 – 1756.
19. Niegoawski D & Eshaghi S (2007) The CorA family: Structure and function revisited. *Cell Mol Life Sci* 64: 2564 – 2574.
20. Puttikhunt C, Nihira T & Yamada Y (1995) Cloning, nucleotide sequence, and transcriptional analysis of the *nusG* gene of *Streptomyces coelicolor* A3(2), which encodes a putative transcriptional antiterminator. *Mol Gen Genet* 247: 118-122.
21. Hagege J, *et al* (1993) Transfer Functions of the Conjugative Integrating Element pSAM2 from *Streptomyces ambofaciens*: Characterization of a kil-kor System Associated with Transfer. *J Bact* 175: 5529-5538.
22. Kim B, Andersen C & Benz R (2001) Identification of a cell wall channel of *Streptomyces griseus*: the channel contains a binding site for streptomycin. *Mol Micro* 41: 665-673.
23. Molloy MP *et al* (2000) Proteomic analysis of the *Escherichia coli* outer membrane. *Eur J Biochem* 267: 2871-2881.

24. Hudson ME, Zhang D & Nodwell JR (2002) Membrane association and kinase-like motifs of the RamC protein of *Streptomyces coelicolor*. *J Bact* 184: 4920-4924.

25. von Heijne J (1992) Membrane protein structure prediction. Hydrophobicity analysis and the positive-inside rule. *J Mol Biol* 225: 487-94.

26. Kontinen VP, Yamanaka M, Nishiyama K & Tokuda H (1996) Roles of the conserved cytoplasmic region and non-conserved carboxy-terminal region of SecE in *Escherichia coli* protein. *J Biochem* 1130: 1124-1130.

27. Nishiyama K & Mizushima S (1992) The carboxyl-terminal region of SecE interacts with SecY and is functional in the reconstitution of protein translocation activity. *Biochem* 267: 7170-7176.

28. Jeong SM, Yoshikawa H & Takahashi H (1993) Isolation and characterization of the *secE* homologue gene of *Bacillus subtilis*. *Mol Microbiol* 10: 133-42.

29. Sureka K, *et al*. (2010) Novel role of phosphorylation-dependent interaction between FtsZ and FipA in mycobacterial cell division. *PLOS one* 5: e8590.

30. Veenendaal AK, Does CV & Driessen AJ (2004) The protein-conducting channel SecYEG. *Biochim Biophys Acta* 1694: 81 – 95.

31. Borgese N, Colombo S & Pedrazzini E (2003) The tale of tail-anchored proteins: coming from the cytosol and looking for a membrane. *J Cell Biol* 161: 1013-1019.

32. Yi L, *et al*. (2003) YidC is strictly required for membrane insertion of subunits a and c of the F(1)F(0)ATP synthase and SecE of the SecYEG translocase. *Biochem* 42: 10537-44.

33. Flett F, Mersinias V & Smith P (1997) High efficiency intergeneric conjugal transfer of plasmid DNA from *Escherichia coli* to methyl DNA-restricting streptomycetes. *FEMS Microbiol Lett* 155: 223-229.

34. Keiser T, *et al* (2000) Practical Streptomyces Genetics. The John Innes Foundation, Norwich.

35. Huang J, Shi J, Molle V, Sohlberg B, Weaver D, et al. (2005) Cross-regulation among disparate antibiotic biosynthetic pathways of *Streptomyces coelicolor. Mol Microbiol* 58: 1276-1287.

36. Sun J, Kelemen GH, Fernandex-Abalos JM & Bibb MJ (1999) Green fluorescent protein as a reporter for spatial and temporal gene expression in *Streptomyces coelicolor* A3(2). *Microbiol* 3: 2221-2227.

APPENDIX A2

TRANSCRIPTIONAL ANALYSIS OF THE ARC2 SERIES ON *S. COELICOLOR*

Author contributions:

This work was performed in collaboration with Vanessa Yoon. I performed the initial microarray experiment and created the *lux* fusion library in *S. coelicolor* for microarray follow up. Vanessa performed the follow up analysis of the *lux* fusions under my supervision.

A2.1 Transcriptional profiling of *S. coelicolor* in the presence of ARC2

To further explore the mechanism by which ARC2 addition leads to the increased production of actinorhodin, *S. coelicolor* was subjected to microarray analysis in the presence and absence of ARC2. Growth and actinorhodin production were monitored for 72 hours to ensure consistent growth between the cultures and strong induction of actinorhodin production in the presence of ARC2 (**Figure A2.1**). RNA was extracted at the time of ARC2 addition and at 1, 2, 4 and 6 hours following addition. Only changes in gene expression at 6 hours post ARC2 addition are reported as this time point produced the strongest changes compared to the DMSO control.

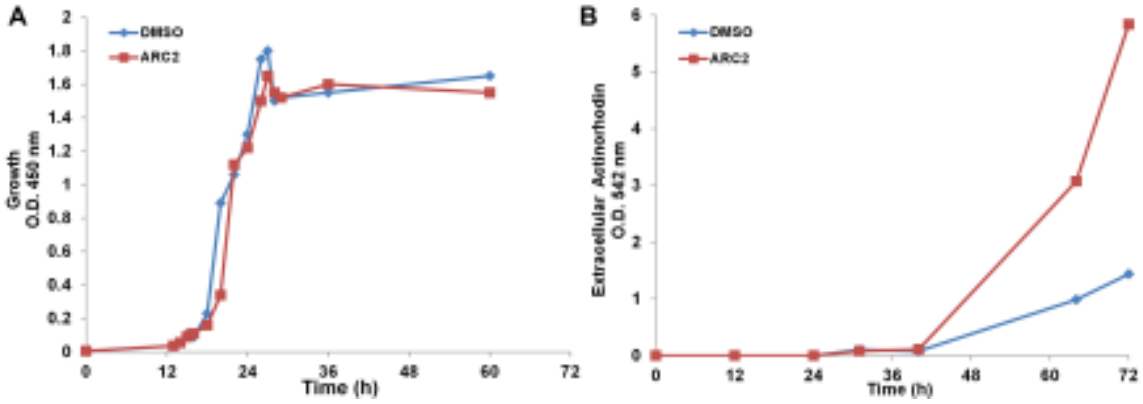

Figure A2.1. Growth and actinorhodin production in the presence of ARC2. A culture of *S. coelicolor* M145 was grown in liquid R5M for 12 hours and split. One received DMSO and the other 25 µM ARC2. RNA was extracted for microarray analysis at the time of subculture (T=12hr), 1, 2, 4, and 6 hours post addition (corresponding to 18 hours culture growth). (**A**) Growth was monitored in the presence and absence of ARC2. (**B**) Extracellular actinorhodin production was measured over 72 hours.

