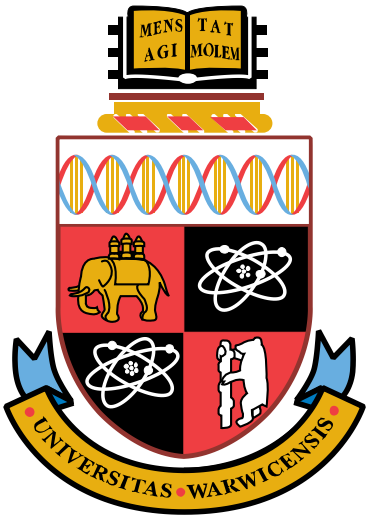


Antibiotic Biosynthesis and its Transcriptional Regulation in *Streptomyces* Bacteria

Shanshan Zhou

Supervisor: Prof. Gregory L. Challis

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Chemistry

Department of Chemistry, University of Warwick

September 2016

Table of Contents

List of Figures

AHFCA1. Lane 1: DNA fragment only. Lanes 2 to 7: AHFCA1 at 0, 0.8, 8, 20, 40 and 100 nmol, respectively. Amounts of DNA and proteins were kept constant (0.8 pmol and 4.0 pmol respectively).

Figure 3.8. Agarose gel electrophoresis analyses of the PCR products for different length of DNA fragments containing MAREs. (LR: low range DNA ladder)

Figure 3.9. Competitive binding of MmfR with two DNA fragments corresponding to the *mmfR-mmfL* and *mmyY-mmyB* intergenic regions. Lanes 1 to 3: DNA fragments only. Lanes 4 to 9: MmfR at 0.1, 0.2, 0.3, 0.4, 0.5 and 0.6 pmol, respectively. The amount of DNA was kept constant (0.05 pmol).

Figure 3.10. Competitive binding of MmfR with two DNA fragments corresponding to the *mmfR-mmfL* intergenic region and upstream of *mmyR*. Lanes 1 to 3: DNA fragments only. Lanes 4 to 9: MmfR at 0.1, 0.2, 0.3, 0.4, 0.5 and 0.6 pmol, respectively. The amount of DNA was kept constant (0.05 pmol).

Figure 3.11. Competitive binding of MmfR with two DNA fragments corresponding to the *mmyY-mmyB* intergenic region and upstream of *mmyR*. Lanes 1 to 3: DNA fragments only. Lanes 4 to 9: MmfR at 0.1, 0.2, 0.3, 0.4, 0.5 and 0.6 pmol, respectively. The amount of DNA was kept constant (0.05 pmol).

Figure 3.12. Interaction of MmfR with the 194-bp DNA fragment of the entire *mmfR-mmfL* intergenic region in response to increasing amounts of natural AHFCAs 1-5. Lane 1: DNA fragment only. Lanes 2 to 9: AHFCAs at 0, 0.8, 4, 8, 14, 20, 40 and 100 nmol, respectively. Amounts of DNA and proteins were kept constant (0.1 pmol and 1.8 pmol respectively).

Figure 3.13. Interaction of MmfR with the 230-bp DNA fragment of the entire *mmyY-mmyB* intergenic region in response to increasing amounts of natural AHFCAs 1-5. Lane 1: DNA fragment only. Lanes 2 to 9: AHFCAs at 0, 0.8, 4, 8, 14, 20, 40 and 100 nmol, respectively. Amounts of DNA and proteins were kept constant (0.1 pmol and 1.8 pmol respectively).

Figure 3.14. (A) Proposed key residues of MmfR interacting with AHFCA2 from solved crystal structure. (B) Key hydrogen bonding interactions in the hydrophilic part of AHFCA2.

Figure 3.15. SDS-PAGE analysis of His_6-MmfR mutants, Y85F, Y85A, Q130E, Y144F and Y144A, after purification. (MWM: molecular weight marker)

Figure 3.16. Measured (top) and deconvoluted (bottom) mass spectra of His_6-MmfR mutants, Y85F (calculated mass = 27819.5 Da), Y85A (calculated mass = 27743.5 Da),

Y144F (calculated mass = 27819.5 Da), Y144A (calculated mass = 27743.5 Da) and Q130E (calculated mass = 27836.5 Da).

Figure 3.17. SDS-PAGE analysis in the attempt to purify the MmfR mutant L110V.

Figure 3.18. Comparison of the effect of AHFCA1 on protein:DNA (*mmfL-mmfR* fragment, 194 bp) dissociation when WT MmfR, MmfR Y85F and Y85A were used. Lane 1: DNA fragment only. Lanes 2 to 9: AHFCA1 at 0, 0.8, 4, 8, 14, 20, 40 and 100 nmol, respectively. Amount of DNA and proteins were kept constant (0.1 pmol and 1.8 pmol respectively).

Figure 3.19. Comparison of the effect of AHFCA1 on protein:DNA (*mmfL-mmfR* fragment, 194 bp) dissociation when WT MmfR and MmfR Q130E mutant were used. Lane 1: DNA fragment only. Lanes 2 to 9: AHFCA1 at 0, 0.8, 4, 8, 14, 20, 40 and 100 nmol, respectively. Amount of DNA and proteins were kept constant (0.1 pmol and 1.8 pmol respectively).

Figure 3.20. Comparison of the effect of AHFCA1 on protein:DNA (*mmfL-mmfR* fragment, 194 bp) dissociation when WT MmfR, MmfR Y144A and Y144F were used. Lane 1: DNA fragment only. Lanes 2 to 9: AHFCA1 at 0, 0.8, 4, 8, 14, 20, 40 and 100 nmol, respectively. Amount of DNA and proteins were kept constant (0.1 pmol and 1.8 pmol respectively).

Figure 3.21. Comparison of the 400 MHz ^{1}H NMR spectra of racemic SCB1 and their diastereoisomers.

Figure 3.22. Interaction of MmfR with the 194-bp DNA fragment of the entire *mmfR-mmfL* intergenic region in response to increasing amounts of AHFCA1 and racemic SCB1. Lane 1: DNA fragment only. Lanes 2 to 7: AHFCA1 or racemic SCB1 at 0, 0.8, 8, 20, 40 and 400 nmol, respectively. Amounts of DNA and proteins were kept constant (0.1 pmol and 1.6 pmol respectively).

Figure 3.23. Agarose gel electrophoresis analysis of the *sgnR* PCR products (604 bp). (LR: low range DNA ladder)

Figure 3.24. Agarose gel electrophoresis analysis of restriction digests (left) using *PstI* to confirm the constructed pET151-*sgnR* plasmid, and PCRs (right) using the constructed pET151-*sgnR* plasmid as template (604 bp). (LR: low range DNA ladder)

Figure 3.25. SDS-PAGE analysis of total (P_{TOT}) and soluble (P_{Sol}) protein fractions resulting from overproduction of His_6-SgnR in *E. coli* BL21star (DE3) (left), and His_6-SgnR after purification (right). (MWM: molecular weight marker)

Figure 4.6. EICs at *m/z* = 325.0675, corresponding to $[M+H]^+$ for the metabolite (top trace) from LC-HRMS analyses of the ethanol extract of *S. coelicolor* M1152/SV-2_E03::*Ssp*I culture supernatant, and synthetic mixture of pyochelins and neopyochelins (bottom trace), including two stereoisomers identical to natural pyochelin from *S. scabies* 87-22 and *P. aeruginosa*.

Figure 4.7. EICs for *m/z* = 325.0675, corresponding to $[M+H]^+$ for isopyochelin, from LC-HRMS analyses of the ethanol extract of *S. coelicolor* M1152/SV-2_E03::*Ssp*I culture supernatant (top trace), the authentic synthetic standard of the isopyochelin stereoisomers from 2-methyl-L-cysteine (middle trace) and the extract to which an approximately equimolar quantity of the synthetic standard has been added (bottom trace).

Figure 4.8. CD spectra of commercially available 2-methyl-L-cysteine and 2-methyl-D-cysteine used in the synthesis of isopyochelin.

Figure 4.9. EICs at *m/z* = 325.0 (corresponding to $[M+H]^+$ for isopyochelin) from LC-HRMS analyses of the synthetic stereoisomers of isopyochelin and the natural metabolites from the ethanol extract of *S. coelicolor* M1152/SV-2_E03::*Ssp*I culture supernatant on a homochiral stationary phase. Comparison with 4"*R*-isopyochelins **135** (left) and 4"*S*-isopyochelins **136** (right) is shown separately.

Figure 4.10. EICs at *m/z* = 353.0988, 339.0832, 210.0583, 224.0740 and 325.0675, corresponding to $[M+H]^+$ for watasemycin, thiazostatin, aerugine, pulicatins A/B and isopyochelin, respectively, from LC-HRMS analyses of the ethanol extract of the Δ*sven0516* mutant culture supernatant (purple traces), in comparison to the *S. coelicolor* M1152/SV-2_E03::*Ssp*I strain (blue traces).

Figure 4.11. EICs at *m/z* = 353.0988, 339.0832, 210.0583, 224.0740 and 325.0675, corresponding to $[M+H]^+$ for watasemycin, thiazostatin, aerugine, pulicatins A/B and isopyochelin, respectively, from LC-HRMS analyses of the ethanol extract of the Δ*sven0508* mutant culture supernatant.

Figure 4.12. Two reported possible absolute configurations of watasemycin and thiazostatin. The absolute stereochemistry with 4"*S* configuration is circled in red.

Figure 4.13. Reassigned stereochemistry of thiazostatin, watasemycin and isopyochelin.

Figure 4.14. EICs at *m/z* = 353.0988, 339.0832, 210.0583, 224.0740 and 325.0675, corresponding to $[M+H]^+$ for watasemycin, thiazostatin, aerugine, pulicatins A/B and isopyochelin, respectively, from LC-MS analyses of the ethanol extract of the Δ*sven0515*

List of Schemes

List of Tables

Acknowledgements

Firstly, I would like to express my deep gratitude to my supervisor, Prof. Gregory L. Challis, for giving me the opportunity to carry out a PhD project under his patience, guidance, support and supervision. His wide knowledge and creative way of thinking through constructive comments brought invaluable guidance to this work. I would like to extend my gratitude to my advisory panel members, Dr. Claudia Blindauer and Prof. Peter J. Sadler, for their helpful discussions and advice throughout my PhD.

Secondly, I would like to thank Dr. Christophe Corre for his guidance and useful discussions in the biology work. I am truly grateful to Dr. Lijiang Song for his kind help and expertise in mass spectrometry, to Dr. Lona Alkhalaf for her helpful comments and advice on my thesis. For their technical advice and helpful comments in the lab, I wish to thank Dr. Daniel Zabala-Alvarez, Dr. Joleen Masschelein, Dr. Matthew Jenner, Dr. Simone Kosol, Dr. Vincent Poon, Dr. Chuan Huang, Dr. Douglas Roberts, Dr. Yousef Dashti, Dr. Mirian Rodriguez-Garcia, Dr. Emzo De Los Santos, Dr. Matias Rey, Dr. David Withall, Dr. John D. Sidda, Daniel Griffiths, Jade Ronan, Ruby Awodi, Rebin Salih, Gideon Idowu, Joshua Cartwright, Rakesh Saroay, Richard Gibson, Marianne Costa, Xinyun Jian, Christian Hobson, Chris Perry, Matt Beech and all of the other people who have made working in the Chemical Biology Research Facility such a great experience.

I wish to thank Nicolas Malet, Peter Harrison and Dr. Yuki Inahashi for their previous work on the project. I also wish to thank Dr. Ivan Prokes and Phil Aston for all of their training and technical support with the NMR and mass spec aspects of the project.

I would also like to thank my friends that support me through my PhD, making these four years memorable: Yijun Zhu, Tiantian Fu, Candace Ho, Hualong Song and many others.

Finally, I also wish to thank my parents and my brother for all of their support throughout my PhD studies. Last but not least, I wish to thank Rubing Wang for his understanding, encouragement and support during the time we were away from each other.

Declaration

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for any degree.

The work presented (including data generated and data analysis) was carried out by the author except in the cases outlined below:

Work in Chapter 4 on chiral LC-MS was performed by Dr. Lijiang Song

Parts of this thesis have been published or accepted for publication by the author:

Zhou, S., Alkhalaf, L., de los Santos, E. L. C. and Challis, G. L., Mechanistic insights into class B radical-*S*-adenosylmethionine methylases: Ubiquitous tailoring enzymes in natural product biosynthesis. *Curr. Opin. Chem. Biol.* **2016**, *35*, 73-79.

Inahashi, Y.,† Zhou, S.,† Bibb, M. J., Song, L., Al-Bassam, M. M., Bibb, M. J. and Challis, G. L., Watasemycin biosynthesis in *Streptomyces venezuelae*: thiazoline C-methylation by a type B radical-SAM methylase. *Chem. Sci.* **2017**, *8*, accepted for publication. († equal contribution)

Abstract

Streptomyces, the largest genus of Actinobacteria, are renowned for their ability to produce a wide variety of specialised metabolites, including many clinically used antibiotics and other bioactive natural products. Biosynthesis of these specialised metabolites is often tightly regulated by transcriptional regulators, some of which are responsive to signalling molecules, e.g., 2-alkyl-4-hydroxymethylfuran-3-carboxylic acids (AHFCAs), thus initiating a series of regulatory events.

The biosynthetic role of MmfL in AHFCA assembly has been investigated *in vitro* using a synthetic substrate. MmfL has been shown to catalyse an AfsA-like reaction, involving condensation of an ACP-bound β-ketothioester and dihydroxyacetone phosphate to form a phosphorylated butenolide intermediate.

The interactions between DNA, AHFCA ligands and ArpA-like repressors, MmfR and SgnR, have been investigated *in vitro* using electrophoretic mobility shift assays. This work leads to a better understanding of the regulation mechanism in specialised metabolite biosynthesis.

Heterologous expression of a putative pyochelin-like gene cluster, *sven0503-sven0517* from *S. venezuelae* ATCC 10712, has previously been shown to result in the production of thiazostatin and watasemycin as the main metabolic products of the cluster. Isopyochelin, a structural isomer of pyochelin, was also identified and characterised by comparison with synthetic standards. Sven0516 has previously been identified as the thiazoline reductase required for the biosynthesis of all of the metabolic products of the cluster. The class B radical SAM methylase Sven0515 was shown to be responsible for the methylation of thiazostatin to give watasemycin. This is the first experimentally-validated example of such a reaction in the biosynthesis of a nonribosomal peptide. From incorporation experiments using stereospecifically deuterium-labelled cysteine, it has been demonstrated that Sven0515-mediated methylation proceeds with abstraction of the *pro-R* hydrogen atom and results in inversion of stereochemistry at the methylating position. The absolute stereochemistry of thiazostatin and watasemycin has been reassigned on the basis of the absolute configuration determined for isopyochelin.

Abbreviations

ACP	acyl carrier protein
AHFCA	2-alkyl-4-hydroxymethylfuran-3-carboxylic acid
AMP	adenosine monophosphate
AMR	antimicrobial resistance
ARE	autoregulatory response element
ATP	adenosine triphosphate
Boc	*t*-butyloxycarbonyl
Bu	butyl
CD	circular dichroism
CDA	calcium-dependent antibiotic
CDMT	2-chloro-4,6-dimethoxy-1,3,5-triazine
CoA	coenzyme A
5′-dAdo	5′-deoxyadenosyl
dAdoH	5'-deoxyadenosine
DCC	*N*,*N*′-dicyclohexylcarbodiimide
DCU	dicyclohexylurea
DCM	dichloromethane
DECP	diethyl cyanophosphonate
DHA	dihydroxyacetone
DHAP	dihydroxyacetone phosphate
DIBAL-H	diisobutylaluminium hydride
DMAP	4-dimethylaminopyridine
DMF	dimethylformamide
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DTT	dithiothreitol
EDC	1-ethyl-3-(3-dimethylaminopropyl) carbodiimide
EDTA	ethylenediaminetetraacetic acid

EIC	extracted ion chromatogram
EMSA	electrophoretic mobility shift assay
ESBL	extended spectrum beta-lactamase
ESI	electrospray ionization
Et	ethyl
GBL	γ-butyrolactone
HEP	2-hydroxyethylphosphonate
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HPLC	high performance liquid chromatography
HRMS	high resolution mass spectrometry
HTH	helix-turn-helix
HPP	2-hydroxypropylphosphonate
IPTG	isopropyl β-D-1-thiogalactopyranoside
LDA	lithium diisopropylamide
Me	methyl
MeCbl	methylcobalamin
MMF	methylenomycin furan
MRSA	methicillin-resistant *Staphylococcus aureus*
MS	mass spectrometry
MS/MS	tandem mass spectrometry
MT	methyl transferase
MWM	molecular weight marker
m/z	mass/charge ratio
NAC	*N*-acetylcysteamine
NAD(P)H	nicotinamide adenine dinucleotide (phosphate)
NMM	*N*-methylmorpholine
NMR	nuclear magnetic resonance
NOESY	nuclear overhauser effect spectroscopy
NRP	nonribosomally biosynthesised peptide
NRPS	nonribosomal peptide synthetase

OD	optical density
PAGE	polyacrylamide gel electrophoresis
PCD	programmed cell death
PCP	peptidyl carrier protein
PCR	polymerase chain reaction
PLP	pyridoxal phosphate
PMSF	phenylmethylsulfonyl fluoride
ppm	parts per million
RiPP	ribosomally biosynthesised and posttranslationally modified peptide
RNA	ribonucleic acid
ROS	reactive oxygen species
RS	radical SAM
RT	room temperature
SAH	*S*-adenosyl homocysteine
SAM	*S*-adenosylmethionine
SCB	*Streptomyces coelicolor* butyrolactone
SDS	sodium dodecyl sulfate
SFM	soya flour medium
SMM	supplemented minimal medium
TBAF	tetrabutylammonium fluoride
TBAI	tetrabutylammonium iodide
TBDMS	*tert*-butyldimethylsilyl
TBDMSOTf	*tert*-butyldimethylsilyl triflate
TE	thioesterase
TEMED	*N*,*N*,*N*,*N*-tetramethylethylenediamine
THF	tetrahydrofuran
TLC	thin layer chromatography
TMS	tetramethylsilane
UV	ultraviolet
VB	virginiae butanolide

VRE	vancomycin-resistant *Enterococcus*
VRSA	vancomycin-resistant *Staphylococcus aureus*
WT	wild type
Xre	xenobiotic response element
YD	yeast-extract dextrose

1. Introduction

1.1 *Streptomyces*

Streptomyces is the largest genus of Actinobacteria and over 600 species of *Streptomyces* bacteria have been described while the number is still increasing every year.[1-2] Found predominantly in soil and decaying vegetation, *Streptomyces* is widely spread around the world. As with the other Actinobacteria, streptomycetes are Gram-positive bacteria with high GC content in their genome. *Streptomyces* is renowned for their ability to produce antibiotics and other bioactive natural products with a wide range of applications in medicine and agriculture. The production of antibiotics, as well as other specialised metabolites, is temporally correlated to the onset of development in the *Streptomyces* life cycle (Figure 1.1).[3]

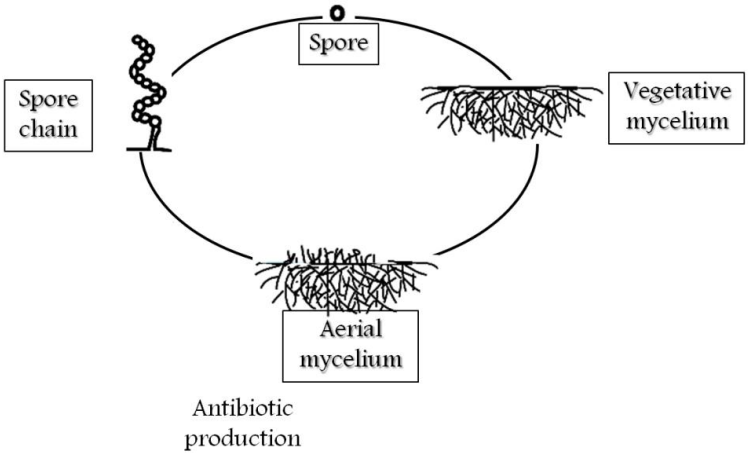


Figure 1.1. The *Streptomyces* life cycle.

The life cycle of the multicellular mycelial *Streptomyces* starts with the germination of a spore that grows out to form vegetative hyphae, after which a process of hyphal growth and branching results in an intricately branched vegetative mycelium. Under adverse conditions, such as nutrient depletion, the vegetative mycelium differentiates to form erected sporogenic structures called aerial hyphae. This is also the moment in the life cycle when most antibiotics are produced. The reason underlying the production of antibiotics at that stage may be due to the need to defend the colony when it is undergoing programmed cell death (PCD).[4] When nutrient depletion occurs, the vegetative or substrate mycelium is autolytically degraded by a PCD-like mechanism so as to provide nutrients for the next round of biomass formation. The resulting nutrients, including amino acids, aminosugars, nucleotides, and lipids, around the lysing substrate mycelium

would inevitably attract motile competing microbes in the habitat. Thus it has been assumed that antibiotics are produced at this time to protect the pool of nutrients.

Genetic studies of *Streptomyces* development regulation have revealed a collection of developmental genes based on analysis of their mutant strains defective in different stages of development.[4] Genes that are required for the formation of aerial hyphae are referred to as *bld* genes, in reference to the bald phenotype of mutants that lack the fluffy aerial hyphae due to their defect in aerial growth. Genes that are required for the formation of mature grey spores on the tips of the white aerial mycelium are called *whi* (white) genes, corresponding to those mutants that are defective in producing the grey spore pigment. Examples of *bld* genes include *bldD*, a highly pleiotropic transcription factor that controls hundreds of development-related genes, and *bldM*, which has been identified to play a critical role in *Streptomyces* differentiation, and many others.[4-6]

The history of antibiotics derived from the genus *Streptomyces* begins with the discovery of streptothricin **1** from *Streptomyces lavendulae* in 1942.[7] Streptothricins are broad antimicrobial agents with a potent inhibitory activity for prokaryotic protein synthesis. However, their cytotoxicity prevents further clinical or veterinary use. It was streptomycin **2** from *Streptomyces griseus*, which came 2 years later, that triggered systematic screening of antibiotics within this genus.[8] Streptomycin is a clinically-used aminoglycoside antibiotic typically used in the treatment of tuberculosis, always in combination with other antituberculosis agents, and sensitive Gram-negative infections. Its acts on the ribosomes by interfering with the binding of aminoacyl-tRNA, causing an erroneous reading of the genetic code, inhibition of initiation of translation of mRNA and aberrant proofreading.[9-11] After the discovery of streptomycin, many other bioactive natural products have been isolated from *Streptomyces* species, including antibacterials (e.g., daptomycin **3** and chloramphenicol **4**), antitumor agents (e.g., doxorubicin **5**), immunosuppressants (e.g., tacrolimus **6** and rapamycin **7**), anthelmintics and insecticides (e.g., avermectin B1a **8** and B1b **9**), herbicides (e.g., phosphinothricin **10**) and antifungal agents (e.g., amphotericin B **11**) (Figure 1.2).[1,12] It is now estimated that over 70% of antibiotics currently used clinically are sourced from *Streptomyces* species.

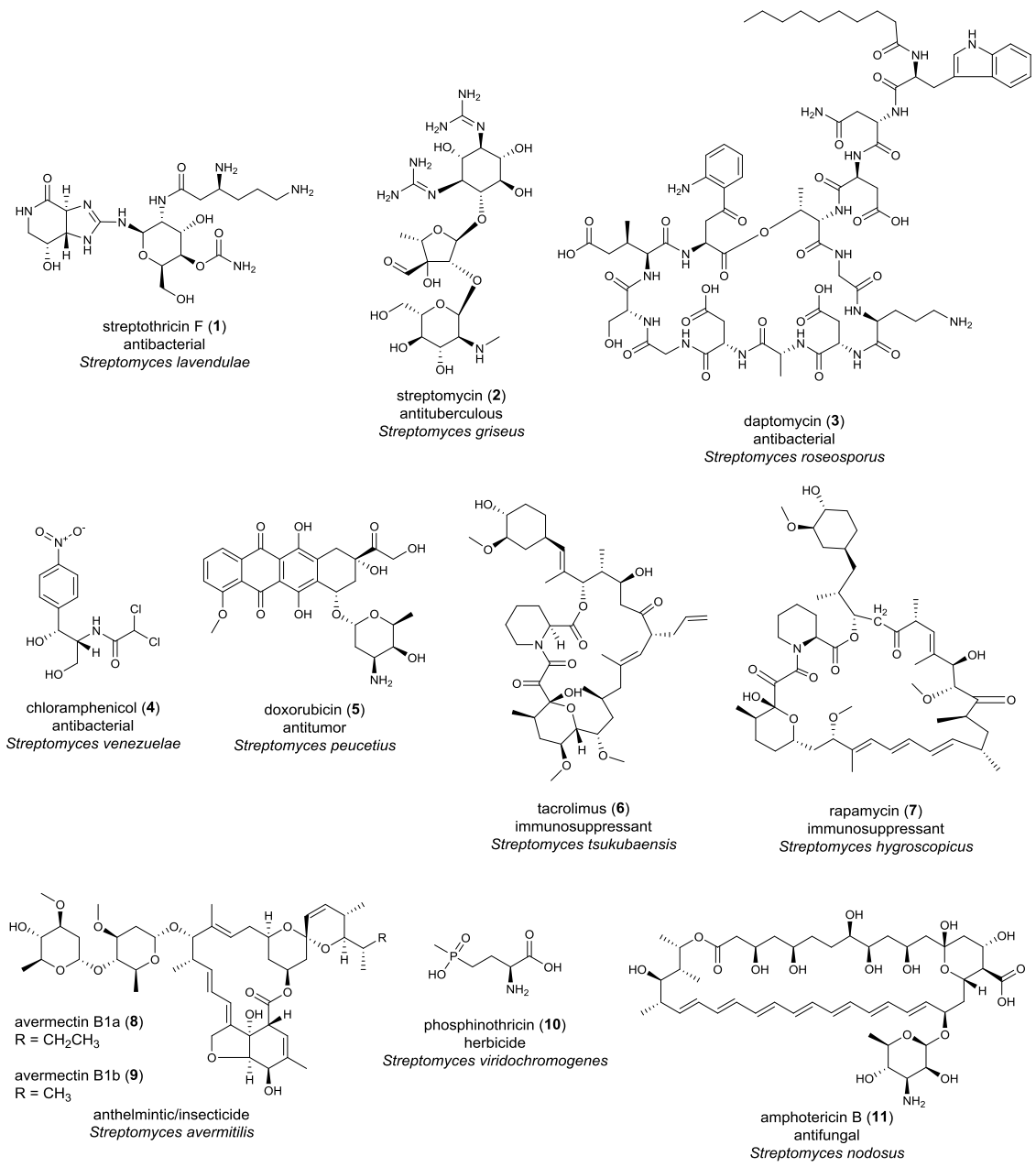


Figure 1.2. Examples of bioactive metabolites that are produced by *Streptomyces* species. The activity and producing organism for each metabolite is shown below its name.

Despite the success of the discovery of antibiotics, and advances in the process of their production, there is an urgent need to discover and develop new antibiotics due to the emergence of antimicrobial resistance (AMR) in microorganisms.[13-14] Rising drug resistance means infectious diseases are increasingly difficult to treat, requiring alternative medications or higher doses, which may be more costly or more toxic. It has been reported that a few infections are now completely untreatable due to AMR, such as gonorrhoea.[15] Some well-known examples of drug-resistant bacteria include methicillin-resistant *Staphylococcus aureus* (MRSA), vancomycin-resistant *S. aureus* (VRSA),

extended spectrum beta-lactamase (ESBL) and vancomycin-resistant *Enterococcus* (VRE). However, compared to the "golden age" of antibiotic discovery, the number of antimicrobial compounds reported per year have substantially declined.[16-17] This decline may be attributed to the reduced probability of finding novel compounds in traditional screening and reduced funding for screening efforts. As the largest antibiotic-producing genus, *Streptomyces* species play a very important role in the fight against emerging multidrug-resistant pathogens. Several species of *Streptomyces* have been particularly well studied, such as *Streptomyces coelicolor* A3(2), the most widely used in genetic studies.

1.1.1 *Streptomyces coelicolor* A3(2)

S. coelicolor A3(2) is a model organism that has been used for more than five decades to study the genetic and biochemical basis for the production of bioactive metabolites.[18] The sequencing of the *S. coelicolor* A3(2) genome, as well as its two plasmids, was completed between 2002-2004.[19-21] Its linear chromosome contains 8.7 Mb with 7,825 predicted open reading frames while the giant linear plasmid SCP1 and the small circular plasmid SCP2 is 365 kb and 31 kb, respectively. Before the DNA sequencing of *S. coelicolor* A3(2), several specialised metabolites had already been discovered (Figure 1.3),[19] including actinorhodins (e.g., **12**), prodiginines (e.g., undecylprodigiosin **13**), calcium-dependent antibiotics (CDAs, e.g., CDA 2a **14**), grey spore pigment,[22] *S. coelicolor* γ-butyrolactones (SCBs, e.g., SCB1 **15**) and methylenomycins (e.g., methylenomycin A **16**),[23-24] the gene cluster of which is carried on the plasmid SCP1.

After the genome sequence of *S. coelicolor* was published, it became apparent that the true potential of this species as a producer of natural products had actually been underestimated.[19] Exploitation of the genome sequence has led to the identification of over 20 biosynthetic gene clusters, including 18 cryptic ones, for specialised metabolites. The products of the cryptic gene clusters were previously unknown, with many not expressed under standard laboratory conditions. A range of metabolic products have been confirmed to be produced from *S. coelicolor* A3(2),[18] such as desferrioxamines (e.g., desferrioxamine E **17**), coelichelin **18**, geosmin **19**, albaflavenone **20**, 2-methylisoborneol **21**, carotenoids (e.g., isorenieratene **22**), hopanoids (e.g., aminotrihydroxybacteriohopane **23**), germicidins (e.g., germicidin A **24**), flaviolin oligomers (e.g., flaviolin dimer **25**), methylenomycin furans (e.g., AHFCA1 **26**) and

coelimycins (e.g., coelimycin P1 **27)** (Figure 1.3). Among all of the above metabolic products, at least six of them, including methylenomycins,[25] actinorhodins,[26] prodiginines,[27] CDAs,[28] coelimycins,[18,29-30] and albaflavenone,[31-32] have shown antibiotic activity while desferrioxamines and coelichelin are known to function as siderophores in *S. coelicolor*.[33-35]

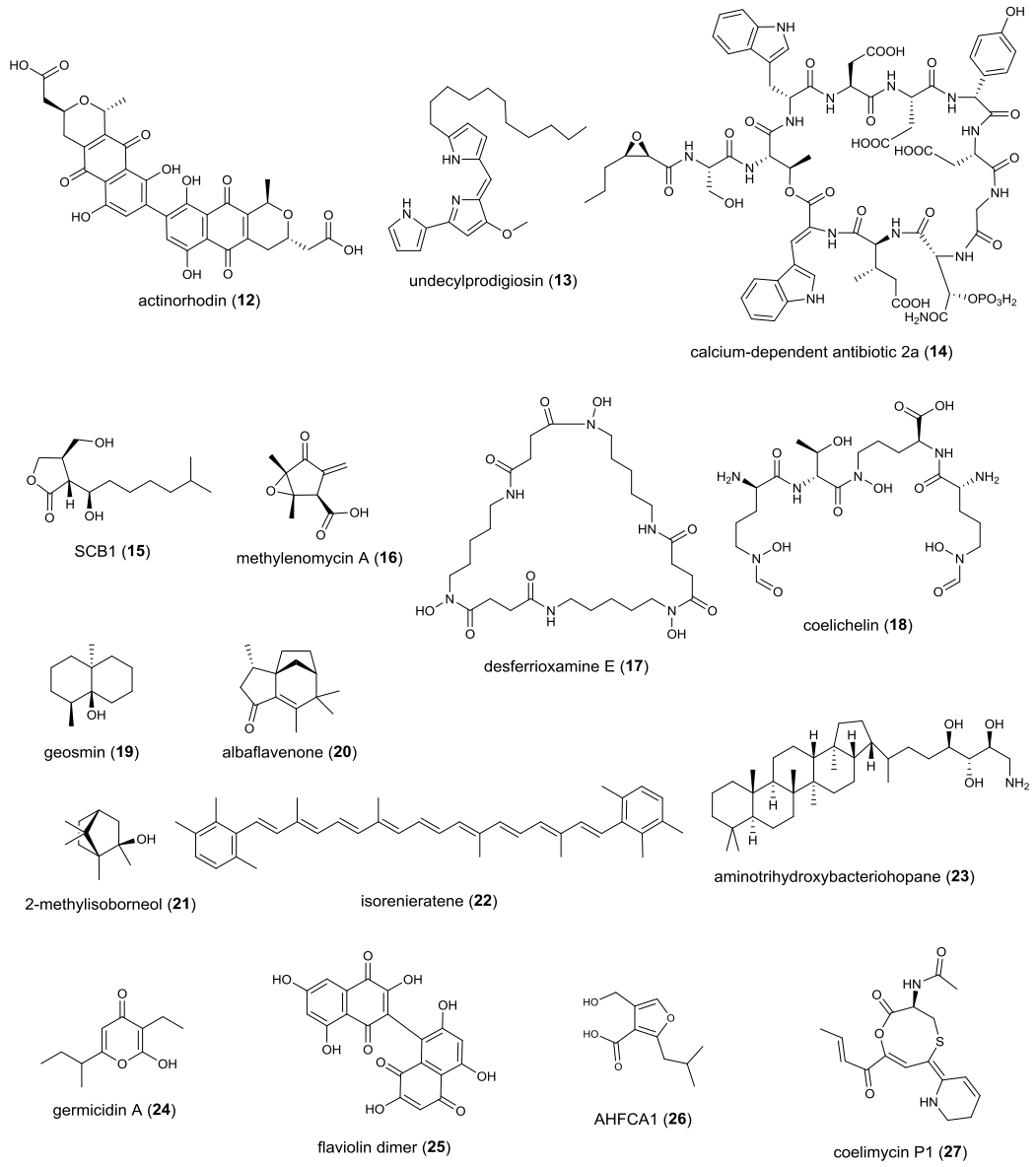


Figure 1.3. Representative structures of specialised metabolites produced by *S. coelicolor*.

Therefore, it appears that the potential of *S. coelicolor* for specialised metabolite production is much greater than originally anticipated from classical bioassay-guided discovery approaches. Indeed, this does not only occur in *S. coelicolor,* but also in other species, like *S. griseus*,[36] *Streptomyces avermitilis* and *Streptomyces venezuelae*,[37-38]

which are all predicted to contain more than 30 natural product biosynthetic gene clusters from analysis of their genome sequences. This predicted number of putative natural products is much higher than has been identified so far. This has prompted extensive research into those cryptic, silent, or sleeping metabolic products and methods for activating their biosynthesis.[39-42]

1.2 *Streptomyces* signalling molecules

1.2.1 Structural types of *Streptomyces* signalling molecules

Of particular interest to many specialised metabolite biosynthetic pathways in bacteria are involvement of diffusible and low molecular weight signalling molecules, also known as 'bacterial hormones' or autoinducers.[43] These molecules control complex regulatory systems in combination with transcriptional regulators, activators or repressors, in a pathway-specific or pleiotropic manner. To date there are two main classes of signalling molecules found in *Streptomyces* species, the γ-butyrolactone (GBL)-type molecules and the furan-type autoinducers, 2-alkyl-4-hydroxymethylfuran-3-carboxylic acids (AHFCAs), which were previously known as methylenomycin furans (MMFs).[44-45] Repressors which recognise these types of signalling molecules all belong to the TetR subfamily of ArpA-like transcriptional repressors and are often involved in regulation of antibiotic biosynthesis. These transcriptional repressors and their regulation involving small signalling molecules are discussed in Section 1.3.

The first example of γ-butyrolactone signalling molecules is autoregulatory factor (A-factor) **28** (Figure 1.5) which was isolated from *Streptomyces* species in 1967.[46] It has been found to induce morphological differentiation, such as aerial hyphae formation, and production of specialised metabolites, including streptomycin and a yellow pigment grixazone, in *S. griseus*.[47-48] Thus, A-factor has been established to have a pleiotropic effect.

Since the discovery of A-factor, other homologues with a γ-butyrolactone have been identified as signalling molecules for specialised metabolism in a number of *Streptomyces* species and related genera.[49] This has thus opened a novel paradigm of the characteristic regulatory system for cellular differentiation in prokaryotes. Based on the currently elucidated structure of GBLs, these molecules can be generally divided into three groups (Figure 1.4) according to the substituent and its stereochemistry at C-6: i)

6-keto type, such as A-factor **28**, which is the only natural example of the 6-keto type of GBLs reported. A-factor readily undergoes epimerisation due to the relative acidity of the proton at C-2 position; ii) IM-2 type, possessing a 6-β-hydroxyl group; iii) VB type, possessing a 6-α-hydroxyl group. Both IM-2 and VB type GBLs belong to the 6-hydroxy type. It has been shown that the absolute stereochemistry of carbons 2, 3 and 6, as well as the length of the alkyl side chains, confers the specificity and affinity of each γ-butyrolactone receptor for its cognate ligand(s).[50-51]

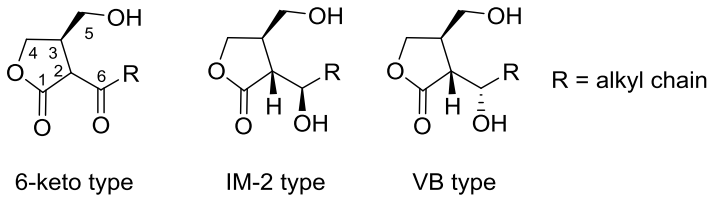


Figure 1.4. Three structural types of GBLs.

There are a range of GBLs isolated and characterised from *Streptomyces* species. They include: factor I **29** from *Streptomyces viridochromogenes*,[52-53] Gräfe factors **30-32** from *Streptomyces bikinensis* and *Streptomyces cyaneofuscatus*;[53-54] virginiae butanolides (VB) A-E **33-37** isolated from *Streptomyces virginiae*;[53,55-56] IM-2 **38** from *Streptomyces lavendulae* FRI-5 and SCBs 1-3 **15**, **39**, **40** from *S. coelicolor* for regulation of both actinorhodin and undecylprodigiosin production.[50,57] Their structures are shown in Figure 1.5 although the absolute stereochemistry of factor I **29** and Gräfe factors **30-32** has not been clearly elucidated.

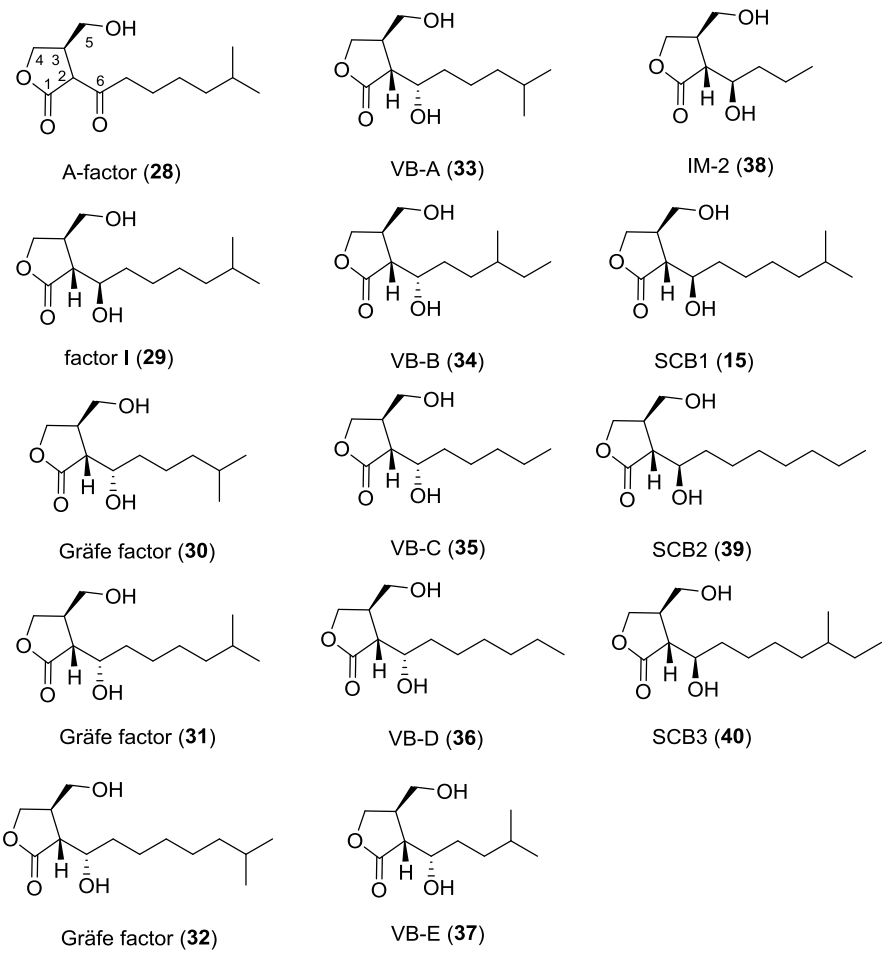


Figure 1.5. Structures of *Streptomyces* GBL-type signalling molecules. The absolute stereochemistry of factor I and Gräfe factors are not clear.

From *S. coelicolor*, the furan type of signalling molecules, AHFCAs 1-5 **26**, **41-44** (Figure 1.6) have been identified and shown to control methylenomycin biosynthesis.[45,58] AHFCA5-7 **44-46** have also been characterised and proposed to function as signalling molecules for regulation of gaburedin biosynthesis in *S. venezuelae* ATCC10712.[59-60]

In addition, another signalling molecule avenolide **47** (Figure 1.6), has been demonstrated to regulate avermectins (e.g., **8** and **9** in Figure 1.2) production in *S. avermitilis*.[61] After its discovery, similar butenolides SRB1 **48** and SRB2 **49** (Figure 1.6) have been reported to induce the production of lankacidin and lankamycin in *Streptomyces rochei*.[62] It seems that they represent a new group of butenolide-type signalling molecules.