Alterations in secondary metabolism by ARC2 addition

Changes in secondary metabolite gene expression are shown in **Table A2.1** (induced) and **Table A2.2** (repressed). Expression of the pathway specific regulator for actinorhodin was induced 1.5 fold and is consistent with the observed increased yield of actinorhodin. This suggests that at least part of the increase in actinorhodin yield is through increased expression of *act*II-4. Increased expression of SCO7221 the polyketide synthase for germicidin production also correlates with observed increases in its yield. Increased transcription was also observed for the biosynthetic genes involved in CDA production. This is in contrast to the observed decrease in CDA yields, suggesting the decreased production of CDA does not result from decreased transcription of the biosynthetic cluster. Decreased expression of prodiginine biosynthesis is consistent with observed yields in the presence of ARC2. The remaining metabolites demonstrating reduced gene expression (SCB1, coelichelin, desferrioxamine and albaflavenone) were not detected by LC-MS analysis.

Table A2.1. Upregulated Biosynthetic Clusters in the Presence of ARC2 (6 hrs post addition)

SCO	Annotation	Fold Difference
Deoxy sugar – 2 genes also downregulated		
SCO0392	SCF62.18, possible methyltransferase, len: 416 aa.	1.5
CDA		
SCO3211	SCE8.04c, trpC2, probable indoleglycerol phosphate synthase, len: 258aa	2.8
SCO3214	SCE8.07c, trpE2, anthranilate synthase component I, len: 511aa	6.3
SCO3215	SCE8.08c, unknown, len: 338 aa	4.5
SCO3217	SCE8.10, cdaR, transcriptional activator protein, len: 638 aa;	6.4
SCO3222	SCE8.15c, possible small secreted protein, len: 169 aa	4.5
SCO3227	SCE8.20c, partial CDS, possible aminotransferase, len: 431 aa	4.0
SCO3229	SCE63.04, probable 4-hydroxyphenylpyruvic acid dioxygenase, len: 371 aa	8.8
SCO3230	SCE63.03c, cdaPSI, CDA peptide synthetase I, len: 7463 aa	2.1
SCO3231	SCE63.02c, cdaPS2, CDA peptide synthetase II, len: 3670 aa	3.8
SCO3232	SCE63.01, partial CDS, cdaPS3, CDA peptide synthetase III, len: >332 aa	3.9
SCO3233	SCE29.02, probable hydrolase, len: 272 aa	2.6
SCO3234	SCE29.03, possible phosphotransferase, len: 300 aa	2.6
SCO3235	SCE29.04c, probable ABC transporter, len: 615aa	3.8
SCO3236	SCE29.05c, possible oxygenase, len: 333 aa	6.0
SCO3239	SCE29.08c, unknown, len: 289 aa	2.1
SCO3241	SCE29.10c, possible isomerase, len: 290 aa	2.0
SCO3243	SCE29.12c, possible myo-inositol phosphate synthase, len: 388 aa	5.5
SCO3244	SCE29.13c, secreted protein, len: 265 aa	5.3
SCO3245	SCE29.14c, possible salicylate hydroxylase (putative secreted), len: 420 aa	2.6
SCO3246	beta-ketoacyl-ACP synthase III;	2.6
SCO3249	SCE29.18c, probable acyl carrier protein, len: 81 aa	2.1
Actinorhodin		
SCO5085	SCBAC28G1.11, actII-4, actinorhodin cluster activator protein, len: 255 aa;	1.5
One gene in hopene 6759-6771		
SCO6766	SC6A5.15, unknown, len: 340 aa;	2.5
Germicidin		
SCO7221	polyketide synthase	2.3

Table A2.2. Down Regulated Biosynthetic Clusters in the Presence of ARC2 (6 hrs post addition)