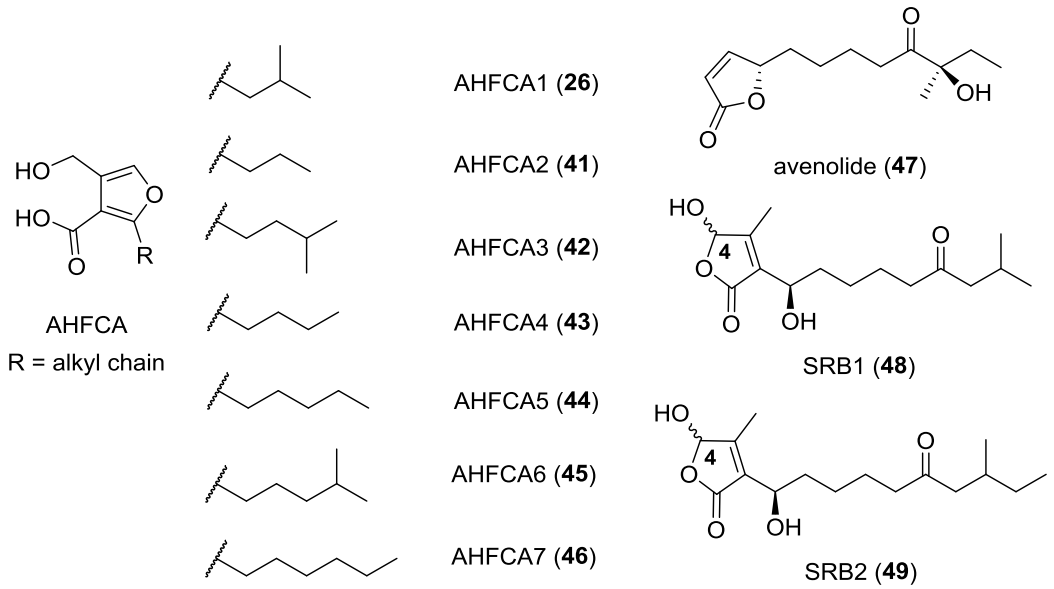


Figure 1.6. Structures of *Streptomyces* AHFCA- and butenolide-type signalling molecules. SRB1 and SRB2 exist as mixtures of interconverting epimers at C-4.[62]

1.2.2 Biosynthesis of γ-butyrolactones

As the most well-studied group of signalling molecules in *Streptomyces*, the biosynthesis of GBLs has attracted a great deal of interests since the discovery of the first γ-butyrolactone molecule A-factor. Investigation into A-factor biosynthesis was first conducted by Horinouchi and co-workers as early as 1984. A-factor production was restored by insertion of the *afsA* gene into an A-factor-deficient *S. griseus* mutant, indicating that AfsA may be an important enzyme required for A-factor biosynthesis.[63-65]

Prior to identification of the specific enzymes responsible for production of GBLs, the biosynthesis of VB-A **33** was investigated by incorporation experiments using stable isotope-labelled precursors in *Streptomyces antibioticus*.[66] The feeding experiments revealed that VB-A is formed from two molecules of acetate and one molecule of both glycerol and isovalerate (Figure 1.7). It was proposed that β-keto acyl CoA is the key precursor, which couples to a C3 unit that is oxidised from the glycerol molecule.

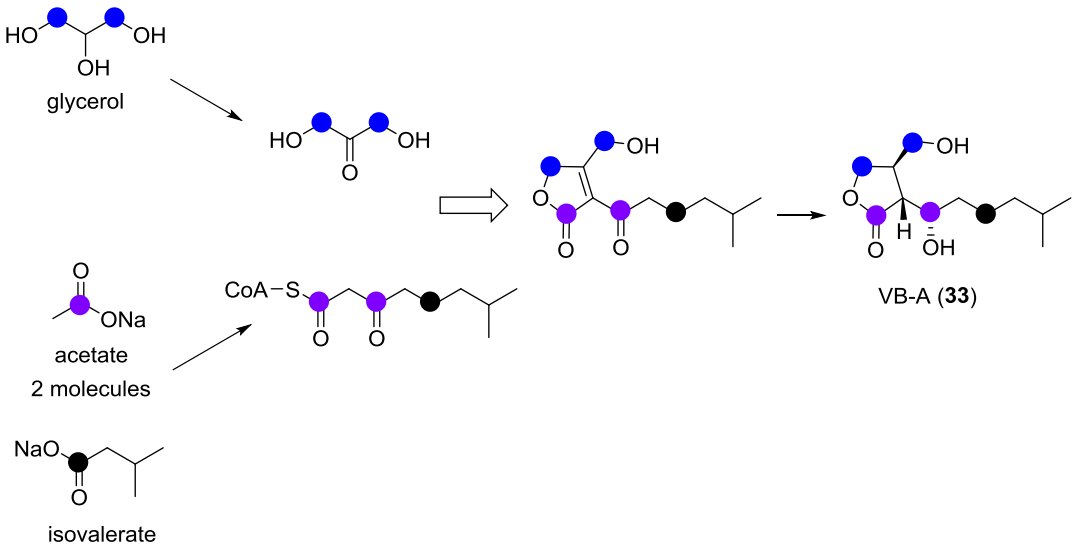
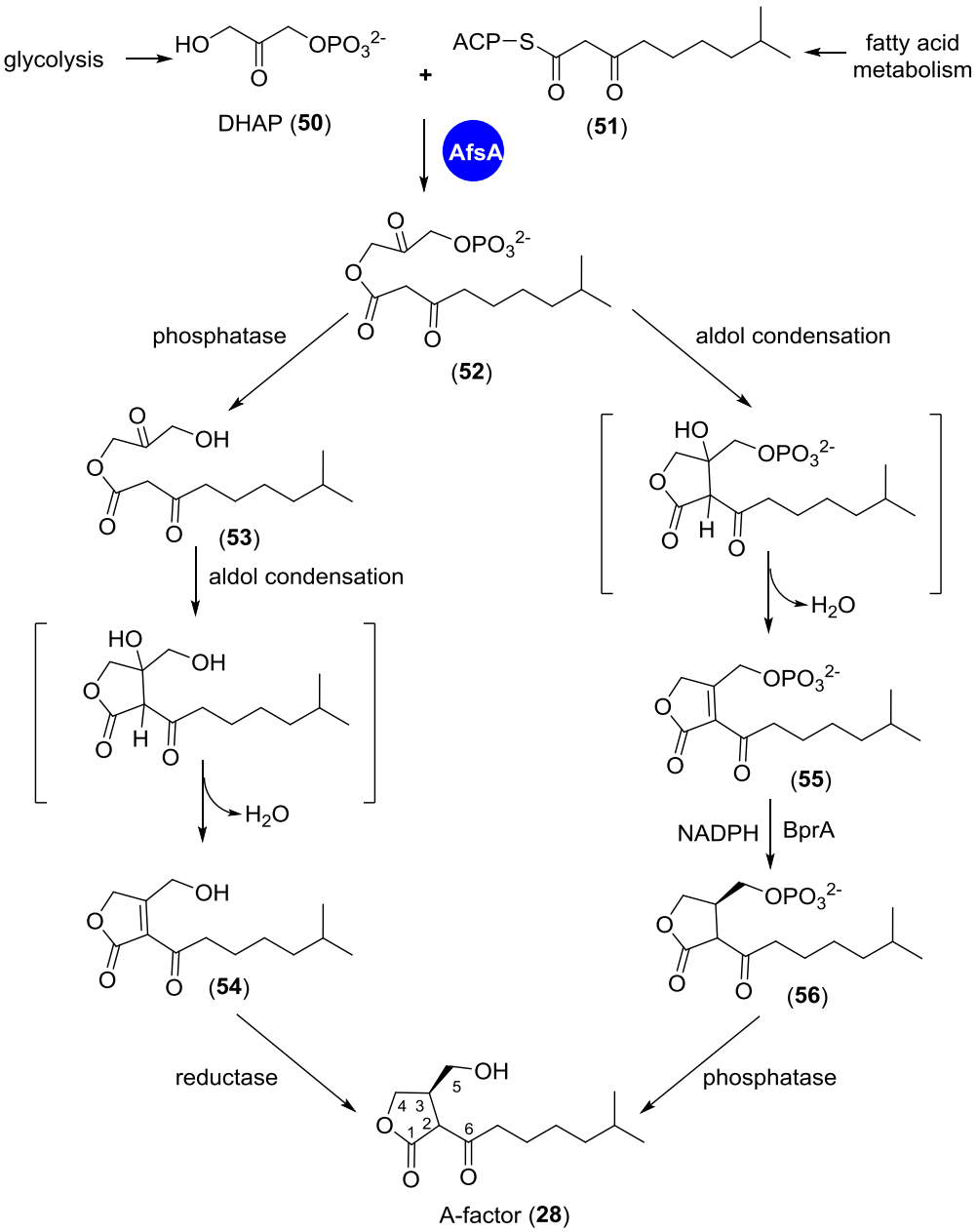


Figure 1.7. Incorporation results using ^{13}C-labelled glycerol, acetate and isovalerate for investigation of VB-A **33** biosynthesis by Sakuda and co-workers.[66]

The biosynthesis of GBLs didn't become much clearer until *in vitro* reconstitution of AfsA by Kato and co-workers in 2007.[67] AfsA has been demonstrated to be the key enzyme required for A-factor biosynthesis, catalysing β-ketoacyl transfer from an acyl carrier protein (ACP)-bound 8-methyl-3-oxononanoyl thioester **51** to the hydroxyl group of dihydroxyacetone phosphate (DHAP) **50** to form an 8-methyl-3-oxononanoyl-DHAP ester intermediate **52** (Scheme 1.1). After formation of the phosphorylated ester intermediate **52**, three further reactions are required for A-factor biosynthesis, including an intramolecular aldol reaction that leads to the butenolide structure **55**, a reduction step for the C2-C3 double bond in the butenolide intermediate and a dephosphorylation step to give the final A-factor **28**. It was also proposed that the three subsequent steps may occur in a different order, i.e. dephosphorylation of the ester intermediate **52** is followed by the aldol condensation and final reduction of the butenolide **54** to A-factor (Scheme 1.1).



Scheme 1.1. Biosynthetic route to A-factor as proposed by Kato and co-workers.

A range of C3 compounds were first tested with His-AfsA, leading to identification of DHAP **50** as the C3 substrate for the AfsA-catalysed reaction. Then the enzymatic assay was undertaken using recombinant His-AfsA, ^{32}P-labelled DHAP and the synthetic 8-methyl-3-oxononanoyl-*N*-acetylcysteamine (NAC), which mimics the corresponding β-ketoacyl-ACP. The reaction resulted in a product with radioactivity on radio-thin layer chromatography (TLC), which was demonstrated to be the ester intermediate, 8-methyl-3-oxononanoyl-DHAP ester **52**, by comparing to its synthetic standard. The butenolide is then formed by a nonenzymatically catalysed intramolecular aldol reaction of the ester intermediate **52**.

BprA (butenolide phosphate reductase), an NADPH-dependent reductase that is encoded by a gene directly downstream of *afsA*, was confirmed *in vitro* to be responsible for reduction of the butenolide **55** to form the butanolide **56**. However, it was shown that the phosphatases and reductases commonly present in bacteria are able to catalyse similar reactions to afford A-factor. For example, introduction of *afsA* into *E. coli* caused the host to produce A-factor analogues, with straight side chains, due to the differences in fatty acid metabolism between *Streptomyces* species and *E. coli*. *Streptomyces* species can produce branched fatty acids from starter units such as isobutyryl-CoA and methylbutyryl-CoA from amino acid degradation, whereas *E. coli* predominantly uses acetyl CoA and thus does not produce branched fatty acids.[68-69]

An AfsA homologue, ScbA in *S. coelicolor*, has been proposed for the biosynthesis of SCBs.[70-71] Takano and co-workers conducted *in silico* analysis of the AfsA family of proteins, including ScbA, and showed that they have similarity to the fatty acid synthesis enzymes FabA and FabZ. Meanwhile, mutation of two predicted active sites in ScbA led to abolishment of the production of SCBs in *S. coelicolor*. Other AfsA homologues have also been found in corresponding γ-butyrolactone producers, e.g. BarX in *S. virginiae* and FarX in *S. lavendulae*.[72-73] It is reasonable to assume that these homologues have similar activity as AfsA in γ-butyrolactone biosynthesis. It was also suggested by Kato and co-workers that a BprA orthologue, ScbB, which has 76% identity in amino acid sequence to BprA, catalyses the reduction of C2-C3 double bond to form the butanolide phosphate **56** in SCB biosynthesis.[67]

Both SCBs in *S. coelicolor* A3(2) and VBs in *S. virginiae* contain a hydroxyl group at C-6 instead of the keto group in A-factor, indicating the presence of additional reductases for this reduction. Due to the absence of such reductase in *S. griseus*, C-6 of A-factor remains as a keto group. One such enzyme, BarS1 in *S. virginiae*, has been characterised. BarS1 was isolated as an NADPH-dependent reductase that reduces the 6-oxo group of the penultimate intermediate in the VB biosynthesis.[74] A BarS1 orthologue, SCO6264, has been proposed to be involved in SCB biosynthesis for a similar reduction reaction in *S. coelicolor*.[70] More investigations are needed to experimentally determine the specific functions of these enzymes.

1.2.3 Biosynthesis of AHFCAs

The furan-type signalling molecules, AHFCAs 1-5, were demonstrated to regulate methylenomycin biosynthesis in *S. coelicolor* by Corre and co-workers in 2008.[45] A putative operon of three genes within the *mmy* cluster, *mmfLHP*, was proposed to direct the biosynthesis of these AHFCA signalling compounds. Deletion of the three genes caused the loss of methylenomycin production.[58] Comparative metabolic profiling has led to the identification and structure elucidation of the AHFCA family members, AHFCAs 1-5 **26**, **41-44** (Figure 1.6), as the products of the *mmfLHP* genes.[45]

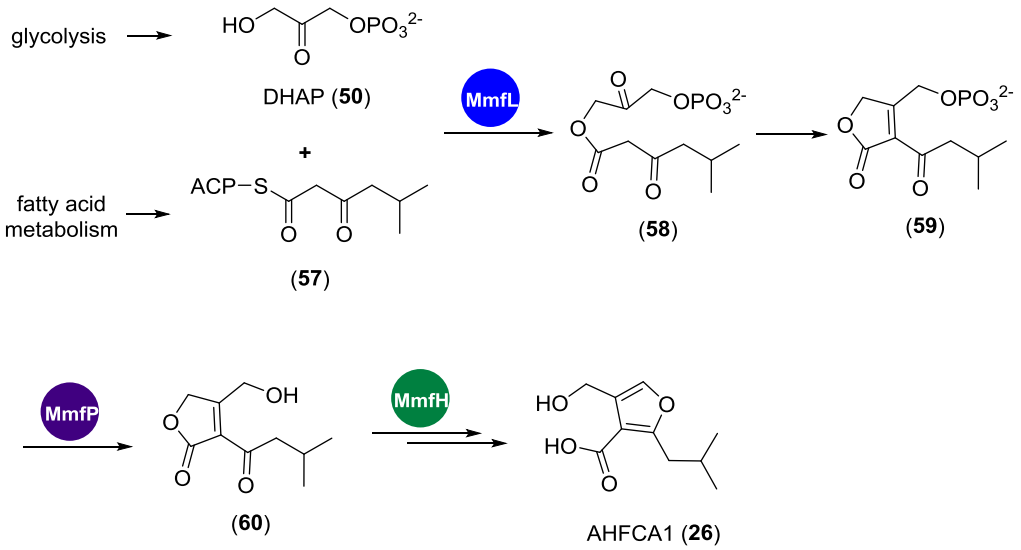
The AfsA homologue, MmfL, has been proposed to play a key role in AHFCA biosynthesis. Methylenomycin production can be restored in an *mmfL*-deficient strain by addition of an EtOAc extract of the culture supernatant from a strain containing *mmfL*. Meanwhile, expression of the *mmfLHP* putative operon in *S. coelicolor* M512, which lacks the SCP1 plasmid, led to the production of the five natural AHFCAs 1-5.

```
MmfL      MNHTNRLLLPAPHDLLF---DGCPPLSFARPLPPADVHKAAAAEVLLTDARPLGENRFAV
ScbA      --------MPEAVVLINSASDANSIEQTALPVPMALVHRTRVQDAFPVSWIPKGGDRFSV
AfsA      --------------------MDAEAEVVHPVGIEMVHRTRPEDAFPRNWVRLGRDRFAV
                                . . *:    **::   :.:  .    * :**:*

MmfL      AALWPRNTFLAHRATSSPCDPLLAAETIRQSAIHLSHTFCDVPIGHHFVLSGLDLDLDLP
ScbA      TAVLPHDHPFFAPVHGDRHDPLLIAETLRQAAMLVFHAGYGVPVGYHFLMATLDYTCHLD
AfsA      EAVLPHDHPFFAPVGDDLHDPLLVAEAMRQAAMLAFHAGYGIPLGYHFLLTELDYVCHPE
           *: *::  :   .  .  **** **::**:*:   *:   :*:*:**::: **   .

MmfL      VWDSGPLPVVLDVTSTKTTTNPRRMARALNADVYVAGLHR-------GRCAIRFEVLAPR
ScbA      HLGVSGEVAELEVEVACSQLKFRG-GQPVQGQVDW-AVRRAGRLAATGTATTRF--TSPQ
AfsA      YLGVGGEPTEIGLEVFCSDLKWRA-GLPAQGRVGW-AVHRGDRLAATGVAATRF--STPK
              .   . : :    :  : *  .   :. *   .::*       * .: **   :*:

MmfL      RYAMIRDRARRAERPAQQAAAGAATALPPETVGFHDDLHVLLATAQGLPDTAWQLRLRRD
ScbA      VYRRMRG-----DFATPTASVPGTAPVPAARAGRTRDEDVVLS--ASSQQDTWRLRVDTS
AfsA      AYRRMRG-----DVPVEGISLPETAPVPASPAGRARVEDVVLS--GTGREGVWELRVDTR
           *  :*      :  .   :   :: :*   .*     .*:*:      : .*.**:

MmfL      HPVLFDHESDHISGMALLEACRQAATALTPPAPGAFGPRQVALTAVASSYQAFGELDSPV
ScbA      HPTLFQRPNDHVPGMLLLEAARQAACLVTGPAPFV-------PSIGGTRFVRYAEFDSPC
AfsA      HLTLFQRPNDHVPGMLLLEAARQAACLVAGPAGIV-------PVEARTRFHRYSEFGSPC
          * .**:: .**: ** ****.****  :: **  .           : :  :.*: **

MmfL      TITTLPAAHGHSPDSGTRTLQLTARQGSRTLITATVTTTTTAGTGSPGPTVPHHGDQTKA
ScbA      WIQATV---RPGPAAGLTTVRVTGHQDGSLVFLTTLSGPAFSG-----------------
AfsA      WIGAVV---QPGTDEDTVTVRVTGHQDGETVFSTVLSGPRAHG-----------------
           * :       .      *:::*.:* .  :: :.::     *

MmfL      VAS
ScbA      ---
AfsA      ---
```

Figure 1.8. Alignment of the amino acid sequences of AfsA from *S. griseus* and the *S. coelicolor* AfsA homologues ScbA and MmfL. Identical and conservatively substituted amino acids are highlighted with asterisk and dots, respectively.

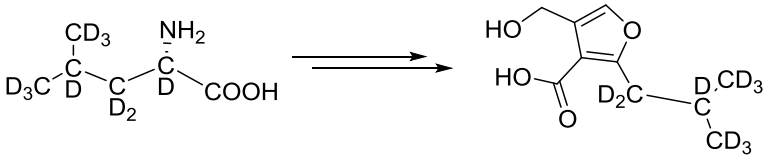
MmfL shares 27% identity and 43% similarity in amino acid sequence to AfsA.[75] Sequence alignment of MmfL with AfsA in *S. griseus* and ScbA in *S. coelicolor*, both of which are key enzymes involved in the biosynthesis of GBL-type signalling molecules, are shown in Figure 1.8.

It was proposed that MmfL catalyses an analogous reaction to AfsA, involving condensation of the β-ketoacyl thioesters derived from fatty acid metabolism with DHAP to form phosphorylated butenolide intermediates. It was also hypothesised that subsequent dephosphorylation, catalysed by MmfP, and further conversion catalysed by MmfH, yield the AHFCA molecules from the phosphorylated butenolide intermediates (Scheme 1.2).



Scheme 1.2. Proposed pathway for the biosynthesis of AHFCAs, taking AHFCA1 as an example.[75]

During the initial elucidation of the structures of AHFCAs, preliminary evidence has been provided for this hypothesis. Incorporation experiments using deuterium-labelled precursors for starter units of fatty acid biosynthesis showed that different fatty acid starter units are incorporated in each of the AHFCAs, yielding the different side chains. One example from the feeding experiments is shown in Figure 1.9.