SCO	Annotation	Fold Difference
Deoxysugar		
SCO0382	SCF62.08, probable UDP-glucose/GDP-mannose family dehydrogenase	-2.5
SCO0383	SCF62.09, unknown, len: 407 aa.	-3.2
Coelichelin		
SCO0489	SCF34.08c, conserved hypothetical protein	-5.7
SCO0491	SCF34.10c, probable ABC-transporter transmembrane protein	-1.7
SCO0492	SCF34.11c, probable peptide synthetase, len: 3643 aa	-11.5
SCO0494	SCF34.13c, probable iron-siderophore binding lipoprotein, len: 350 aa	-3.4
SCO0496	SCF34.15c, probable iron-siderophore permease transmembrane protein	-3.1
SCO0498	SCF34.17c, probable peptide monooxygenase, len: 451 aa	-34.9
SCO0499	SCF34.18, possible formyltransferase, len: 315 aa	-15.2
Desferrioxamine		
SCO2782	SCC105.13, probable pyridoxal-dependent decarboxylase	-4.4
SCO2783	SCC105.14, probable monooxygenase	-4.3
SCO2784	SCC105.15, probable aceytltranferase	-1.6
One gene from CDA		
SCO3218	SCE8.11c, small conserved hypothetical protein, len: 71aa	-4.9
Albaflavenone		
SCO5222	SC7E4.19, possible lyase, len: 361 aa	-1.6
Prodiginines		
SCO5877	SC2E9.18, redD, transcriptional regulator, len: 350 aa	-23.9
SCO5878	SC2E9.19, redX, polyketide synthase with two beta-ketoacyl synthase	-30.5
SCO5879	SC2E9.20, redW, acyl-coa dehydrogenase	-7.3
SCO5880	SC2E9.21, redY, unknown function	-8.5
SCO5883	SC3F7.03c, redU, unknown, len	-2.5
SCO5884	SC3F7.04c, unknown, len: 296 a	-5.8
SCO5885	SC3F7.05c, membrane protein, len: 146 aa	-6.7
SCO5886	SC3F7.06c, redR, probable 3-oxoacyl-[acyl-carrier-protein] synthase II	-7.8
SCO5887	SC3F7.07c, redQ, probable acyl carrier protein	-63.9
SCO5888	beta-ketoacyl-ACP synthase III	-14.2
SCO5889	SC3F7.09, redO, unknown	-15.1
SCO5890	SC3F7.10, redN, 8-amino-7-oxononanoate synthase	-51.6
SCO5891	St3F7.11, redM, probable peptide synthase	-20.8
SCO5892	SC3F7.12, redL, probable polyketide synthase	-11.8
SCO5893	SC3F7.13, redK, probable oxidoreductase	-9.1
SCO5894	SC3F7.14, redJ, probable thioesterase	-2.7
SCO5895	SC3F7.15, redI, possible methyltransferase	-18.9
SCO5896	SC10A5.01, redH, probable phosphoenolpyruvate-utilizing enzyme	-26.0
SCO5897	SC10A5.02, redG, probable oxidase	-37.7
SCO5898	SC10A5.03, redF, probable membrane protein	-22.5
GBL		
SCO6266	SCAH10.31, scbA	-13.1
Hexaketide 6273-6288		
SCO6282	SC1G7.08c, probable 3-oxoacyl-[acyl-carrier protein] reductase	-9.5

Genes with increased expression in the presence of ARC2

64 genes showed increased expression in the presence of ARC2 with at least 1.5 fold increases compared to the DMSO control (**Table A2.3**). Genes chosen for follow up confirmation are marked with asterisks. These include 4 regulatory proteins, genes responsible for the reduction of nitrate to nitrate, superoxide dismutase and glycerol uptake. Of particular interest is the increased expression of AfsS (AfsR2) as this regulator is known to stimulate actinorhodin production (1).

Table A2.3. Genes induced in the presence of ARC2 (6 hrs post addition)

SCO	Annotation	Fold Difference
Regulators		
SCO0168*	SCJ1.17, possible regulator, len: 190 aa, similar to M. tuberculosis	6.0
SCO0174*	SCJ1.23c, possible DNA-binding protein, len: 323 aa.	1.9
SCO0204*	SCJ12.16c, probable luxR family two-component RR, len: 233 aa.	2.6
SCO4156	SCD84.23c, probable two-component systen response regulator, len: 248 aa;	3.1
SCO4159	SCD84.26c, glnR, transcriptional regulatory protein, len: 267 aa;	2.2
SCO4198	2SCD46.12, probable DNA-binding protein, len: 141 aa;	5.5
SCO4425*	SCD6.03c, afsR2, regulator of antibiotic production, len: 63 aa;	3.0
SCO7014	SC1H10.03c, probable LacI-family transcriptional protein, len: 337 aa	2.7
FAS/Acetyl-CoA		
SCO5385*	2SC6G5.29, possible 3-hydroxybutyryl-coA dehydrogenase, len: 282 aa;	3.7
SCO5399	Catalyzes the synthesis of acetoacetyl coenzyme A	2.0
SCO4800	SCD63A.11, icmB, isobutyryl CoA mutase, small subunit, len: 138 aa;	1.6
Nitrate → Nitrite		
SCO0212	SCJ12.24c, unknown, len: 157 aa.	3.8
SCO0213	SCJ12.25c, possible nitrate/nitrite transporter protein, len: 412 aa.	14.4
SCO0216	SCJ12.28, narG2, probable nitrate reductase alpha chain, len: 1233 aa.	13.4
SCO0217	SCJ12.29, narH2, probable nitrate reductase beta chain, len: 522 aa.	13.1
SCO0218	SCJ12.30, narJ2, probable nitrate reductase delta chain, len: 267 aa.	4.2
SCO0219	SCJ12.31, narI2, possible nitrate reductase gamma chain, len: 256 aa.	1.8
Glycerol		
SCO1659	SCI52.01, glpF, probable glycerol uptake facilitator protein, len: 264 aa	5.7
SCO1660	SCI52.02, glpK, glycerol kinase, len: 512 aa	6.0
SCO1661	SCI52.03, probable glycerol-3-phosphate dehydrogenase, len: 538 aa;	6.5
SCO1662	SCI52.04, conserved hypothetical protein, len: 333 aa;	3.6
Oxidative Stress		
SCO2633	SC8E4A.03, sodF, superoxide dismutase [Fe-Zn] (EC 1.15.1.1), len: 213 aa	12.0