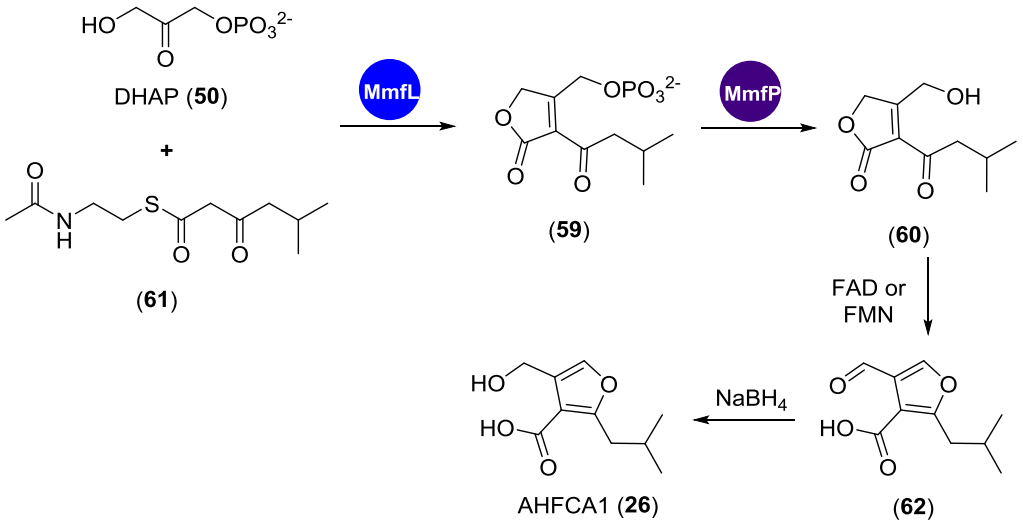


deuterium-labelled L-leucine deuterium-labelled AHFCA1 (**26**)

Figure 1.9. An example of the feeding experiments using deuterium-labelled precursors to yield the AHFCAs.

Davis and co-workers proposed that the divergence between GBL and AHFCA biosynthesis may arise due to the orientation of DHAP relative to the β-ketothioester precursors in the reactions catalysed by different AfsA-like enzymes.[76] This hypothesis implied that AHFCA biosynthesis would not proceed *via* a butenolide intermediate. However, subsequent incorporation experiments using ^{13}C-labelled glycerols labelled at the pro-*R* and pro-*S* carbon atoms confirmed the orientation of DHAP with respect to the β-ketothioester, thus supporting that AHFCA biosynthesis proceeds *via* a butenolide intermediate.[75]

More recently, the three enzymes, MmfL, MmfH and MmfP, were studied *in vitro* to establish their specific roles in the assembly of AHFCAs (PhD thesis by N. Malet).[77] Each enzyme was overproduced in *E. coli* and purified as recombinant His_6-tagged protein. MmfL was shown to catalyse the condensation of DHAP **50** and the synthetic NAC β-ketothioester **61** to give a putative phosphorylated butenolide product **59** (Scheme 1.3), which could not be detected due to its high polarity. Treatment of the product of the MmfL-catalysed reaction with alkaline phosphatase yielded a dephosphorylated product. However, experimental evidence is needed for structure confirmation of this dephosphorylated product as the butenolide **60**. MmfP has been demonstrated to be a Mg^{2+} dependent phosphatase capable of catalysing the dephosphorylation of the presumed phosphorylated butenolide **59** of the MmfL-catalysed reaction to yield butenolide **60** in *in vitro* enzymatic assays. Interestingly, FAD was shown to be capable of transforming the butenolide intermediate **60** into an aldehyde analogue **62** of AHFCA1, which could be reduced using sodium borohydride to yield AHFCA1 **26** (Scheme 1.3).[77] However, more work is needed to ascertain the function of each enzyme in AHFCA biosynthesis.



Scheme 1.3. Summary of the *in vitro* enzymatic studies on AHFCA biosynthesis by N. Malet.[77]

1.3 Transcriptional regulation of specialised metabolites in *Streptomyces*

1.3.1 TetR transcriptional repressors

Signalling molecules, such as GBLs and AHFCAs, are involved in regulation of corresponding specialised metabolite biosynthesis upon recognition and interaction with their cognate receptor proteins. Their receptors are known as the ArpA-like transcriptional repressor proteins. The ArpA class of the TetR family is an example of a transcriptional repressor involved in regulation of specialised metabolite biosynthesis. Such repressors are widespread in *Streptomyces* bacteria.

The TetR family of regulators act as repressors in transcription and belong to the one-component system in which a single polypeptide contains both a sensory domain and a DNA-binding domain.[78] This family is named after the member of this group that has been most completely characterised genetically and biochemically, the tetracycline resistance (TetR) protein, which was first discovered in *E. coli*.[79] TetR proteins generally function as homodimers, with each subunit consisting of a ligand-binding domain and a highly conserved helix-turn-helix (HTH) DNA-binding domain at the *N*-terminus. Members of this family are best known for their roles as regulators of antibiotic efflux pumps, which contribute significantly to antibiotic self-resistance mechanisms. Apart from this, they are also known to regulate numerous aspects of bacterial physiology,

including metabolism of amino acids, lipids, and other carbon and nitrogen-containing compounds.

TetR tightly regulates expression of *tetA* by binding specifically to the operator located in the *tetR-tetA* intergenic region (Figure 1.10). Tetracycline **63** binding to the repressor initiates conformational changes in TetR, thus reducing the affinity of its DNA binding domain for the *tetA* promoter within the *tetA-tetR* intergenic region. Dissociation of TetR from the DNA allows RNA polymerase to bind to the *tetA* promoter, resulting in transcription of *tetA*. The translated TetA protein functions as efflux pump for tetracycline, thus conferring self-resistance to tetracycline. Since TetR repressors are homodimeric, they often interact with semi-palindromic or dyad sequences.[80] The HTH DNA-binding domain binds to major grooves of DNA.

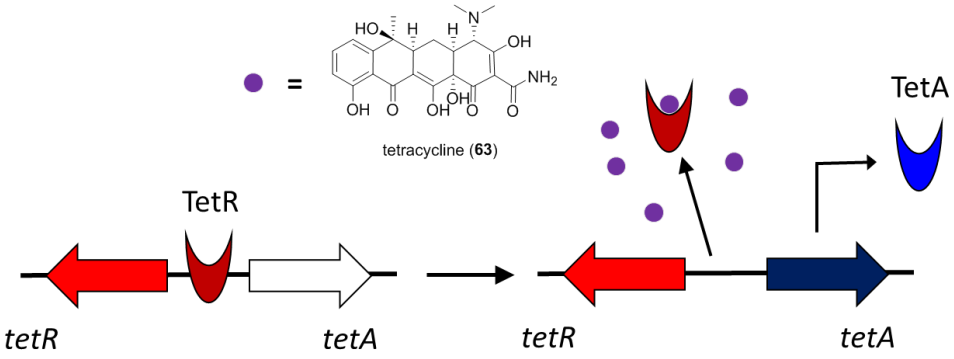


Figure 1.10. Mode of action of TetR. TetR repressor binds to the *tetA* promoter and represses *tetA* expression. The binding of tetracycline to TetR releases it from the promoter, thus allowing transcription of *tetA*. The TetA protein exports tetracycline out of the cell.

1.3.2 ArpA-like transcriptional repressors

As a cognate receptor for A-factor, which was discovered from *S. griseus* in the 1960s, the A-factor receptor protein (ArpA) from *S. griseus* was the first and most extensively studied ArpA-family repressor in *Streptomyces* bacteria. The ArpA class of proteins belongs to the TetR family and act as transcriptional repressors for regulation of specialised metabolites biosynthesis.

The ArpA-like repressors are widespread in *Streptomyces* bacteria. Similar to the tetracycline/TetR system, when the intracellular concentration of A-factor reaches a critical point, it binds to ArpA and causes conformational change in the protein, thus leading to release of ArpA from its cognate operator sequence, which is known as an

autoregulatory response element (ARE) located in the promoter of *adpA*. Therefore, it results in the transcription and translation of *adpA*. The expressed AdpA is a pleiotropic transcriptional activator implicated in the regulation of specialised metabolites production and morphological differentiation (Figure 1.11). In the streptomycin biosynthetic gene cluster of *S. griseus*, StrR functions as a pathway-specific transcriptional activator required to transcribe other genes within the cluster.[81] AdpA was identified to bind the operator region of *strR* and activate its expression, which leads to induction of streptomycin production.[82] Meanwhile, AdpA indirectly activates expression of *griR*, a pathway-specific activator gene that regulates biosynthesis of the yellow pigment grixazone.[83] It is also required for expression of other members of the *adpA* regulon, some of which are required for morphological differentiation.[84]

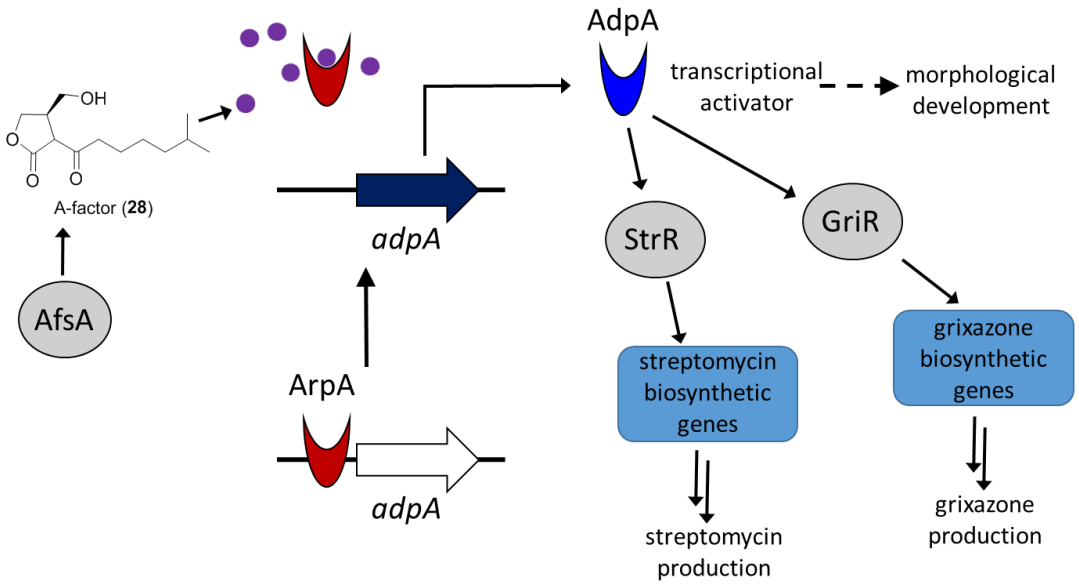


Figure 1.11. The A-Factor/ArpA regulatory cascade in *S. griseus* leading to transcription of *adpA* and subsequent morphological development and production of specialised metabolites, such as streptomycin and grixazone.

Analogous A-factor/ArpA systems in *Streptomyces* bacteria have been identified by genome mining, such as the AHFCA/SgnR/GbnR system regulating gaburedin biosynthesis in *S. venezuelae*,[59] the AHFCA/MmfR/MmyR system regulating methylenomycin biosynthesis in *S. coelicolor*,[45,58] and the SCB/ScbR/ScbR2 system controlling the biosynthesis of coelimycins in *S. coelicolor*.[71,85]

1.3.3 Regulation of methylenomycin biosynthesis in *S. coelicolor*

Methylenomycin A **16** (Figure 1.3), an unusual epoxycyclopentenone antibiotic, is one of the antibiotics produced by *S. coelicolor* A3(2) and its biosynthetic gene cluster, containing a total of 21 genes, are located on the 365 kb linear plasmid SCP1. Organisation of the methylenomycin gene cluster, including biosynthetic, resistance and regulatory genes, in *S. coelicolor* is shown in Figure 1.12. Methylenomycin was also reported to be produced by *Streptomyces violaceoruber* SANK 95570,[86] in which its gene cluster is more than 99% identical at DNA level to that in *S. coelicolor*.[20]

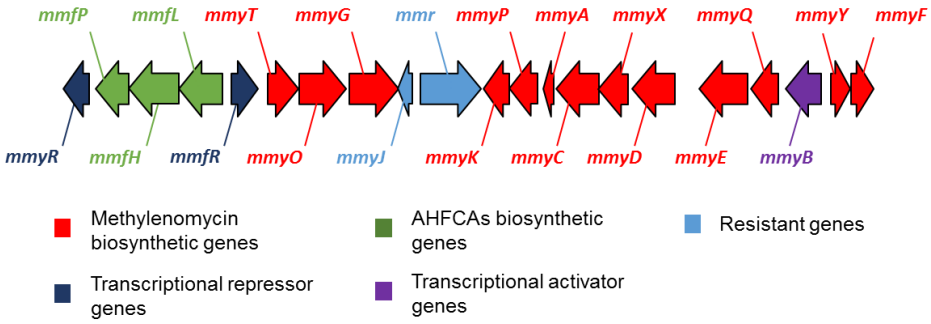


Figure 1.12. Organisation of the methylenomycin biosynthetic gene cluster in *S. coelicolor* A3(2).

It was noted that there are genes encoding for two repressors, *mmfR* and *mmyR*, in the methylenomycin (*mmy*) gene cluster in *S. coelicolor*. The products of both of them show sequence similarity to ArpA except that MmyR lacks two conserved residues that are present in ArpA that probably play a role in DNA- and ligand-binding. Deletion of these two repressors gives distinct phenotypes. Deletion of *mmyR* resulted in overproduction of methylenomycin, whereas deletion of *mmfR* had the opposite effect, causing substantial loss of production. However, the presence of the *mmfR* gene in the absence of *mmfLHP* led to greatly reduced methylenomycin production. Further studies showed that methylenomycin production was restored in the *mmfR*-containing strain when supplied with AHFCAs, whereas this was not observed in the *mmyR*-carrying strain. This indicates that MmfR, not MmyR, is sensitive to AHFCAs.[58] Thus, the cassette of five genes, *mmfR*, *mmfL*, *mmfH*, *mmfP* and *mmyR*, forms the AHFCA-dependent signalling system involved in the regulation of methylenomycin biosynthesis.

Similarly to the A-factor/ArpA regulation system, AREs are identified in the *mmy* gene cluster and known as MAREs, corresponding to a general consensus proposed by Folcher and co-workers as shown in Table 1.1.[87] These sequences are highly conserved short (18

bp) and almost palindromic DNA fragments, located in the intergenic regions of *mmfL-mmfR* and *mmyY-mmyB* as well as the region upstream of *mmyR*.

Table 1.1. ARE sequences in the *S. coelicolor* methylenomycin biosynthetic gene cluster and the *S. venezuelae* gaburedin biosynthetic gene cluster (W = A/T, Y = C/T).

Name	Intergenic region	Sequence (18 bp)
ARE consensus		AWACNNACYNNNCGGTTT
MARE1	*mmfR-mmfL*	ATACCTTCCCGCAGGTAT
MARE2	*mmyY-mmyB*	AAACCTTCGGGAAGGTTT
MARE3	upstream of *mmyR*	ATACCTTCCCGAGGGTAT
ARE1	*sgnR-sgnL*	ATACATTCTCGAAGGTAT
ARE2	upstream of *gbnA*	ATACCTTCGTGAAGGTTT
ARE3	*gbnR-sgnH*	ATACCTTCCAGACCGAAT

Binding of the repressor protein(s) to the MAREs is proposed to prevent the transcription of the adjacent genes. Previous electrophoretic mobility shift assays (EMSAs) have shown that MmfR is able to bind the *mmfL-mmfR* intergenic DNA sequence containing MARE1 and can be released by the addition of AHFCAs.[77] Recently, the crystal structure of MmfR with AHFCA2 bound has been solved by the Corre and Challis groups (Figure 1.13). This provides a clear picture of the mode of interaction of MmfR with AHFCA ligands. MmyR, which is 35% identical and 56% similar to MmfR in amino acid sequence, is also proposed to bind to the same ARE sequences. However, in contrast to MmfR, preliminary results showed that MmyR is not responsive to the same AHFCA signalling molecules.

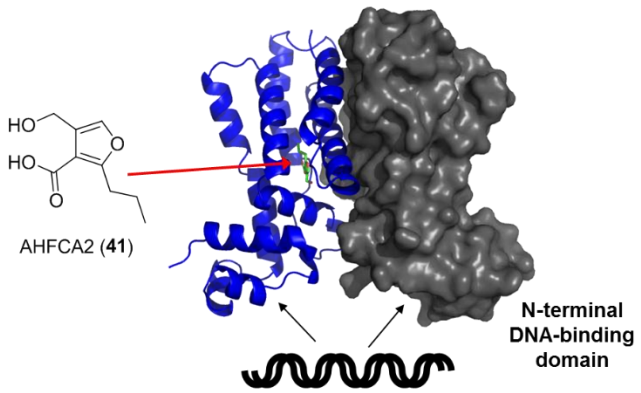


Figure 1.13. X-ray crystal structure of MmfR with AHFCA2 bound.

The *mmyB* gene encodes a protein with a likely DNA-binding motif related to that of the xenobiotic response element **(**Xre) family of transcriptional regulators. The in-frame deletion of *mmyB* results in the loss of methylenomycin production, suggesting that MmyB is a transcriptional activator for the biosynthesis of methylenomycin. It has been proposed that MmyB may be the direct target of the putative AHFCA/MmfR/MmyR regulation system.[58] From bioinformatics analyses of the *mmy* gene cluster, there are three short sequences (known as 'B-boxes') proposed to be the binding sites for MmyB, two of which are located in the operator regions of *mmyB-mmyY* and one situated upstream of *mmyT*.[58] These sequences are distinct from the ARE sequences. The putative function of MmyB is binding to the "B-box" sequences and activating transcription of the structural genes in the *mmy* gene cluster, including the putative *mmyTOG*, *mmyBQEDXCAPK* and *mmyYF* operons, leading to methylenomycin production. It was noted that both *mmfL* and *mmyB* genes contain the TTA codon, accounting for the dependence of methylenomycin production on *bldA*, which encodes the only tRNA for the rare UUA leucine codon. Therefore, BldA is required for translation of *mmfL* and *mmyB*. BldA is a global regulator of morphological differentiation and metabolite production, which predominantly accumulates during late exponential and stationary phase.[88]

A regulatory cascade has been proposed for the regulation of methylenomycin biosynthesis in *S. coelicolor* (Figure 1.14). Transcription of *mmyB*, *mmfL*, *mmyR* and itself is tightly regulated in the presence of MmfR repressor by binding to AREs. In presence of mature BldA, *mmfL* is translated to a functional protein and assembles the AHFCA signalling molecules. When the AHFCAs are produced to reach a critical concentration, they bind to the MmfR repressor and cause derepression of MmfR binding to the operator within the *mmfL-mmfR* intergenic region, leading to a significant increase of the production of AHFCAs. The increased concentration of AHFCAs leads to derepression of MmfR on the *mmyB* promoter and thus the transcriptional activator MmyB can be produced. MmyB then binds to its operator regions ('B-boxes'), initiating the transcription of the structural genes to produce methylenomycins. MmyR repressor is proposed to act as an off-switch to stop the production of methylenomycin compounds.

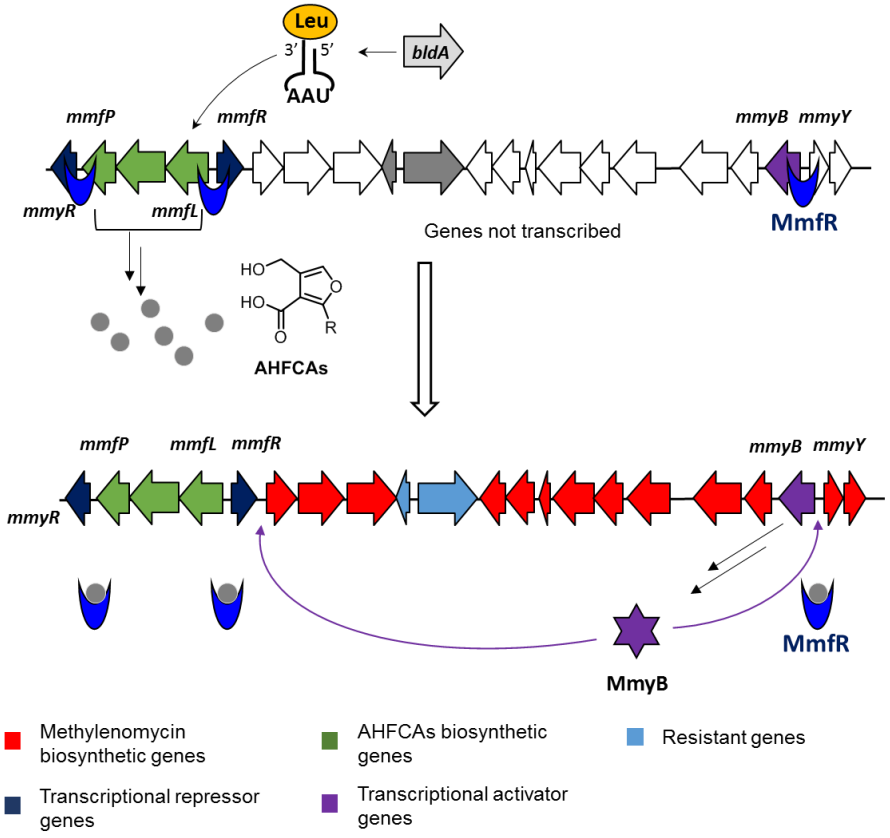


Figure 1.14. Proposed signalling cascade with MmfR involved in methylenomycin biosynthesis. MARE sequences for MmfR binding are located in the intergenic regions of *mmfL-mmfR*, *mmyY-mmyB* and the region upstream of *mmyR*. MmyR is also proposed to bind the same MAREs.

Analogous AHFCA-dependent signalling gene clusters have been identified from other *Streptomyces* bacteria, such as in *S. venezuelae*.[59] Sequence analysis of the gaburedin (*gbn)* gene cluster revealed the presence of the *mmyR* homologue *gbnR*, as well as the *mmfR* homologue *sgnR* and the *mmfLPH* homologues *sgnLPH*. Three highly conserved ARE sequences identified from the *gbn* cluster are almost identical in sequence to the ARE consensus from the *mmy* cluster and located in similar positions as those MAREs present in the *mmy* gene cluster, including the region upstream of *gbnA*, the intergenic regions of *sgnR-sgnL* and *gbnR-sgnH* as shown in Figure 1.15. These sequences are proposed to be the AREs for binding by the MmfR homologue SgnR. Genetic inactivation of the analogous *mmyR* repressor, *gbnR*, in *S. venezuelae* led to the discovery of a novel group of natural products, the gaburedins.[59]

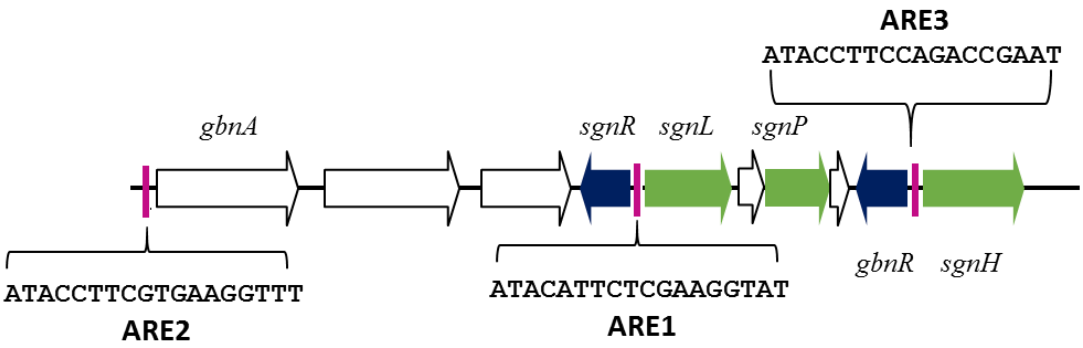


Figure 1.15. Location of AREs, which are shown in purple, in the *S. venezuelae* gaburedin biosynthetic gene cluster.

1.4 Siderophores

Siderophores are low-molecular-weight (< 10 kDa) iron-chelating compounds produced by grasses and microorganisms such as bacteria and fungi.[89-90] Their name is from the ancient Greek words *sidero* 'iron' and *phore* 'carriers' meaning 'iron carriers'. Iron is an essential element that plays an important role in a diverse array of biological processes for eukaryotes and prokaryotes. The indispensable nature of iron is due to its involvement in several key metabolic processes including amino acid synthesis, tricarboxylic acid cycle activity, DNA replication, cellular respiration and electron transport.[91] Due to the interconversion between the Fe^{3+} (ferric) and Fe^{2+} (ferrous) redox states, it is able to participate in single electron transfers and mediate common redox reactions as a versatile redox-active catalyst. However, free bioavailable iron in solution is extremely limited (10^{-8}-10^{-9} M) and below the range that supports microbial growth (~10^{-6} M).[92] This is due to its very low solubility under physiological conditions and consequent formation of ferric oxyhydroxide precipitates, though iron is actually one of the most abundant elements on Earth.

In response to iron limitation in their environment, many bacteria can produce different structural types of siderophores and excrete them into the extracellular environment, where the siderophore acts to sequester and solubilise the iron. The problem of iron deficiency is also sometimes faced by plants,[93] especially for those grown in calcareous soils with high pH, which decrease the affinity of plants for iron and thus impede iron uptake mechanism. A great variation can be seen in siderophore structures produced by bacteria, accompanied by different levels of iron chelating affinity. Siderophores usually

form a stable complex preferentially with Fe^{3+} compared to other naturally occurring abundant metal ions. Based on the co-ordinating groups that chelate the ferric iron, they are generally classified into several groups, including the most common catecholates, hydroxamates, carboxylates, and some containing distinct binding groups such as salicylic acid, oxazoline or thiazoline nitrogen, or having a mix of these chemically different classes of co-ordinating groups such as pyoverdins.[89]

The catecholate type of siderophore is characterised by the presence of phenolate or 2,3-dihydroxy benzoate binding groups. Examples include the three catecholate siderophores produced by *Azotobacter vinelandii* in iron-limited medium: the monocatecholate aminochelin **64**, the dicatecholate azotochelin **65** and the tricatecholate protochelin **66** (Figure 1.16).[94-97] A typical triscatechol siderophore is enterobactin **67**, a cyclic trimer composed of 2,3-dihydroxy-*N*-benzoylserine produced by bacteria of the family Enterobacteriaceae, including *Escherichia coli* and many others such as *Aerobacter aerogenes* and *Erwinia herbicola*.[98-99]

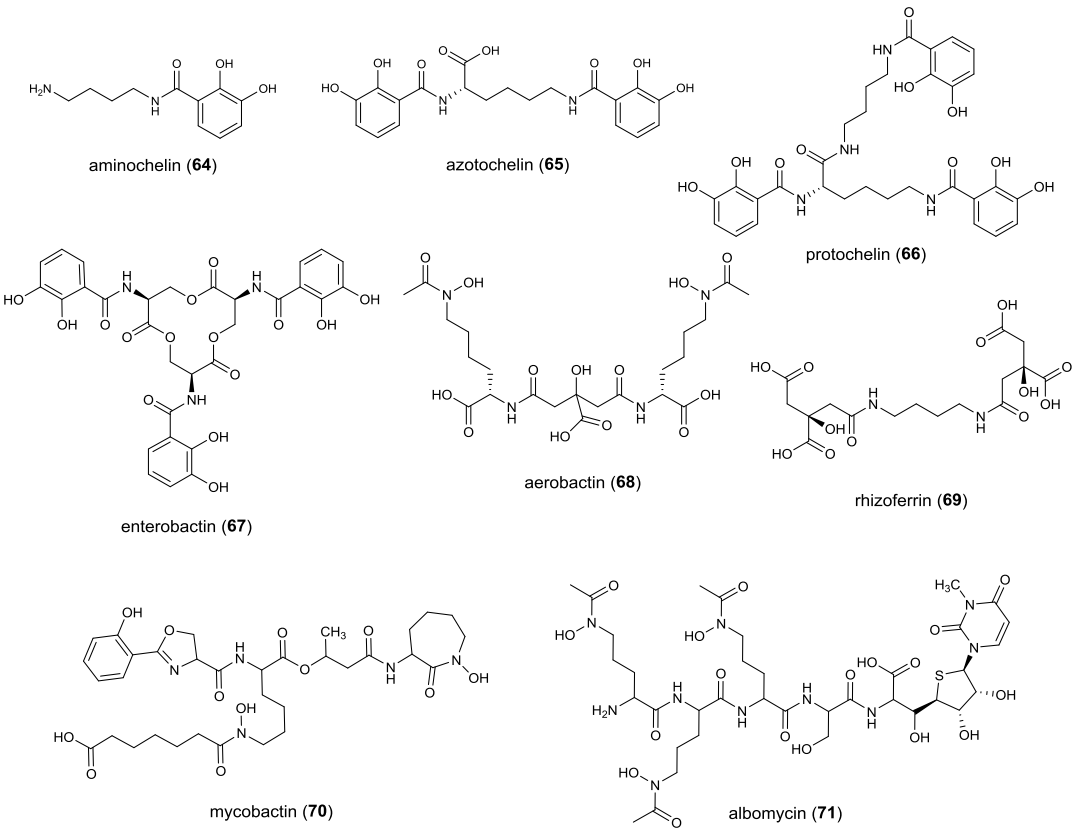


Figure 1.16. Representative examples of microbial siderophores.

Hydroxamate siderophores, produced mainly by fungi as well as some *Streptomyces* and *Pseudomonas* species, are mostly trihydroxamates, or dihydroxamates, leading to

formation of hexadentate, tetradentate and bidendate ligands. This type of siderophores include the trihydroxamate ferrichromes and the dihydroxamate aerobactin **68**, as well as the desferrioxamines mentioned previously (e.g., desferrioxamine E **17** in Figure 1.3) secreted by many *Streptomyces* species including *S. coelicolor* A3(2). Another type of siderophores chelate ferric iron by their hydroxyl carboxylate and carboxylate groups such as rhizoferrin **69**, which is produced by fungi, and its enantiomer, enantio-rhizoferrin produced in some bacteria.[100] In addition, certain siderophores have mixed ligands besides the above iron-binding groups, like mycobactin **70**, which is isolated from *Mycobacteria* with the presence of two hydroxamates, a phenolate and an oxazoline nitrogen in its structure (Figure 1.16).[101]

1.4.1 Bioactivity of siderophores

Microbial siderophores are not only important in microbial physiology but also have wide applications in various fields, such as agriculture to improve soil fertility and biocontrol, environmental applications and medicinal applications.[89] Clinically, siderophore-based drugs can be used to efficiently treat iron overload diseases. A siderophore from *Streptomyces pilosus*, desferrioxamine B, is marketed as mesylate salt under the trade name Desferol for removing excess iron from the body in thalassemia treatment. Remarkably, the potency of some antibiotics has been elevated by binding into the iron-binding functional groups of siderophores, for example, antimalarial siderophores combating iron-dependent malaria parasites, and sideromycins (e.g., albomycin **71** in Figure 1.16), a group of iron-chelating antibiotics produced by *Streptomyces* species exhibiting good antimicrobial activity.[102] Inspired by this, one promising strategy has been developed as so-called 'Trojan Horse' antibiotics, which are covalently bonded siderophore-antibiotic conjugates.[103] The siderophore part of the molecule can bind to ferric ion and be recognised by the membrane transport system, facilitating the uptake of the antibiotic warhead into the cell.

1.4.2 Biosynthesis of siderophores

There are two main pathways for siderophore biosynthesis, one of which involves the use of nonribosomal peptide synthetases (NRPSs) multienzymes while the other is NRPS independent.[104] NRPSs are large multi-modular enzymes that produce peptides in a sequential fashion, without the need for an RNA template. The order in which the constituent amino acids are incorporated into the peptide product is dictated by the order

of the NRPS modules.[105] An NRPS chain extension unit normally contains three domains: the adenylation (A) domain, the peptidyl carrier protein (PCP) domain, which is also known as the thiolation (T) domain, and the condensation (C) domain.

The A domain of each module of the NRPS recognises the substrate and catalyses the adenylation of its carboxyl group for activation by reaction with ATP. This activated ester is then covalently linked as its thioester on the PCP domain. It uses terminal thiol of a post-translationally installed phosphopantetheine arm to capture the activated carboxyl group of the adenylate. The C domain catalyses the direct transfer to another acyl amino acid intermediate on the adjacent downstream module to form a peptide bond. In some NRPS modules, C domain is replaced by a heterocyclisation (Cy) domain that catalyses heterocycle formation by a reaction of β-amino thiol group in the substrate attached to the PCP domain of the upstream module.

The growing chain is covalently bound to the PCP domain in successive modules throughout the assembly process. There is also a thioesterase (TE) domain, which is usually present in the final module. Hydrolysis or cyclisation results in the release of the assembled chain from the NRPS.

1.5 2-Hydroxyphenothiazolines

As well as having antioxidant, antibacterial and antifungal activity, many 2-hydroxyphenylthiazolines have iron-chelating properties and act as siderophores (Figure 1.17). They are therefore produced and secreted by some pathogens as virulence factors.[105-107] Members of the family include anguibactin **72**, a siderophore produced by *Vibrio anguillarum* 775, which is responsible for the septicemic disease vibriosis in fish,[108-109] yersiniabactin **73**, a siderophore that is an important virulence factor for some pathogenic *Yersinia* strains such as *Yersinia enterocolitica*[110] and pyochelin **74,** a siderophore first isolated from the Gram-negative *Pseudomonas aeruginosa*, which can opportunistically cause infections in plants and animals, including humans.[111] In *Pseudomonas* species, two siderophore-mediated systems are utilised for iron uptake, one mediated by the siderophore pyochelin and the other one mediated by the fluorescent siderophore pyoverdin.[89] Another member of the 2-hydroxyphenylthiazoline family, aerugine **75**, also identified from the *Pseudomonas* species, showed antifungal activity against *Colletotrichum orbiculare* as well as antioomycete activity against *Phytophthora capsici* and *Pythium ultimum*.[112-113]

Interestingly, *Pseudomonas protegens* (formerly *Pseudomonas fluorescens*) Pf-5 and CHA0 produce the enantiomer of pyochelin, enantio-pyochelin **76** (Figure 1.17), which has also been shown to function as a siderophore.[114] Some members of the 2-hydroxyphenylthiazoline family are produced by *Streptomyces* species, including pulicatins A-E **77**-**81**,[115] thiazostatin **82** and watasemycin **83** (Figure 1.17),[116-117] which are discussed in detail in Section 1.5.3. More recently, new members of this family, including ulbactins F/G **84**, have been isolated from a *Brevibacillus* sp. and shown to have anticancer activities.[118] Despite having a structure similar to pyochelin **74** and enantio-pyochelin **76**, it is not currently known whether these compounds are also able to function as siderophores.

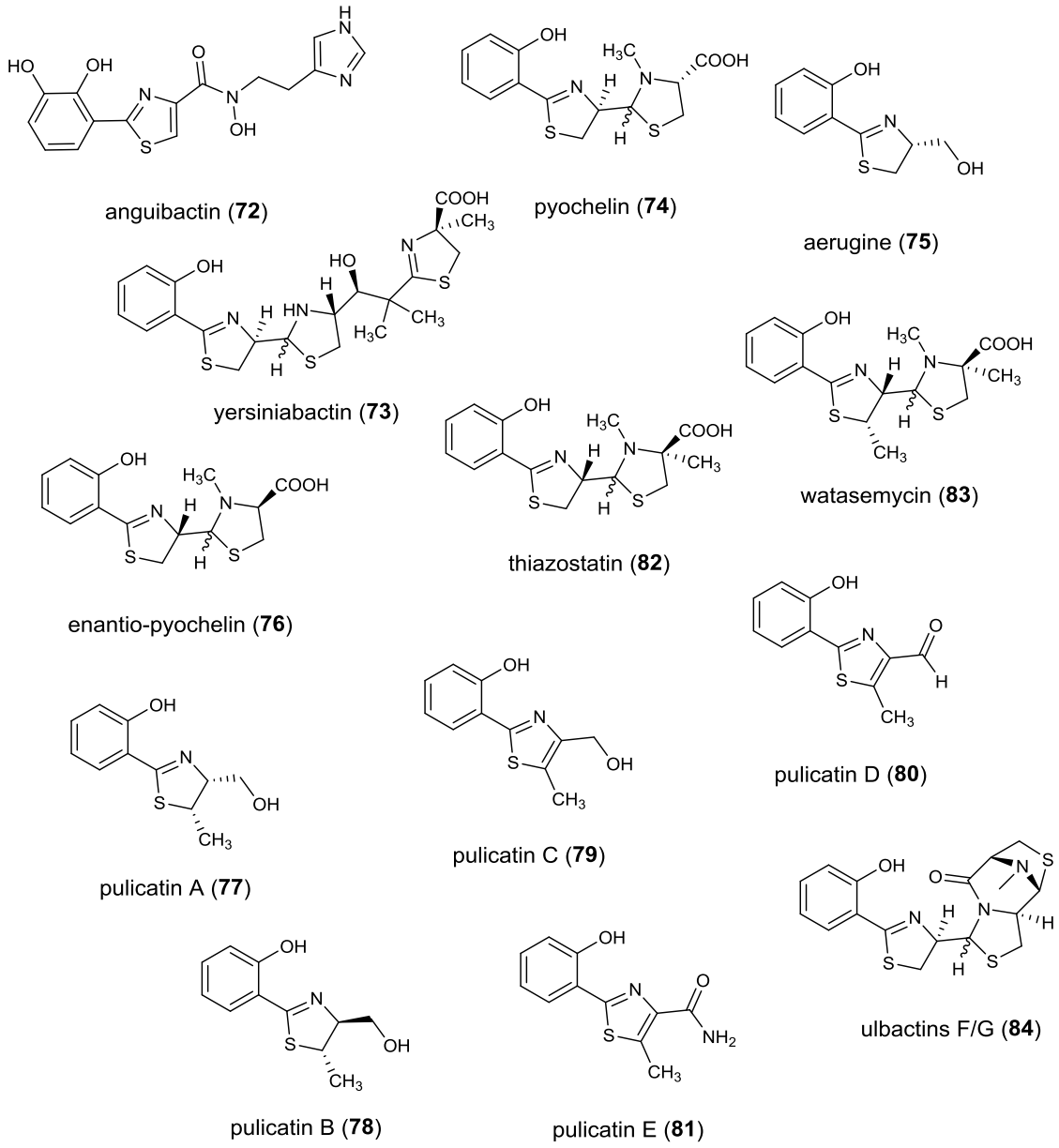


Figure 1.17. Representative members of the 2-hydroxyphenylthiazoline family compounds found in bacteria.

1.5.1 Pyochelin and enantiopyochelin

4'*R*, 4"*R*-Pyochelin **74** is a mixed-type siderophore containing a phenolate, a carboxylate, a thiazoline and a thiazolidine nitrogen atom. It was first isolated from *P. aeruginosa*,[111] and later found to be also produced by some members of the *Burkholderia cepacia* complex,[119] and plant-associated *P. fluorescens*,[120] as well as the Gram-positive *Streptomyces scabies* 87-22.[121] It should be noted that pyochelin is normally present as an equilibrium mixture of two diastereomeric forms due to isomerism at C-2". Apart from pyochelin, this also occurs in other thiazolidine derivatives including yersiniabactin

73,[110] ulbactins F/G **84**,[118] thiazostatin **82** and watasemycin **83** (Figure 1.17).[116-117] A possible mechanism for the epimerisation was proposed by Seto *et al.* for thiazostatin (Figure 1.18).[116]

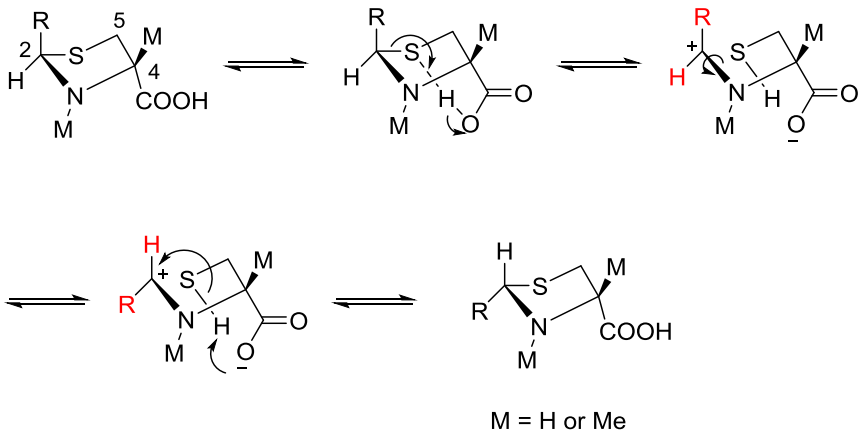


Figure 1.18. Proposed mechanism for epimerisation at C-2 of 2,4-disubstituted thiazolidine compounds.[116]

As one of the two important siderophore-mediated iron-uptake systems in *Pseudomonas* species, the structure, iron-chelating mode and biosynthesis of pyochelin have been clearly elucidated. The isolation of a 1:1:1 complex of Fe^{3+}, pyochelin and cepaciabactin, a bidendate ligand, suggests that four of the ligand sites of Fe^{3+} are satisfied by pyochelin while the remaining two bind cepaciabactin.[122] This has been further supported by the X-ray structure analysis of the 1:1 complex of pyochelin and Fe^{3+}, in which four of the six octahedral coordination sites of Fe^{3+} are occupied by the phenolate and the carboxylate oxygen and by the two nitrogen atoms of the two heterocycle rings while the remaining two octahedral sites are occupied by solvent molecules. It was found that the complexed form only corresponds to one of the two diastereoisomers, 4'*R*, 2"*R*, 4"*R*-pyochelin. This is also the case in another siderophore yersiniabactin **36**, which forms the Fe^{3+} complex only with one of its two stereoisomers containing the same *R* configuration at those chiral centres as pyochelin.[110,123]

Enantio-pyochelin **76**, the optical antipode of pyochelin produced from *P. protegens*, is found to promote growth under iron-limiting conditions in *P. protegens*, thus acting as a siderophore in the same way as pyochelin in *P. aeruginosa*. However, the interspecies studies indicate that enantio-pyochelin did not promote growth in *P. aeruginosa* and neither did pyochelin in *P. protegens*, suggesting that the pyochelin or enantio-pyochelin-mediated iron uptake machinery is highly stereospecific in both species.[114]

The reason underlying this is considered to prevent potential competitors occupying the same ecological niche from stealing heterologous iron-siderophore complexes.[124-125]

Interestingly, pyochelin has been shown to have antibiotic activity against several *Xanthomonas* species and *S. aureus*, which is not a result of iron competition, but due to its ability to generate reactive oxygen species (ROS).[126] However, analysis of the spectrum of the antibacterial activity of pyochelin revealed that bacteria unable to produce catecholate siderophores are sensitive to pyochelin while catecholate siderophore producers are resistant by the ability of catecholate siderophores, such as enterobactin **67** (Figure 1.16), to reduce ROS.[126] In addition, pyochelin has been investigated as a candidate for the Trojan horse prodrug strategy to develop potent antibiotics by conjugation with fluoroquinolones.[127]

1.5.2 Biosynthesis of pyochelin and enantio-pyochelin

The biosynthesis of both pyochelin **74** in *P. aeruginosa* and enantio-pyochelin **76** in *P. protegens* belongs to the NRPSs system, which has many similarities (Figure 1.19).