Table A2.3 Con'd

Membrane Proteins		
SCO2164	SC5F7.37c, possible integral membrane efflux protein, len: >822aa	3.9
SCO3110	SCE41.19c, possible ABC transport system integral membrane, len: 854 aa	3.2
SCO3111	SCE41.20c, possible ABC transport system ATP-binding protein, len: 262 aa	4.0
SCO3999	2SC10A7.03, possible lipoprotein, len: 376 aa	1.9
SCO4054	2SCD60.20, possible integral membrane protein, len: 228 aa.	1.5
SCO4584	SCD20.02, possible membrane protein, len: 420 aa.	1.6
SCO4852	SC5G8.20c, possible integral membrane protein, len: 135 aa.	3.2
SCO4903	2SCK8.29c, possible membrane protein, len: 67aa;	2.2
SCO4963	2SCK31.23, possible ABC transporter ATP-binding protein, len: 335 aa;	2.0
SCO5578	SC7A1.22, sugar transporter, len: aa	4.0
SCO6091	SCBAC1A6.15c, probable integral membrane protein, len: 756 aa	6.9
SCO7153	SC9A4.15, probable sugar transporter, len: 472 aa;	3.4
SCO7536	SC8G12.12, possible integral membrane protein, len: 739 aa.	5.6
Hypothetical		
SCO0169	SCJ1.18, conserved hypothetical protein, len: 217 aa. Duplication region	9.0
SCO0200	SCJ12.12c, unknown, len: 301 aa. Similar to several hypothetical	9.7
SCO2912	SCE19A.12c, hypothetical protein, len: 47 aa	1.8
SCO3929	SCQ11.12c, hypothetical protein, len: 450 aa	1.6
SCO4199	2SCD46.13, unknown, len: 213aa; contains mostly hydrophilic residues.	4.0
SCO5252	2SC7G11.14, hypothetical protein, len: 194 aa	1.7
SCO5264	2SC7G11.26c, unknown, len: 360 aa.	2.4
SCO5265	2SC7G11.27c, hypothetical protein, len: 333 aa;	2.7
SCO5398	SC8F4.02c, conserved hypothetical protein, len: 146aa;	2.9
SCO7467	SCBAC14E8.07, unknown, len: 375 aa	3.8
Secreted		
SCO3286	SCE15.03, secreted protein, len: 677 aa	2.4
SCO7717	SC8D11.08c, possible secreted protein, len: 137 aa;	6.6
Miscellaneous Enzymes		
SCO0179	SCJ1.28c, possible zinc-containing dehydrogenase, len: 356 aa.	2.6
SCO0199	SCJ12.11c, probable alcohol dehydrogenase, len: 340 aa.	9.9
SCO0201	SCJ12.13c, possible integral membrane protein, len: 223 aa.	17.3
SCO0985	catalyzes the transfer of a methyl group from 5-methyltetrahydrofolate	1.6
SCO1487	catalyzes the transfer of the carbamoyl moiety from carbamoyl phosphate	3.3
SCO1706	SCI30A.27c, probable aldehyde dehydrogenase, len: 462 aa	2.6
SCO3127	SCE66.06c, ppc, phosphoenolpyruvate carboxylase, len: 911 aa	2.9
SCO4157	SCD84.24c, probable protease, len: 347 aa;	3.7
SCO4683	converts 2-oxoglutarate to glutamate;	2.7
SCO4814	involved in de novo purine biosynthesis	3.1
SCO5515	SC8D9.27, serA, D-3-phosphoglycerate dehydrogenase, len: 529 aa;	3.2
SCO5522	catalyzes the oxidation of 3-isopropylmalate to	2.2
SCO6097	SCBAC1A6.21c, cysN, sulfate adenylyltransferase subunit 1, len: 451 aa;	1.6
SCO6199	SC2G5.20, possible secreted esterase, len: 743aa;	3.5
SCO6663	catalyzes the formation of ribose 5-phosphate and xylulose 5-phosphate	2.0
SCO7468	SCBAC17A6.01c, possible flavin-containing monooxygenase, len: >327aa	2.2
SCO7638	SC10F4.11c, eno2, enolase, len: 434 aa;	3.6

Genes whose expression decreased by ARC2 addition

132 genes showed decreased expression in the presence of ARC2 (**Table A2.4**). Genes chosen for follow up analysis are marked with asterisks. These include NADH dehydrogenase, γ-butyrolactone production (*scbA*), and genes involved in acetyl-CoA production from branched chain amino acid degradation. Branched chain amino acid degradation is a major supplier of acetyl-CoA for actinorhodin production (2). The decreased expression of genes involved in acetyl-CoA production suggests that there is a buildup in these metabolites within the cell and further production is unnecessary. However, this buildup in acetyl-CoA needs to be confirmed experimentally.

Table A2.4. Genes repressed by the presence of ARC2 (6 hrs post addition)

SCO	Annotation	Fold Difference
Regulators		
SCO2426	SCC42.07, possible regulatory protein	-2.2
SCO2517	SCC121.20c, possible two-component system response regulator (ecrA1/A2)	-3.7
SCO2518	SCC121.21c, possible two-component sensor kinase (ecrA1/A2)	-15.1
SCO3579	SCH17.13c, probable transcriptional regulator, len: 112 a	-1.9
SCO3763	SCH22A.14c, probable RNA polymerase ECF sigma factor	-3.7
SCO6008	SC7B7.05, probable transcriptional repressor protein	-2.6
Acetyl-CoA Production		
SCO1345	3-ketoacyl-(ACP) reductase (FabG2)	-2.3
SCO1346*	3-oxoacyl-(ACP) reductase (FabG3)	-3.2
SCO1393	SC1A8A.13, acsA, acetoacetyl-CoA synthetase, len: 658 aa	-1.7
SCO1428	SC6D7.11, acd, acyl-CoA dehydrogenase (EC 1.3.99.-), len: 391 aa	-4.5
SCO1198	SCG11A.29c, probable acyl-CoA dehydrogenase, len:393 aa	-3.8
SCO1199	SCG11A.30c, probable oxidoreductase	-1.7
SCO2008	SC7H2.22, possible branched chain amino acid binding protein,	-40.5
SCO2009	SC7H2.23, probable branched chain amino acid transport permease,	-58.5
SCO2010	SC7H2.24, probable branched chain amino acid transport permease	-31.4
SCO2011	SC7H2.25, probable branched chain amino acid transport ATP-binding pro	-36.7
SCO2012	SC7H2.26, probable branched chain amino acid transport ATP-binding pro	-18.4
SCO2774	SCC105.05c, acdH2, probable acyl-CoA dehydrogenase, len: 385 aa	-1.9
SCO2776	SCC105.07, accD1, acetyl/propionyl CoA carboxylase, beta subunit	-29.0
SCO2777	SCC105.08, accC, acetyl/propionyl CoA carboxylase alpha subunit	-11.4
SCO2778	SCC105.09, hmgL, hydroxymethylglutaryl-CoA lyase	-12.1
SCO2779	SCC105.10, acdH, acyl-CoA dehydrogenase	-26.0
SCO3247	SCE29.16c, possible acyl CoA oxidase, len: 600 aa	-2.2
SCO3563	forms acetyl-CoA from acetate and coenzymeA	-2.7
SCO3829	SCGD3.30c, bkdC2, probable dihydrolipoamide acyltransferase component	-2.0
SCO3830	SCGD3.31c, bkdB2, probable branched-chain alpha keto acid dehydrogenase	-2.0
SCO3831*	SCGD3.32c, bkdA2, probable branched-chain alpha keto acid dehydrogenase	-3.5