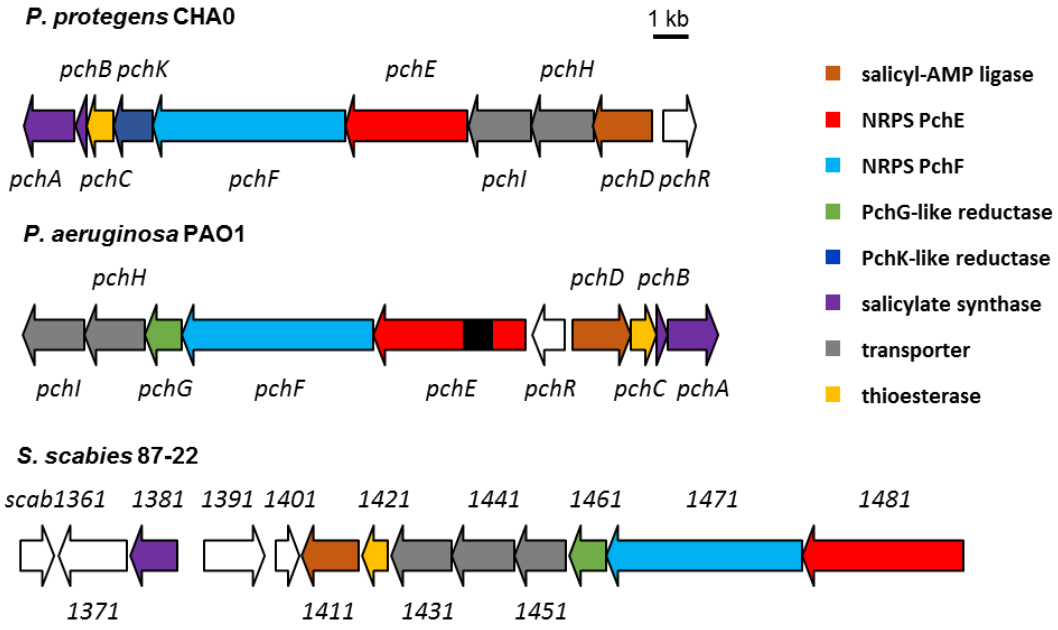


Figure 1.19. Organisation of the pyochelin biosynthetic gene clusters in *P. protegens* CHA0, *P. aeruginosa* PAO1 and *S. scabies* 87-22. Pyochelin biosynthetic gene homologues are colour-coded based on conserved function. Genes presumably involved in pyochelin export are grey, while non-homologous genes are white. The *pchE* gene is smaller in *P. protegens* CHA0 than in *P. aeruginosa* PAO1 due to the absence of an MT-like coding region (coloured in black in *pchE* in *P. aeruginosa* PAO1).

The biosynthetic genes for pyochelin **74** in *P. aeruginosa* are organised in two operons, *pchDCBA* and *pchEFGHI*, with *pchR* encoding a transcriptional regulator. It has been well established that pyochelin **74** is biosynthesized from salicylate and two molecules of L-cysteine (Figure 1.19 and 1.20).[128-133] Two enzymes, PchA, an isochorismate synthase, and PchB, an isochorismate pyruvate-lyase, are responsible for the biosynthesis of salicylate, which is then adenylated by the A domain, PchD. The gene *pchC* encodes an external thioesterase, which belongs to the type II TE family.[134] PchC has been demonstrated *in vivo* to play a proofreading role by removing wrongly charged molecules from the PCP domains of PchE and PchF.[133] The adenylated salicylate is loaded onto the N-terminal PCP domain of the peptide synthetase PchE and condensed with an activated L-cysteine, which is linked at the C-terminus of PchE. This process, including condensation, epimerisation and cyclisation reactions catalysed by PchE, leads to the formation of the first thiazoline moiety. The peptide synthetase PchF catalyses condensation of the PchE-bound 2-hydroxyphenylthiazolinyl intermediate with a second molecule of L-cysteine. PchF catalyses an analogous reaction to PchE, resulting in the second thiazoline ring. Reduction by the thiazoline reductase PchG and *N*-methylation by the MT domain of PchF gives the pyochelin scaffold. It should be noted that the first L-cysteine molecule incorporated is converted to its D-isoform by the methyl transferase (MT)-like domain in PchE while the second cysteine retains its L-configuration, leading to pyochelin **74** with 4'*R* and 4''*R* configuration on the thiazoline and thiazolidine rings, respectively. The internal TE domain in PchF catalyses hydrolysis to release pyochelin **74** from the biosynthetic enzymes.

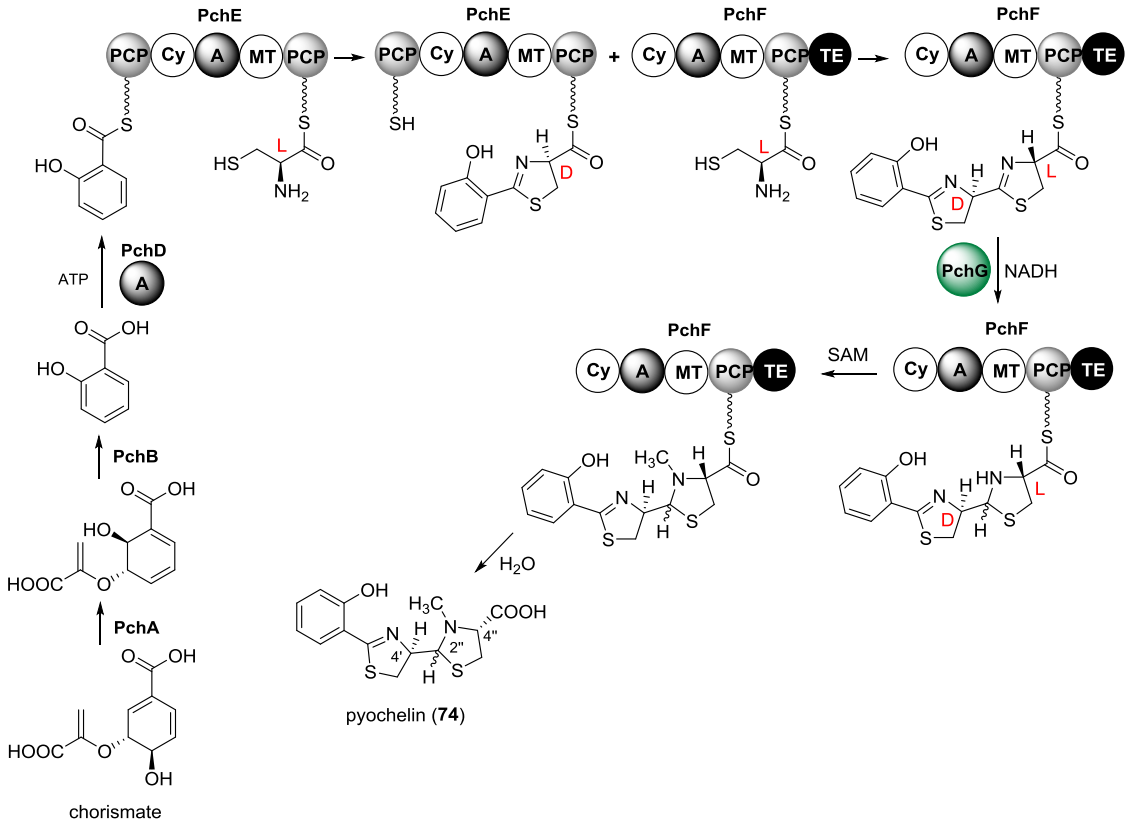


Figure 1.20. Proposed pathway for the biosynthesis of pyochelin in *P. aeruginosa*.

The biosynthetic pathway of enantio-pyochelin **76** in *P. protegens* is very similar to pyochelin **74** biosynthesis in *P. aeruginosa*, utilising homologues of the six proteins PchABCDEF (Figure 1.19 and 1.21),[114] although the organisation of the two biosynthetic gene clusters is slightly different. Their main differences focus on the following two aspects: i) PchK, which is proposed to be the thiazoline reductase responsible for reduction of the second thiazoline, is not homologous to the reductase PchG that is present in the *P. aeruginosa* pyochelin biosynthetic gene cluster. However, the underlying mechanism of reduction (and epimerisation) by PchK, which leads to the stereocenter with D-configuration in the second thiazoline ring, is still unknown, possibly also involving the use of other enzyme(s). ii) The PchE homologue in *P. protegens* is smaller than in *P. aeruginosa* due to the absence of an MT-like domain, which is believed to be the epimerisation region responsible for creation of the D-configured stereocenter in the first thiazoline of pyochelin. Collectively, these differences explain why *P. protegens* produces enantio-pyochelin **76** with the opposite stereochemistry to pyochelin **74**. It is noteworthy that the PchR regulators have been proposed to be one of the

determinants for the stereospecificity of iron uptake mediated by pyochelin and enantio-pyochelin in *Pseudomonas* species.[135]

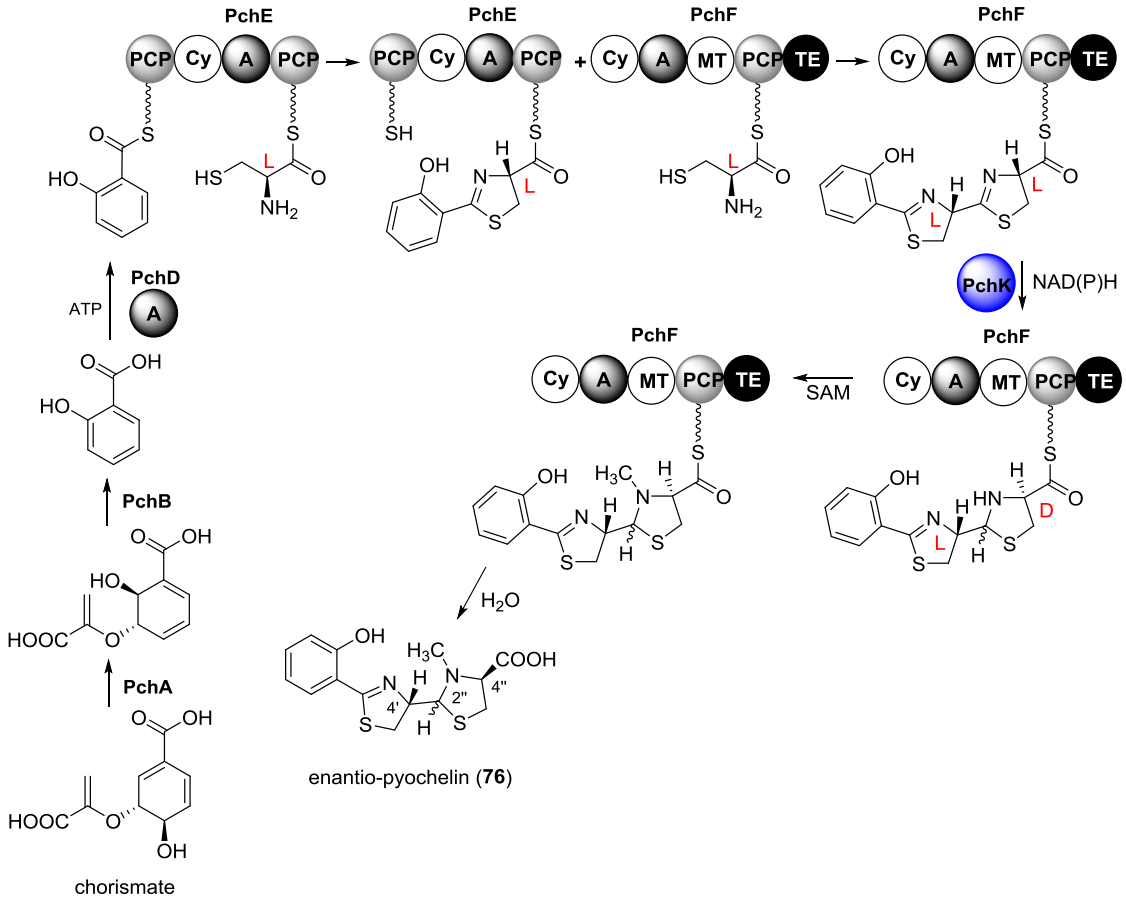


Figure 1.21. Proposed pathway for the biosynthesis of enantio-pyochelin in *P. protegens*.

Pyochelin **74** is also produced by a Gram-positive bacterium, *S. scabies* 87-22, the biosynthetic cluster of which contains six proteins, SCAB1381, SCAB1421, SCAB1411, SCAB1481, SCAB1471 and SCAB1461, functionally related to PchABCDEFG in *P. aeruginosa*.[121] SCAB1381 is proposed to be the salicylate synthase responsible for the formation of the salicylate moiety from chorismate. This is in contrast to the pathways in *Pseudomonas* species, in which two enzymes PchA and PchB are required. The other difference between two species is, a TetR-family protein (SCAB1401) and an AfsR-family protein (SCAB1371) are involved in the transcriptional regulation of the pyochelin biosynthetic gene cluster in *S. scabies* 87-22, whereas the *Pseudomonas* species utilises an AraC-family protein (PchR).[121]

1.5.3 Pulicatins, thiazostatin and watasemycin

Pulicatins A-E **77-81** (Figure 1.17) is a group of thiazoline metabolites isolated from a cone snail *Conus pulicarius* associate, *Streptomyces* sp. CP32, and have been reported to have neuroactive activities. The absolute stereochemistry of pulicatin A **77** and B **78** has been fully elucidated based on studies using nuclear overhauser effect spectroscopy (NOESY), circular dichroism (CD) and their derivatization to Mosher's esters, namely (*R*)-(+) or (*S*)-(-) α-methoxy-α-(trifluoromethyl) phenylacetyl ester.[115] The configuration of the stereogenic centre C-5' has been established as *S* as shown in Figure 1.17.

Thiazostatin **82** is a close analogue of pyochelin which differs only by an additional methyl group at C-4". It was first reported in 1989 to be isolated from *Streptomyces tolurosus* 1368-MT1 during screening for antioxidant substances.[116] As other 2,4-disubstituted thiazolidine compounds, it presents as a pair of diastereoisomers due to epimerisation at C-2", with antioxidant activity. However, thiazostatin was not reported with antimicrobial or iron-chelating activity since its discovery, probably due to lack of published data on this compound.

Watasemycin **83**, which possesses a methyl group at C-5' of its thiazoline ring instead of a hydrogen atom, is structurally closely related to thiazostatin and pyochelin with similar epimerisation at C-2". This compound was first reported in 2002 by Furumai *et al.* to be isolated from *Streptomyces sp.* TP-A0597, accompanied by isolation of thiazostatin **82**.[117] Watasemycin has been reported with modest antibiotic activity against Gram-positive and negative bacteria and yeast. Recently, watasemycin and thiazostatin, as well as some other 2-hydroxylphenylthiazolines, have been identified to be the metabolic products of a biosynthetic gene cluster from *S. venezuelae* ATCC10712.[136] More details about investigations into watasemycin biosynthesis in *S. venezuelae* are described in Chapter 4.

The relative stereochemistry of thiazostatin and watasemycin has been proposed based on their NOESY studies. The proposed configuration for thiazostatin is 4'*S*, 4"*R* or 4'*R*, 4"*S* while for watasemycin is 4'*S*, 5'*S*, 4"*R* or 4'*R*, 5'*R*, 4"*S* (Figure 1.22).[117] However, their absolute configuration hasn't been determined yet.

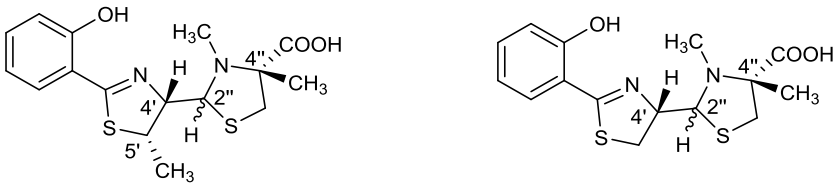


Figure 1.22. Reported relative stereochemistry of watasemycin and thiazostatin.

1.6 *S*-Adenosylmethionine (SAM)-dependent methylation

S-Adenosylmethionine (SAM) is a versatile small molecule used in many biological reactions (Figure 1.23). It is a commonly used methyl donor in numerous biologically important methylation reactions, including DNA, RNA and protein methylation, as well as in the biosynthesis of bacterial specialised metabolites. The 3-amino-3-carboxypropyl group of SAM can also be transferred to different acceptor molecules, such as RNA and proteins. In addition to serving as a methyl donor and a 3-amino-3-carboxypropyl donor, SAM is also involved in several other types of group transfer reactions, e.g., the adenosyl group transfer in fluorinase and chlorinase.[137-139]

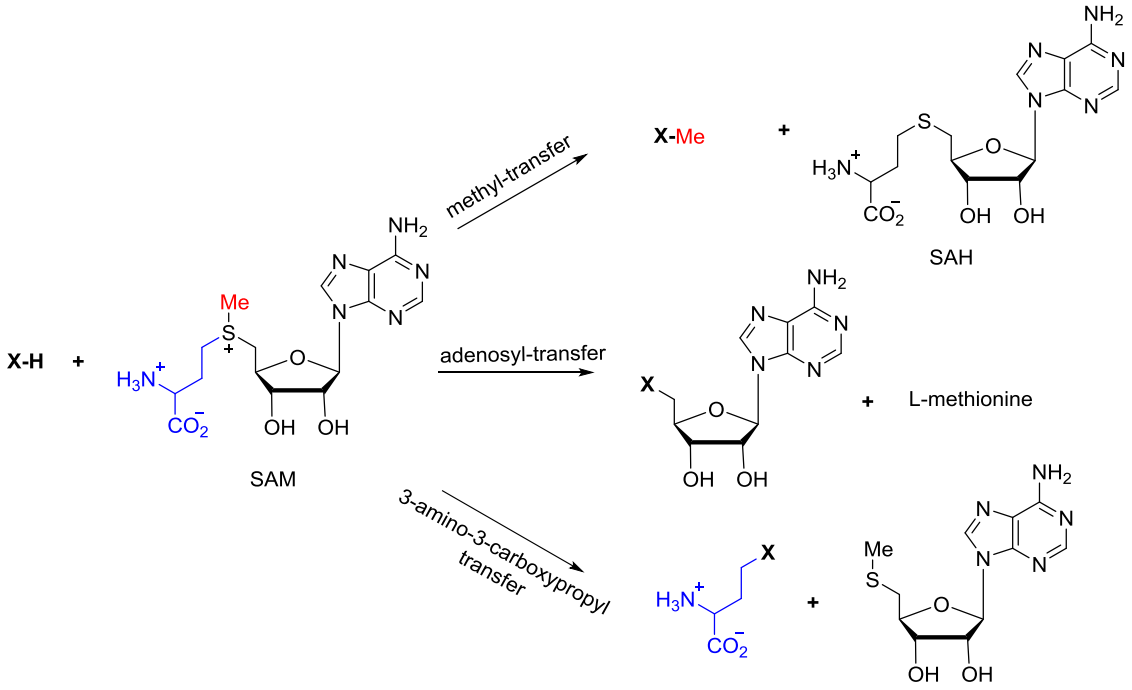


Figure 1.23. Three typical SAM-dependent reactions with methyl, 3-amino-3-carboxypropyl or adenosyl group transfer.

In classical SAM-dependent methylation reactions, a simple S_N2-type nucleophilic substitution mechanism is utilised, in which a nucleophile such as a heteroatom (N, O or

S), directly attacks the methyl group attached to the positively charged sulfonium ion in SAM, with concomitant formation of *S*-adenosyl homocysteine (SAH). However, SAM-dependent C-methylations are more complicated and have been proposed to involve different mechanisms. Substrates that are more nucleophilic or easily converted to a nucleophile can be methylated *via* simple nucleophilic mechanisms, for example, the SAM-dependent formation of cyclopropane on unsaturated fatty acids in bacteria or enzyme-assisted nucleophile formation in DNA cytosine C-5 methylation.[140] The methyl group at C-4" of watasemycin and thiazostatin is presumably generated by nucleophilic methylation of the α-carbon of cysteine with SAM.[117] The mechanism of introducing a methyl group at the inert C-5' position of watasemycin is unclear due to the fact that nucleophilic substitution is not likely to take place at such an unreactive carbon centre.