Table A2.4 Con't…

SCO4006	Activates fatty acids by binding to coenzyme A	-4.3
SCO5415	SC8F4.19, icmA, isobutyryl-CoA mutase A, len: 566 aa	-1.9
SCO6195	SC2G5.16, probable acetyl-coenzyme A synthetase	-8.7
SCO6196	Activates fatty acids by binding to coenzyme A	-9.5
SCO6731	acetyl-CoA acetyltransferase	-5.1
SCO6732	SC5F2A.15, possible fatty acid oxidative multifunctional enzyme	-2.3
SCO6788	acetyl-CoA acetyltransferase	-6.0
SCO7469	SCBAC17A6.02c, phenylacetate-CoA ligase	-13.1
SCO7470	SCBAC17A6.03c, possible phenylacetic acid degradation protein PaaI	-6.1
SCO7471*	SCBAC17A6.04, possible phenylacetic acid degradation protein PaaA	-15.6
SCO7472	SCBAC17A6.05, possible phenylacetic acid degradation protein PaaB	-4.5
SCO7473	SCBAC17A6.06, possible phenylacetic acid degradation protein PaaC	-2.0
SCO7474	SCBAC17A6.07, possible phenylacetic acid degradation protein PaaD	-3.2
SCO7475	SCBAC17A6.08, possible phenylacetic acid degradation NADH oxidored PaaE	-4.6
Glutamine → glutamate		
SCO4047	phosphoribosylformylglycinamidine synthase	-3.4
SCO4048	phosphoribosyl formylglycinamidine synthase I	-4.0
SCO5774	SC4H8.13c, gluD, probable glutamate permease	-2.0
SCO5775	SC4H8.14c, gluC, probable glutamate permease	-1.8
SCO5776	SC4H8.15c, gluB, probable glutamate binding protein	-1.8
SCO5777	SC4H8.16c, gluA, probable glutamate uptake system ATP-binding protein	-2.1
NADH Dehydrogenase		
SCO4562	NADH dehydrogenase alpha subunit	-9.0
SCO4563	NADH dehydrogenase beta subunit	-5.0
SCO4564	NADH dehydrogenase subunit C	-3.5
SCO4565	NADH dehydrogenase delta subunit	-10.6
SCO4566	ATP synthase subunit E	-4.9
SCO4567	NuoF, NADH dehydrogenase subunit	-4.8
SCO4568	NADH dehydrogenase gamma subunit	-4.3
SCO4569	NADH dehydrogenase subunit H	-2.5
SCO4974	NADH dehydrogenase subunit M	-2.1
SCO4979	phosphoenolpyruvate carboxykinase (GTP)	-3.3
Secretions		
SCO0297	SC5G9.06, possible secreted protein, len: 425 aa;	-37.0
SCO0974	SCBAC19F3.01c, possible secreted protein, len: >237aa	-4.8
SCO1105	2SCG4.21c, possible secreted protein, len: 500 aa	-3.7
SCO2383	SC4A7.11, possible secreted protein, len: 1545 aa	-10.1
SCO2527	SCC42.08c, probable secreted arabinase	-1.9
SCO2795	2SCC13.03, probable sugar binding secreted protein	-2.6
SCO2920	SCE19A.20c, probable secreted protease	-11.0
SCO2921	SCE19A.21, membrane protein, len: 99 aa	-1.7
SCO2978	SCE50.06, possible secreted protein, len: 465 aa	-1.9
SCO5259	2SC7G11.21c, atrB, permease, len	-1.9
SCO5260	2SC7G11.22c, atrA, secreted protein	-1.7
SCO6108	SCBAC1A6.32, fusH, esterase, len: 520 aa	-8.2
SCO6109	SCBAC1A6.33, probable secreted hydrolase	-3.0
SCO1824	SCI8.09, ssp, secreted subtilisin-like protease, len: 512 aa	-1.8
SCO1230	2SCG1.05c, possible secreted tripeptidylaminopeptidase, len: 541 aa	-4.4
SCO5009	SCK15.11, probable secretory protein, len: 445 aa	-1.7
SCO7015	SC1H10.04c, possible secreted glycosyl hydrolase	-3.8
*SCO7014 – LacI repressor – on induced list		
SCO6197	SC2G5.18c, possible secreted protein	-10.1
SCO7037	SC4G1.03, possible secreted protein	-2.9
SCO7233	SC2H12.32, possible secreted protein	-1.7

Table A2.4 Con't…

Membrane Proteins		
SCO1899	SCI7.17, possible integral membrane sugar transport protein	-3.2
SCO1900	SCI7.18, possible integral membrane sugar transport protein	-1.8
SCO2404	SC4A7.32, probable sugar-binding receptor	-2.5
SCO2405	SC4A7.24, probable sugar-transport ATP binding protein	-2.2
SCO7011	SC8F11.37c, possible membrane transport protein	-23.0
SCO7012	SC1H10.01c, possible binding protein dependent transport protein	-44.5
SCO7013	SC1H10.02C, possible sugar-binding lipoprotein	-27.2
SCO7028	SC1H10.17, possible sugar-binding lipoprotein	-1.7
SCO6979	SC8F11.05, probable solute-binding lipoprotein	-7.3
SCO6980	SC8F11.06, probable ABC transport protein	-14.1
SCO6981	ABC transporter ATP-binding protein	-6.5
SCO6005	SC7B7.02, possible lipoprotein	-19.5
SCO6006	SC7b7.03, probable integral membrane transport protein	-9.4
SCO6007	SC7b7.04, probable integral membrane transport protein	-9.2
SCO6009	SC7B7.06, probable solute-binding protein of transmembrane transport system	-1.9
SCO5458	SC3D11.15, possible lipoprotein, len: 400aa	-1.8
SCO5667	SC8B7.11c, probable ABC-transporter polyamine-binding lipoprotein	-2.5
SCO5668	SC8B7.10c, probable polyamine ABC-transporter ATP-bi nding protein	-1.9
SCO5670	SC8B7.08c, probable polyamine ABC-transporter integr al membrane protein	-3.6
Hypothetical		
SCO0257	SCF20.03, hypothetical protein, len: 333 aa;	-3.2
SCO0268	SCF1.10, unknown, len:53 aa	-3.7
SCO0685	SCF15.06c, unknown, len: 135 aa.	-1.6
SCO0725	3SC5B7.03, unknown	-1.7
SCO1432	SC6D7.07, possible membrane protein, len: 479 aa	-1.5
SCO1640	SCI41.23c, unknown, len: 453 aa	-2.5
SCO1700	SCI30A.21c, membrane protein, len: 104 aa	-2.3
SCO2270	SCC75A.16c, possible membrane protein	-3.9
SCO2384	SC4A7.12, unknown, len	-1.5
SCO2519	SCC121.22c, probable membrane protein	-17.0
SCO2668	SC6D10.11, hypothetical protein, len: 452 aa	-2.1
SCO2861	SCE20.35, unknown, len: 240 aa.	-4.6
SCO2979	SCE50.07, possible integral membrane transport protein	-3.1
SCO3657	SCH10.35c, hypothetical protein, len: 136 aa	-1.5
SCO3663	SCH44.03c, possible membrane protein, len: 63 aa	-1.8
SCO4213	2SCD46.26, unknown, len: 120 aa	-4.2
SCO4254	SCD8A.27, hypothetical protein, len: 956 aa	-1.9
SCO4399	SCD10.31, conserved hypothetical protein, len: 461 aa	-1.8
SCO4440	SCD6.18, hypothetical protein, len: 252 aa	-10.8
SCO4515	SCD35.22c, hypothetical protein, len: 401 aa	-1.8
SCO4516	SCD35.23c, unknown, len: 119 aa	-2.2
SCO5230	SC7E4.27c, possible integral membrane protein, len: 94 aa	-3.1
SCO5275	2SC7G11.37, possible ATP/GTP binding protein, len: 1007 aa	-2.0
SCO5329	SC6G9.04c, hypothetical protein, len: 565 aa	-1.8
SCO5332	SCBAC5H2.01, unknown (fragment), len: >33 aa SC6G9.01c	-2.1
SCO5389	2SC6G5.33, unknown, len: 130 aa	-3.0
SCO5899	SC10A5.04, unknown, len: 295 aa	-1.5
SCO6209	SC2G5.30, unknown, len: 191 aa	-3.8
SCO6210	SC2G5.31, unknown, len: 133 aa	-17.2
SCO6569	SC3F9.04, probable secreted solute binding protein	-3.1
SCO6544	SC5C7.29, possible membrane protein, len: 226 aa	-5.7
SCO6621	SC1F2.18, unknown, len: 424 aa	-1.9
SCO6632	SC4G2.06, possible membrane protein	-1.5
SCO6637	SC4G2.11c, unknown, len	-1.7
SCO6638	SC4G2.12c, unknown, len: 605 aa	-1.5

Table A2.4 Con't…

SCO6976	SC8F11.02c, unknown, len: 298 aa	-2.5
SCO6982	SC8F11.08, unknown	-1.8
SCO7608	SC2H2.06, hypothetical protein	-1.9
Miscellaneous Enzymes		
SCO0256	SCF20.02, possible short chain oxidoreductase, len: 265aa	-1.8
SCO0259	SCF1.01, partial CDS, probable alcohol dehydrogenase, len: >167aa	-1.5
SCO0560	SCF73.07c, cpeB, catalase/peroxidase, len: 740 aa	-1.5
SCO0762	SCF81.21c, sti1, protease inhibitor precursor, len: 144 aa	-3.9
SCO1174	SCG11A.05, thcA, aldehyde dehydrogenase, len:534 aa	-9.0
SCO2117	SC6E10.11, trpE1, probable anthranilate synthase	-3.7
SCO2267	SCC75A.13, probable heme oxygenase	-4.0
SCO2914	SCE19A.14c, probable amino acid permease	-4.1
SCO3835	SCH69.05, possible dehydrogenase, len: 563aa	-23.2
SCO4295	SCD95A.28, scoF4, cold shock protein, len: 67 aa	-7.7
SCO5229	SC7E4.26c, probable permease, len: 4532 aa	-3.8
SCO6068	SC9B1.15c, cvnB6, unknown	-2.5
SCO6069	SC9B1.16c, cvnA6, possible large secreted protein	-2.0
SCO6204	C2G5.25c, probable catalase	-11.4
SCO6206	SC2G5.27c, possible oxidoreductase	-3.1
SCO6211	SC9G1.01, possible uricase	-27.6
SCO6212	SC9G1.02, possible permease	-2.7
SCO6213	SC9G1.03, probable hydrolase	-2.8
SCO6214	SC9G1.04, probable permease	-3.2
SCO6243	malate synthase	-1.9
SCO6247	SCAH10.12, possible allantoinase	-5.4
SCO6248	allantoicase	-2.7
SCO6499	SC1E6.08, gvpO, probable gas vesicle synthesis protein	-9.5
SCO6501	SC1E6.10, gvpF, probable gas vesicle synthesis protein	-1.7

Confirmation in the effects on gene expression by ARC2 addition

Promoters of the genes of interest from the microarray were fused upstream of the *luxCDABE* operon to measure the changes in gene expression in the presence of ARC2 (**Table A2.5**). A reporter based strategy for confirmation was opted for over RT-PCR as we envision using this reporter library to assess the effects of the remaining ARC molecules from the original HTS screen. As well, the multiple concentrations and multiple time point analysis required for understanding the ARC2 effect would be too labour intensive for a RNA-based method.