More recently, it has become clear that SAM is also able to methylate inert carbon or phosphorous atoms *via* radical intermediates, which probably involves the use of additional cofactors. These reactions are known to be catalysed by radical SAM (RS) enzymes,[141] which initiate a diverse set of reactions *via* generation of a radical intermediate. One molecule of SAM receives an electron transferred from the reduced [4Fe-4S] cluster and undergoes homolytic cleavage of the C5'-S bond to generate the 5′-deoxyadenosyl (5′-dAdo) radical (Figure 1.24). This initiation step is common to all RS methylases. Several reviews have summarised these RS methylases and classified them into different groups, mainly depending on their protein architectures, cofactor requirements and mechanisms of catalysis.[141-143] So far, five classes of RS methylases have been described, including class A, B, C, D and a new class E, which was identified very recently.[142-144]

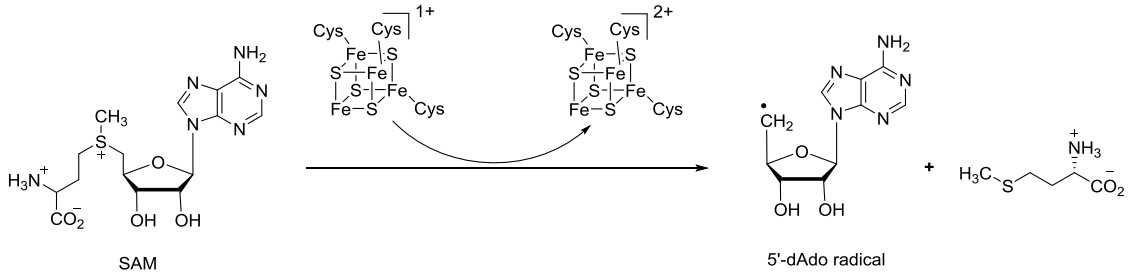


Figure 1.24. Generation of 5′-dAdo radical intermediate from SAM as an initiation step for methylation reactions catalysed by RS methylases.

Class A RS methylases are characterised by two strictly conserved cysteine residues and methylation is carried out by a protein-based methylene radical, resulting in only two of

the three protons from the methyl group of SAM being transferred to the product. These methyltransferases perform methylations on sp^2-hybridised carbon centres and are commonly involved in RNA methylation, e.g., RlmN and Cfr,[145-146] which are best characterised examples of this class to catalyse methylation on adenosine of RNAs.

Class C methylases are annotated with a second domain having high sequence similarity to coproporphyrinogen III oxidase, in addition to the SAM-binding domain, and catalyse methylation of sp^2-hybridised carbon atoms in aromatic heterocycles. These enzymes appear to be involved in the biosynthesis of many antitumor and antibiotic bacterial metabolites such as bleomycin and nosiheptide.[141]

The only example of class D enzyme is MJ0619 in methanopterin biosynthesis, which uses methylenetetrahydrofolate as an alternative methyl group source.[147] A family of NifB proteins have been reported by Hu *et al.* to catalyse the essential step of radical SAM-dependent carbide insertion in the assembly of nitrogenase metallocofactor.[144] This group of enzymes has been treated as a new class of RS enzymes, the class E methylases.

1.6.1 Class B radical SAM methylases

Class B RS methylases are of particular interest among all RS methylases due to their ability to catalyse methylation on phosphinate phosphorous atoms and sp^2- and sp^3-hybridised carbon centres, representing the largest and most versatile class of RS methylases.[141,148] The methylation reactions catalysed by class B enzymes are involved in the biosynthesis of a wide variety of microbial natural products, including aminoglycosides (e.g., gentamycin G418 **85** and fortimicin A **86**),[149-150] β-lactams (e.g., thienamycin **87**),[151] polyketides (e.g., chondrochloren B **88**),[152] ribosomally biosynthesised and posttranslationally modified peptides (RiPPs) (e.g., thiostrepton A **89**, bottromycin A2 **90** and polytheonamide **91**),[153-158] nonribosomally biosynthesised peptides (NRPs) (e.g., watasemycin **83**),[136] phosphonates (e.g., phosphinothricin **10**),[159] enediynes (e.g., calicheamicin **92** and kedarcidin **93**),[160-161] aminocoumarins (e.g., clorobiocin **94**) and terpenes (e.g., 2-methylhopene **95**) (Figure 1.25).[162-163]

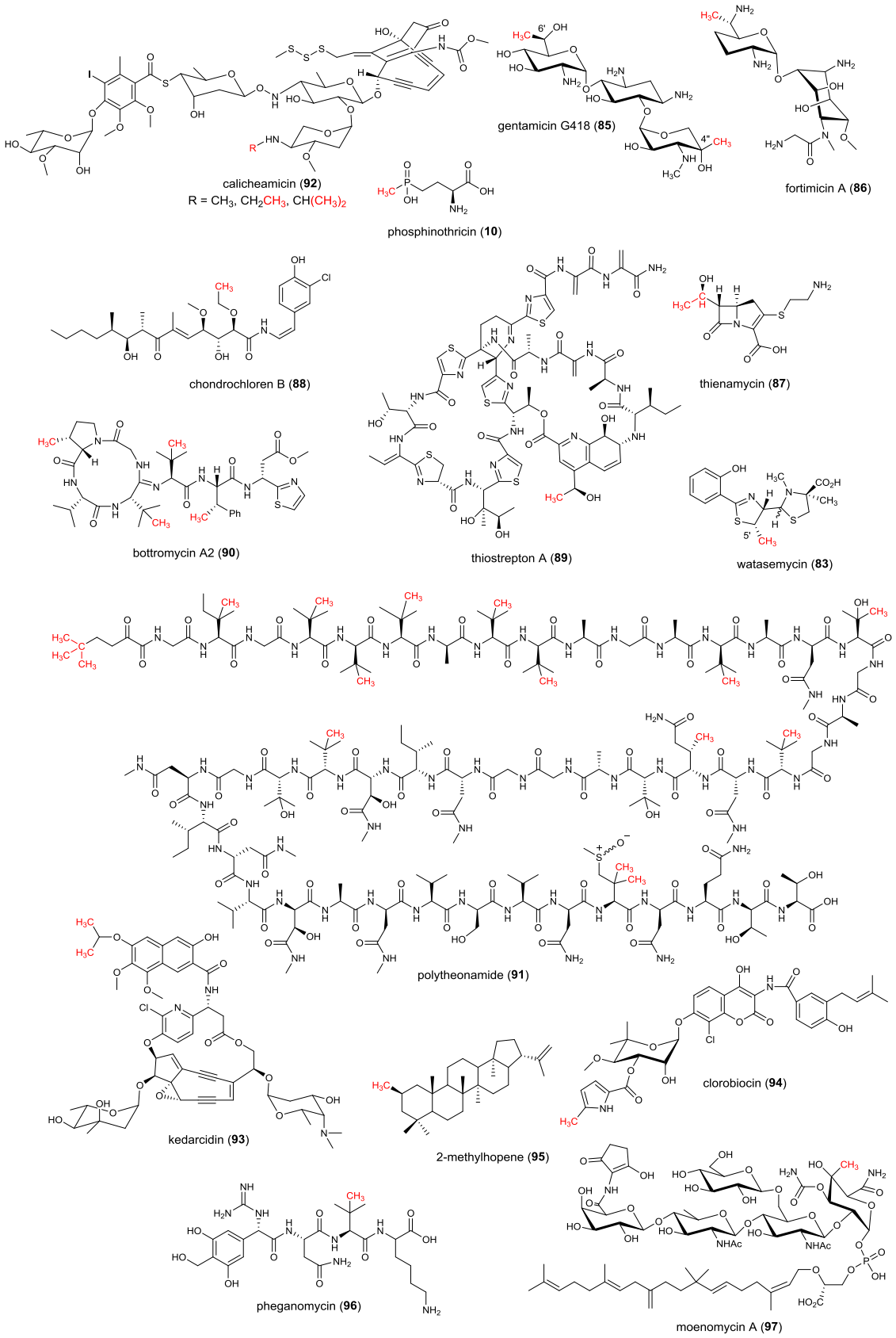


Figure 1.25. Specialised metabolites for which class B RS methylases are known or predicted to be involved in their biosynthesis, with corresponding functional groups coloured in red.

These enzymes are distinguished in architecture from other classes of RS methylases by an *N*-terminal methylcobalamin (MeCbl)-binding domain in addition to the core radical SAM domain. The most common reactions catalysed by class B RS methylases involve the methylation of un-activated sp^3-hybridised carbon centres. Examples include the elongation of a methoxy group to an ethoxy group, catalysed by CndI in chondrochloren B **88** biosynthesis, and the β-position methylation of the valine moiety catalysed by a putative class B enzyme in pheganomycin **96** biosynthesis (Figure 1.25).[152,164]

The catalytic mechanism for class B RS methylases is proposed as shown in Figure 1.26. Same as other RS methylases, the 5'-dAdo radical is firstly generated from the common initiation step utilising a [4Fe-4S] cluster. Then the 5'-dAdo radical abstracts a hydrogen atom from the target carbon centre to give a substrate radical, which attacks the methyl moiety of MeCbl, leading to homolytic cleavage of the C-Co bond. This results in the production of cob(II)alamin and the product modified at the methylating position. The inactive cob(II)alamin can be returned to MeCbl by obtaining an electron and subsequently reacting with another molecule of SAM.

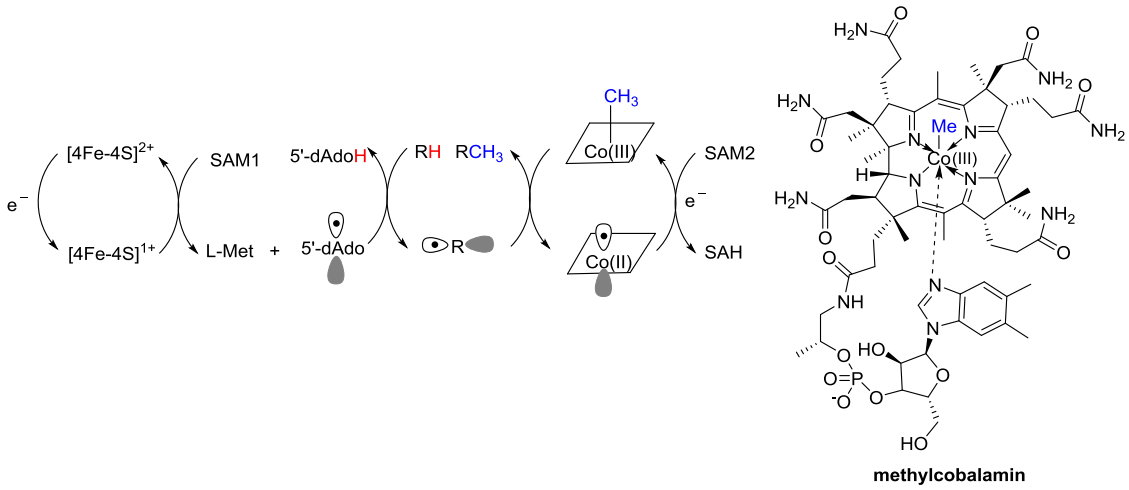


Figure 1.26. Proposed catalytic mechanism of class B RS methylases (left) and the structure of MeCbl (right).[143,148]

1.6.1.1 Stereochemical course of C-methylation by class B radical SAM methylases

It has been proposed that C-methylations catalysed by class B RS methylases proceed with inversion of configuration at the reacting carbon centre.[148] Clear examples of this process include the GenD1-catalysed methylation at C-4" in gentamicin biosynthesis and

the methylation by MoeK5 in moenomycin **97** biosynthesis (Figure 1.25).[149,165] The inverted stereochemistry suggests that SAM and MeCbl are positioned on opposite sides of the substrate in the active site of class B RS methylases (Figure 1.27).[148] Such an arrangement dictates that the substrate radical must undergo inversion of configuration before it can attack MeCbl. Further details using stereospecifically isotope-labelled precursors to decipher the stereochemical course are discussed in Chapter 4.

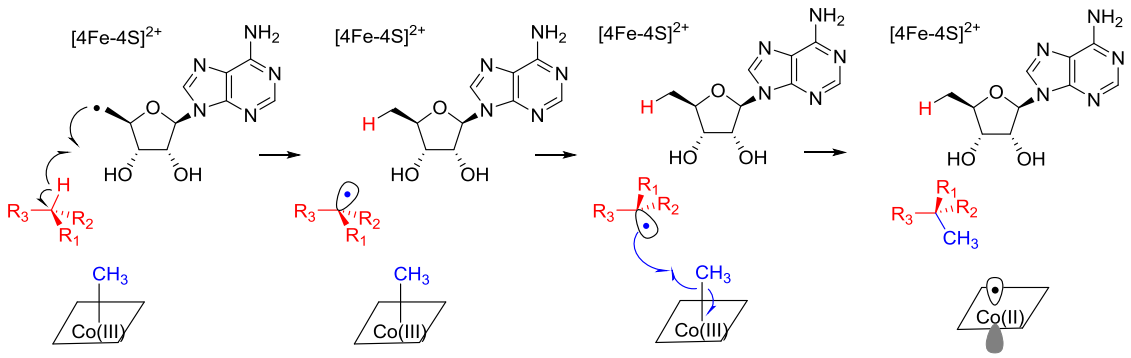


Figure 1.27. Proposed organisation of the active site in class B RS methylases.[148]

1.7 Aims of project

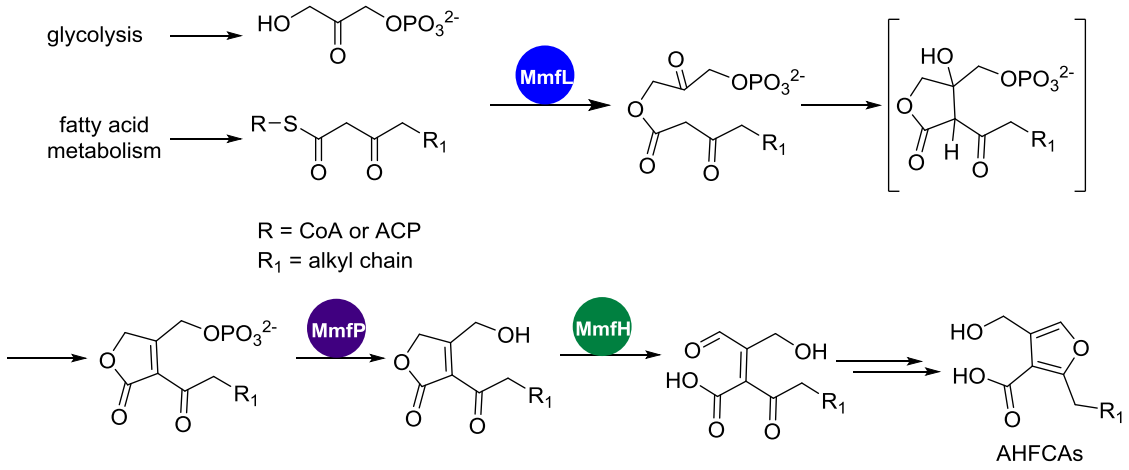
This work had three main aims: i) to establish the role of MmfL in AHFCA assembly in *S. coelicolor* A3(2); ii) to study the regulation mechanism of specialised metabolite biosynthesis by AHFCAs and ArpA-like repressor proteins in *Streptomyces*; iii) to characterise the watasemycin biosynthetic pathway in *S. venezuelae* ATCC10712. The following objectives were pursued.

- Investigate the function of MmfL *in vitro* for AHFCA assembly by synthesising a synthetic standard of the putative dephosphorylated butenolide product of the MmfL-catalysed reaction.
- Study the *in vitro* interaction between MmfR, its cognate operator sequences and AHFCAs or SCB1 using EMSAs.
- Identify the key residues involved in the ligand recognition by MmfR using site-directed mutagenesis based on analysis of the X-ray crystal structure of MmfR with AHFCA2.
- Investigate the *in vitro* interaction between the MmfR homologue SgnR, its cognate operator sequences and corresponding AHFCA signalling molecules from the gaburedin biosynthetic gene cluster in *S. venezuelae*.
- Identification and characterisation of the pyochelin-like metabolic product from the watasemycin biosynthetic gene cluster in *S. venezuelae* ATCC10712.
- Establish the function of several enzymes, including Sven0508, Sven0516 and Sven0515, in the biosynthesis of 2-hydroxylphenylthiazolines in *S. venezuelae*.
- Study the stereochemical course of the methylation at C-5' in watasemycin biosynthesis.

2. Investigation of the role of MmfL in AHFCA biosynthesis

2.1 Introduction to previous studies

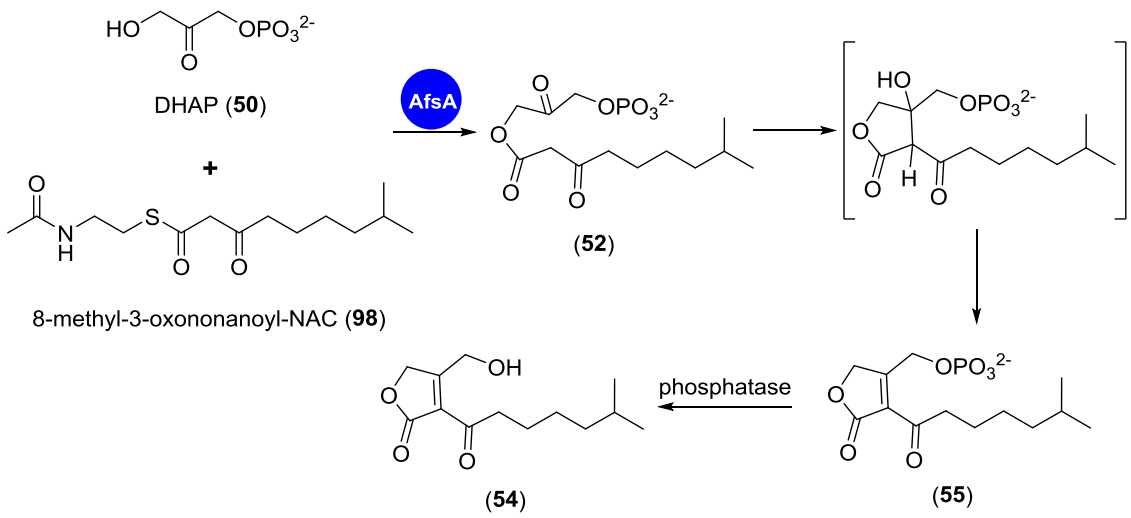
MmfL, MmfP and MmfH are a cassette of three enzymes proposed to be responsible for the biosynthesis of AHFCA signalling molecules that control expression of the methylenomycin biosynthetic gene clsuster, as discussed in Section 1.2.3. During the biosynthetic process, MmfL is proposed to catalyse an analogous reaction to AfsA,[67] including the condensation of various ACP-bound β-ketothioester intermediates in fatty acid biosynthesis with DHAP derived from glycolysis, followed by spontaneous intramolecular aldol condensation and dehydration, resulting in the corresponding phosphorylated butenolides. Subsequent dephosphorylation of the butenolides catalysed by MmfP, followed by MmfH-catalysed oxidation of the butenolide would yield an acyclic aldehyde intermediate, which undergoes cyclisation, dehydration and reduction to yield the AHFCAs (Scheme 2.1).



Scheme 2.1. Proposed biosynthetic pathway for AHFCAs, involving different β-ketoacyl-ACP thioesters.

Based on the sequence similarity between MmfL and AfsA as shown in Figure 1.8 in Chapter 1, as well as the work reported by Kato *et al*.,[67] the role of MmfL in AHFCA biosynthesis may be similar to AfsA in the γ-butyrolactone A-factor biosynthesis in *S. griseus*. During investigation into the function of AfsA, Kato and co-workers used an 8-methyl-3-oxononanoyl-NAC **98** to mimic the corresponding β-ketoacyl-ACP as the substrate for AfsA *in vitro* (Scheme 2.2). The enzymatic reaction of AfsA with NAC β-ketothioester **98** and DHAP **50** results in the 8-methyl-3-oxononanoyl-DHAP ester **52**, which was confirmed by comparison to a synthetic standard. The product of the AfsA reaction is proposed to undergo an aldol condensation and dehydration nonenzymatically

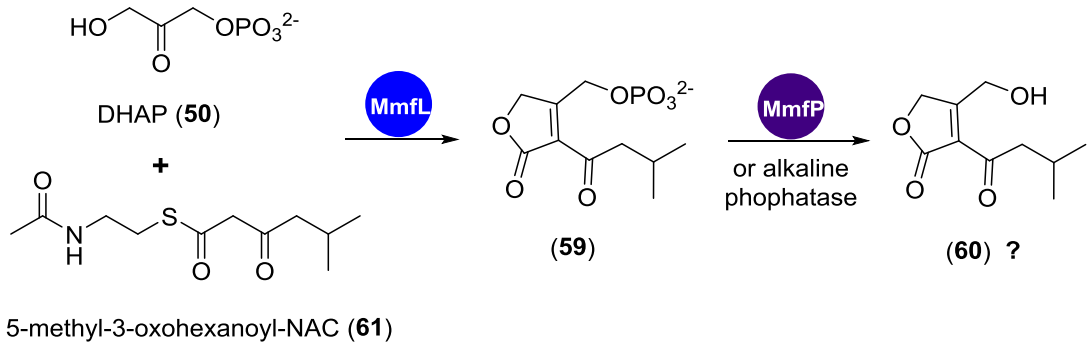
to afford the phosphorylated butenolide **55**, which can be observed by LC-MS as the dephosphorylated form **54**, after phosphatase treatment.



Scheme 2.2. Enzyme assays performed by Kato and co-workers for investigation of AfsA function.[67]

Initial investigations into the biosynthetic pathway of AHFCAs were undertaken in order to establish the specific roles of MmfL, MmfP and MmfH in AHFCA assembly (work by N. Malet).[77] The three enzymes were overproduced in *E. coli* and purified as recombinant *N*-terminal His_6-tagged proteins. The reaction catalysed by MmfL is proposed to be the first step in AHFCA biosynthesis, resulting in a phosphorylated butenolide intermediate.