In addition to the changes from the microarray, genes for fatty acid biosynthesis (*fabD*, *fabF* and *fabG*) and a few other regulatory genes from secondary metabolism were included (*afsK*, *afsR* and *absA1*). Genes involved in fatty acid biosynthesis were added to assess the ARC2 effect on its target pathway. *asfK* and *asfR* were included to explore more genes in the AfsK/R/S system (see Chapter 1) and *absA1* was added as a control as its expression was unchanged in the presence of ARC2 during the microarray analysis (**Table A2.5**).

Reproducible changes were observed in the expression of the pleotropic regulators *afsS* and *scbA*, genes involved in branched chain amino acid degradation and NADH dehydrogenase. The involvement of these genes in the ARC2 effect remains to be explored. ARC2 addition resulted in decreased expression of two genes in fatty acid biosynthesis, *fabG* and *fabD* (**Table A2.5**).

Table A2.5. Summary of the effect of ARC2 on various processes

Process	Name	Fold Change Microarray	Fold Change *lux*
Secondary Metabolism			
	Geosmin (SCO6073)	0.72	0.2(20hrs)*
	Unknown PKS (SCO6739)	0.59	0.2(20hrs)*
	Unknown lantibiotic (SCO0269)	.55	-
	CPK	n/d	-
	Dessferrioxamine (SCO2782)	n/d	-
	Prodiginine (RedD)	0.43	0.6(28hrs)*
	Actinorhodin (ActII-4)	1.59	3(28hrs)*
	Germicidin (germA)	n/d	-
Regulators of Secondary Metabolism			
	AfsS	1.5	15(16hrs)*
	AfsR	n/d	-
	AfsK	n/d	-
	AbsA1	n/d	-
	ScbA	0.39	0.2(20hrs)*
	EcrA1 (SCO2519)	0.43	-
Branched Chain Amino Acid Degradation			
	SCO7471	0.5	0.4(20hrs)*
	SCO3831	0.5	0.4(20hrs)*
Fatty Acid Synthesis			
	SCO5144	1.6	0.2(20hrs)
	SCO5385	2.23	-
	SCO1815 FabG	n/d	0.2(20hrs)
	SCO2390 FabF	n/d	-
	SCO2347 FabD	n/d	0.2(20hrs)
	SCO1346 FabG3	0.6	-
	SCO1750	0.79	0.2(20hrs)*
Electron Transport Chain			
	NuoA SCO0216	0.11	0.2(20hrs)*
	SCO0216	2.18	0.6(20hrs)
Miscellaneous			
	SCO0174	2.04	0.2(20hrs)
	SCO0168	1.7	0.6(20hrs)
	SCO0204	1.54	0.4(20hrs)
	SodF – oxidative stress	1.8	-
	glycerol transport (SCO1660)	3.5	-

* denotes reproducible change between microarray and *lux*
n/d denotes no detectable expression
- denotes no change in gene expression in comparison to DMSO

Involvement of ScbA in the mechanism of action of ARC2

One reproducible change in gene expression was the decreased expression of the γ-butyrolactone synthase ScbA. Deletion of this gene results in strong induction of actinorhodin production by an unknown mechanism (3, 4). To explore the possible involvement of ScbA in the ARC2 effect, we obtained a *S. coelicolor scbA* knockout strain (4). The ability of ARC2 to increase the expression of the pathway specific activator for actinorhodin production ActII-4 was unaffected in this mutant strain, suggesting ScbA is not involved in the ARC2 effect (**Figure A2.2**).

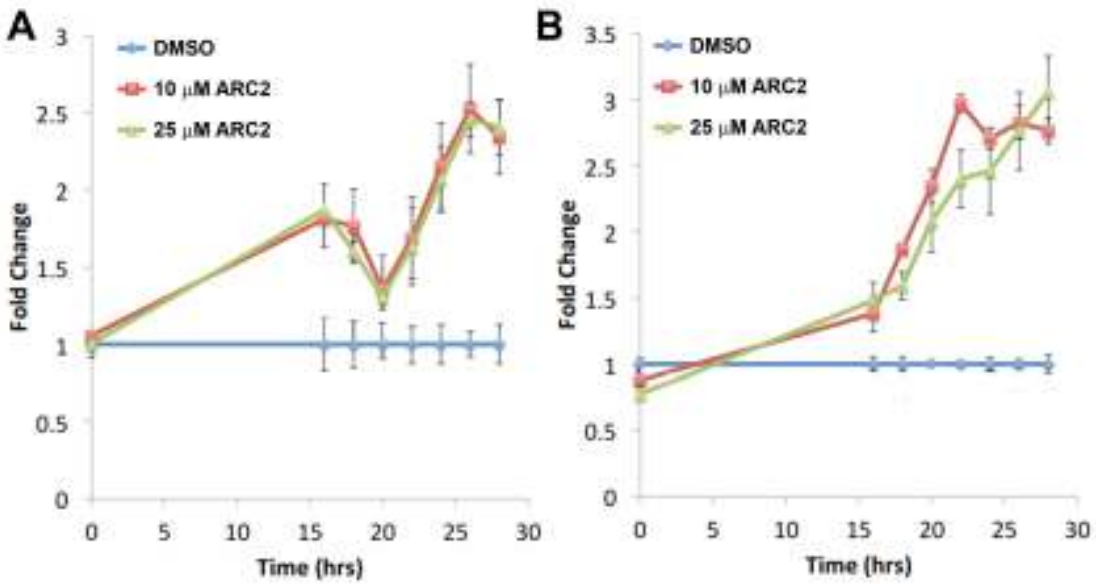

Figure A2.2. Expression of ActII-4 in the presence of ARC2. *S. coelicolor* harbouring the *actII*-4-*lux* fusion was grown for 8 hours followed by addition of either 10 μM or 25 μM ARC2. An equivalent volume of DMSO was added as a control. Luminescence driven by the *actII*-4 promoter was monitored in liquid culture beginning after 16 hours after ARC2 addition in **(A)** M145 and **(B)** *scbA*.

Involvement of AfsS in the mechanism of action of ARC2

The strongest change observed in follow up analysis was the induction of *afsS* expression, resulting in ~15 fold induction within 16 hours of ARC2 addition. This stimulation appears to precede wild type induction of *afsS* expression and may be involved in the early production of actinorhodin observed by ARC2 addition. However, further analysis is needed to explore the involvement of AfsS in the ARC2 effect. A deletion mutant of AfsS is currently being created to test this. If AfsS is involved in this process reduced stimulation of actinorhodin would be expected when ARC2 is applied to this mutant strain.

A complete understanding of the ARC2 Effect

The ARC2 series inhibits a well conserved process, fatty acid biosynthesis which makes the ARC2 series an attractive strategy to alter the secondary metabolic profiles of the actinomycetes. In order to exploit the ARC2 series' effect on secondary metabolism to the fullest, a complete understanding of its mechanism of action is necessary. We believe the effect of ARC2 addition is a result of both increased precursor availability and the induction of *afsS* expression. The possible involvement of AfsS expression is an exciting development as it would improve the activity of the ARC2 series beyond simple precursor availability. The current rationale is that weak fatty acid inhibition leads to improved precursor availability and results in increased yields of acyl-CoA dependent metabolites. If

the ARC2 effect is limited to a buildup in precursors only those secondary metabolites being expressed that depend on these precursors would be influenced. However, the involvement of AfsS would add a second layer to the ARC2 effect and stimulate the production of any AfsS regulated metabolite, extending the utility of the ARC2 series beyond acyl-CoA dependent metabolites.

A2.2 Material and Methods

A2.2.1 Microarray analysis. *S. coelicolor* M145 was inoculated into 1.1L of liquid R5M and incubated shaking at 30^{o}C for 16 hours, the culture was then split in two – one culture receiving 25 μM ARC2 and the other used as the DMSO solvent control. Cultures continued to grow shaking at 30^{0}C for 72 hours. Growth was measured as O.D. 450nm every 2 hours. Actinorhodin production was measured up to 72 hours post ARC2 addition. RNA was harvested from cells collected at subculture (12hr), 1, 2, 4 and 6 hours after subculture. RNA was extracted as previously described (5). RNA samples were sent to London Regional Genomics Center for microarray analysis. For microarray analysis, RNA was converted to cDNA by reverse transcription, biotinylated, fragmented and hybridized to a custom-designed Affymetrix GeneChip array using Affymetrix published protocols. This array was created following previously reported arrays; briefly, this array contained 226,576 oligonucleotide probes each 25 bp in length. 8,205 probes targeting protein-coding regions, 10,834 covering the

intergenic regions and 3,672 probes targeting non-coding RNAs (6). The hybridized GeneChips were stained and washed using an Affymetrix Fluidics Station 450 and scanned using the Affymetrix Scanner 30007G. Transcriptional fold-change of each gene was calculated as the ratio between DMSO and ARC2 treated cells.

A2.2.2 Cloning of the *lux* fusions for microarray follow up. Promoter fusions for the genes of interest from the microarray where introduced upstream of the *luxCDABE* operon in the pF*lux* plasmid. The fusions were primarily ordered from IDT, corresponding to 400bp upstream of the translational start site for each gene. Promoter fusions for *absA1*, *actII-4*, *afsK*, *afsR*, *afsS*, *scbA* were created by PCR amplification from the *S. coelicolor* M145 chromosome and introduced into pF*lux* by restriction digest using EcoRV and NotI. All *lux* fusions were subsequently introduced into *S. coelicolor M145* by conjugal transfer using the mating *E. coli* strain ET12567 containing the pUZ8002 plasmid (7).

A2.2.3. Growth curves. *S. coelicolor* harbouring the individual *lux* fusions were inoculated in liquid R5M to an O.D. 450 nm of 0.1 and grown for 8 hrs at 30^{o}C. Strains were subcultured into 96 well white plates containing a 5 mm glass bead in each well. 150 µL of culture was added into each well and either DMSO, 10 25 or 50 µM ARC2. DMSO concentrations were kept constant at 2%. Luminescence was read at time of ARC2 addition and every 2 hours beginning 16 hours after subculture.

A2.3 References

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

2. Stirrett K, Denoya C & Westpheling J (2009) Branched-chain amino acid catabolism provides precursors for the type II polyketide antibiotic, actinorhodin, via pathways that are nutrient dependent*. J Ind Microbiol Biotechnol* 36: 129-137.

3. Butler MJ*, et al* (2003) Deletion of *scbA* enhances antibiotic production in *Streptomyces lividans. Appl Microbiol Biotechnol* 61: 512-516.

4. D'Alia D*, et al* (2011) Deletion of the signalling molecule synthase ScbA has pleiotropic effects on secondary metabolite biosynthesis, morphological differentiation and primary metabolism in *Streptomyces coelicolor* A3(2)*. Microb Biotechnol* 4: 239-251.

5. Kieser T, Bibb M, Buttner MJ, Chater KF & Hopwood DA (2000) *Practical Streptomyces Genetics,* (The John Innes Foundation, Norwich), pp 613.

6. Nieselt K*, et al* (2010) The dynamic architecture of the metabolic switch in *Streptomyces coelicolor. BMC Genomics* 11: 10.

7. Flett F, Mersinias V & Smith CP (1997) High efficiency intergeneric conjugal transfer of plasmid DNA from *Escherichia coli* to methyl DNA-restricting streptomycetes*. FEMS Microbiol Lett* 155: 223-229.