Enzymatic reactions were performed for His_6-MmfL using a synthetic substrate, the NAC β-ketothioester **61**, which is a mimic of the proposed β-ketoacyl-ACP thioester precursor of AHFCA1 (Scheme 2.3). The MmfP-mediated dephosphorylation could be replaced by treatment with alkaline phosphatase *in vitro*, similar to that used in the AfsA study. The putative phosphorylated butenolide product **59** could not be detected due to its high polarity. Further treatment of **59** with alkaline phosphatase yields a dephosphorylated product, which can be observed by LC-MS. However, structure elucidation of this product failed in previous work. On one hand, purification of the putative butenolide **60** for subsequent NMR spectroscopic analysis from a scaled-up enzymatic reaction appears extremely difficult because of its inherent instability. On the other hand, attempts to synthesise the butenolide **60** as a reference also failed.

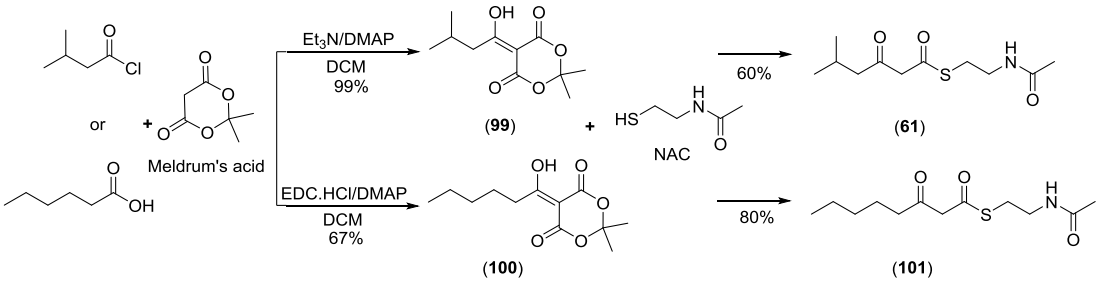


Scheme 2.3. Enzymatic reactions performed by N. Malet with synthetic 5-methyl-3-oxohexanoyl-NAC for investigation of the function of MmfL.[77]

In order to determine the specific biosynthetic role of MmfL in the assembly of AHFCAs, further studies need to be carried out, mainly focusing on structure elucidation of the presumed dephosphorylated butenolide. The key point was preparation of an authentic standard of butenolide to use it as a reference for characterisation of the product from an MmfL-catalysed reaction.

2.2 Synthesis of the N-acetylcysteamine (NAC) β-ketothioester substrate analogue

The substrates of MmfL are predicted to be DHAP, which is a commercially available compound, and various ACP-bound β-ketothioesters. Two NAC β-ketothioesters with different alkyl chains **61** and **101**, corresponding to the β-ketoacyl-ACP thioester precursors of AHFCA1 and AHFCA5, respectively, were prepared in this study as substrate analogues for MmfL (Scheme 2.4).

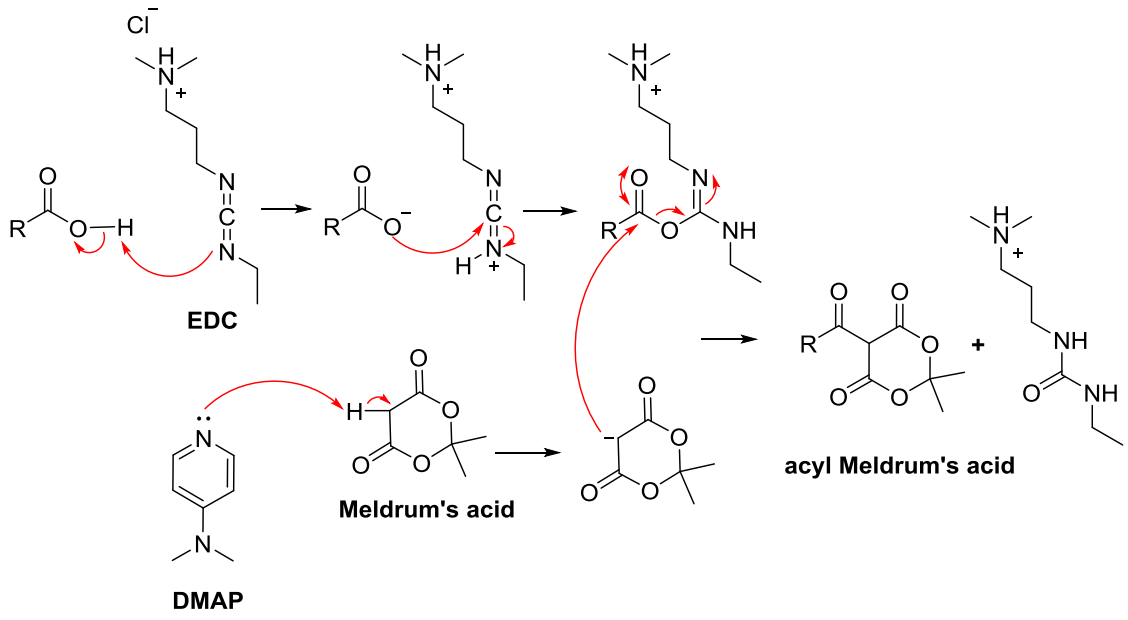


Scheme 2.4. Synthesis of NAC β-ketothioesters **61** and **101** as substrate analogues of MmfL.

Synthetic procedures for such NAC β-ketothioesters are well described in the literature.[67,77,166] They typically involve the acylation of Meldrum's acid followed by a reaction with NAC to yield the desired β-ketothioester. Due to availability of the acid or acid chloride in our lab, 3-methylbutyryl chloride and hexanoic acid were used for

synthesis of the corresponding acyl Meldrum's acids, which led to the NAC β-ketothioester precursors of AHFCA1 and AHFCA5, respectively. It should be noted that the prepared NAC β-ketothioesters exist in a mixture of the keto and enol form due to their tendency for keto-enol tautomerism, which can be easily distinguished from ^{1}H NMR spectra.

The combination of 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) and 4-dimethylaminopyridine (DMAP) was used to couple the carboxylic acid to Meldrum's acid,[167] while triethylamine and DMAP were adopted for acylation with acid chloride. EDC is capable of activating the carboxylic acid groups and offers facile removal of the urea by-product by aqueous washing due to the pendant dimethylamine moiety bestowing water-solubility. Moreover, coupling of the carboxylic acid with Meldrum's acid by EDC simplifies the preparation of acyl Meldrum's acid since the acylation takes one step, comparatively, acid chloride formation and subsequent acylation requires two steps. The general mechanism for the coupling of the carboxylic acid with Meldrum's acid induced by EDC and DMAP is described in Scheme 2.5.

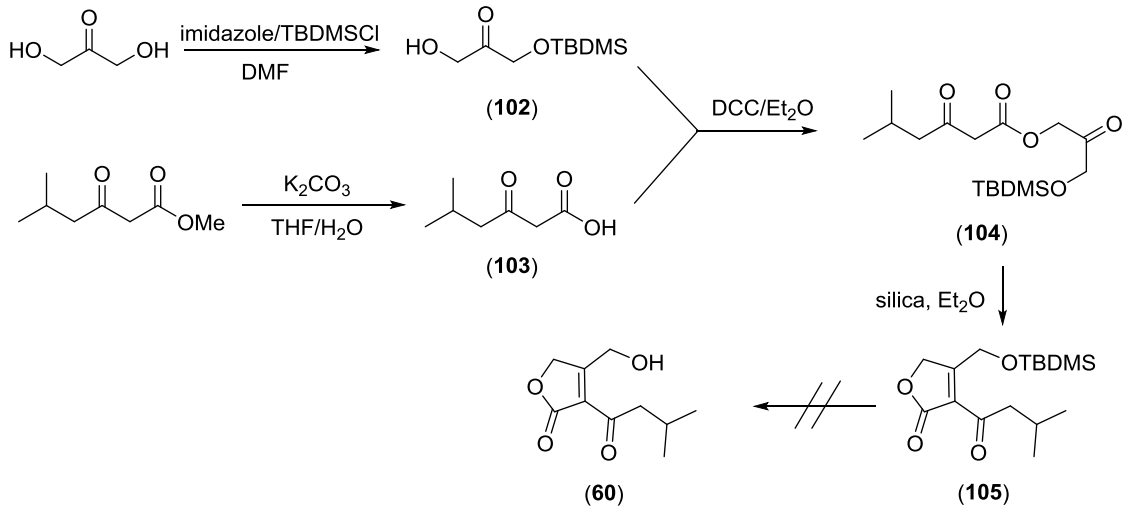


Scheme 2.5. General mechanism for the coupling of the carboxylic acid with Meldrum's acid using EDC and DMAP.

2.3 Synthesis of the butenolide authentic standard

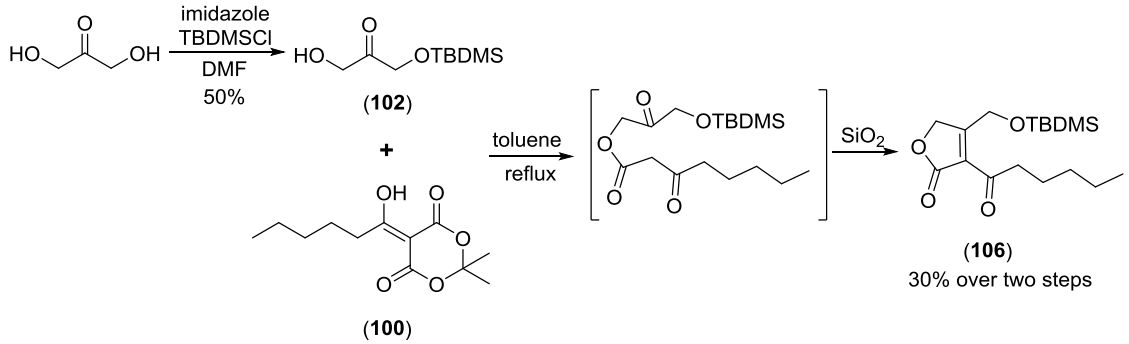
In order to confirm the proposed butenolide structure of the dephosphorylated product, the synthesis of an authentic standard of **60** was required. The *tert*-butyldimethylsilyl

(TBDMS) protected butenolide **105** was chemically synthesised by N. Malet but he was unable to complete the final deprotection step as shown in Scheme 2.6.



Scheme 2.6. Attempted synthesis of the dephosphorylated putative product **60** of the MmfL-catalysed condensation of DHAP and NAC β-ketothioester **61** by N. Malet.[77]

Additionally, although the synthetic approach was able to yield the TBDMS-protected butenolide **105**, there is still room for improvement in the synthesis. Firstly, hydrolysis of the β-ketoester to form the β-ketoacid **103** had a low yield with a long reaction time of nearly five days, and considerable amount of starting β-ketoester remained. Secondly, for commercially unavailable β-ketoesters, additional synthetic steps would be required to prepare them.



Scheme 2.7. Synthesis of the TBDMS-protected butenolide **106**.

In order to improve the synthesis of the TBDMS-protected butenolide, an alternative synthetic route was adopted in this study as shown in Scheme 2.7. This synthetic methodology was reported by Sello and co-workers for the synthesis of A-factor,[168]

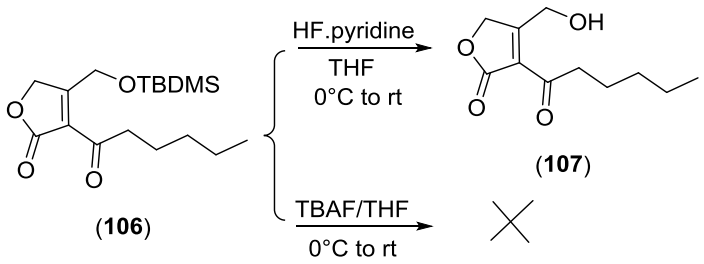
which was also used in the synthesis of γ-butyrolactone SCB1 in this work as described in Section 3.5.1.

The synthesis proceeds *via* a TBDMS-protected butenolide intermediate. The appropriate acyl Meldrum's acid is refluxed in toluene in the presence of TBDMS mono-protected dihydroxyacetone, yielding the corresponding β-ketoester, which cyclises during purification by flash chromatography on silica. This key step involves a tandem reaction sequence of an esterification and intramolecular Knoevenagel condensation.

During the synthesis, it was noticed that the TBDMS-protected butenolides are not very stable and may deteriorate during storage even at -20 ℃, particularly for those with shorter and branched alkyl chains such as the protected butenolide **105** with an isopropyl side chain corresponding to AHFCA1. This also explains the reason for the difficult purification during preparation of these compounds. Thus most of our efforts were focused on the TBDMS-protected butenolide **106** corresponding to AHFCA5, which is an unbranched pentyl group on the side chain.

As the instability of the butenolide structure is well known from reported literature and previous studies,[77,169] attempts were focused on deprotection of the silyl group using very mild conditions. Initial studies tested the effectiveness of tetrabutylammonium fluoride (TBAF) as a desilylating agent to remove the protecting group (Scheme 2.6). The reaction with TBAF was performed in THF at 0 ℃ for 30 min (Scheme 2.8). LC-MS was used to analyse reaction products because of the rapid degradation of the butenolide product. Only a tiny amount of the peak with *m/z* corresponding to the deprotected butenolide **107** was detected by LC-MS. This is probably due to generation of strongly basic ammonium alkoxides during TBAF treatment that are incompatible with base sensitive compounds.

We then explored the feasibility of using hydrogen fluoride-pyridine complex, which is a milder reaction condition (Scheme 2.8). The deprotection was performed in THF at 0 ℃ for 40 min. It resulted in a much higher peak at *m/z* = 235.0941 on LC-MS, corresponding to $[M+Na]^+$ of the deprotected butenolide **107**, as shown in trace A of Figure 2.1.



Scheme 2.8. Attempts to deprotect the TBDMS-protected butenolide **106**.

The instability of protected butenolide **106** is clearly demonstrated by the fact that a sample of it showed a low level of deprotected butenolide **107** by LC-MS (trace B of Figure 2.1). This highlights the importance of conducting LC-MS analysis with fresh samples in the studies related to the catalytic activity of MmfL.

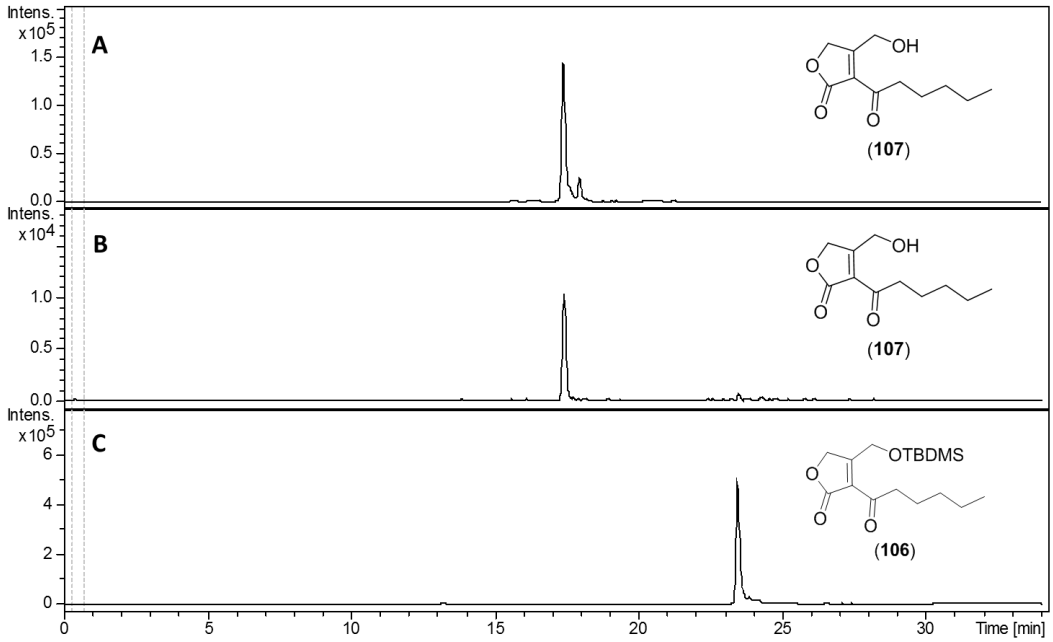


Figure 2.1. EICs at *m/z* = 235.0941 (A and B), corresponding to $[M+Na]^+$ for butenolide **107**, (A) from LC-MS analysis of the deprotection by hydrogen fluoride-pyridine, (B) from a sample of the synthetic TBDMS-protected butenolide **106**, and *m/z* = 349.1806 (C), corresponding to $[M+Na]^+$ for TBDMS-protected butenolide **106**.

Since NMR characterisation of **107** was not possible, MS/MS was used to further characterise the compound. The fragmentation pattern resulting from the $[M+H]^+$ parent ion in the ESI-MS/MS spectrum of the synthetic butenolide **107** from deprotection by hydrogen fluoride-pyridine complex correlates well with its predicted fragmentation as shown in Figure 2.2. Thus we conclude that the deprotection reaction from the TBDMS-protected butenolide **106** by hydrogen fluoride-pyridine produces butenolide **107**.

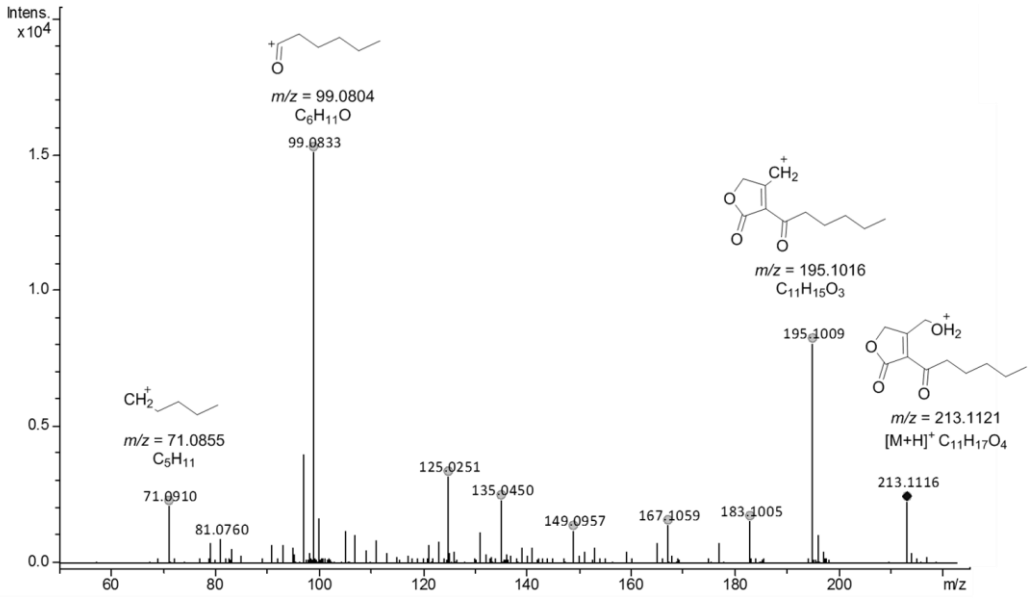


Figure 2.2. Positive ion mode ESI-MS/MS spectrum of the synthetic butenolide **107** from deprotection by hydrogen fluoride-pyridine complex with predicted fragmentation annotated.

2.4 Overproduction, purification and characterisation of recombinant MmfL

A glycerol stock of *E. coli* BL21 star (DE3) cells containing the pET151-*mmfL* plasmid was obtained from Nicolas Malet. His_6-MmfL was purified by nickel affinity chromatography as described in Section 6.3.16. SDS-PAGE analysis of the purified His_6-MmfL protein (~41 kDa) confirmed it had the expected molecular weight (Figure 2.3).

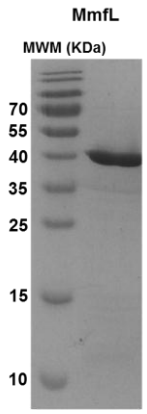


Figure 2.3. SDS-PAGE analysis of His_6-MmfL (~41 kDa) after purification. (MWM: molecular weight marker)

The molecular weight of His_6- MmfL was further confirmed by LC-HRMS (Figure 2.4).

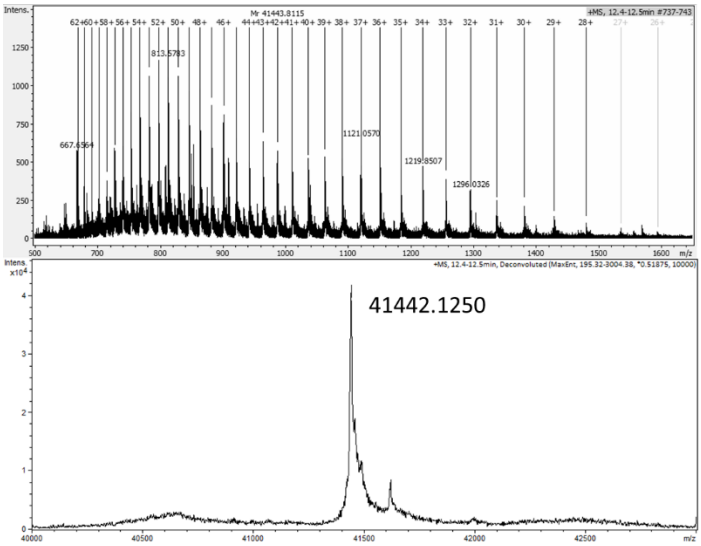
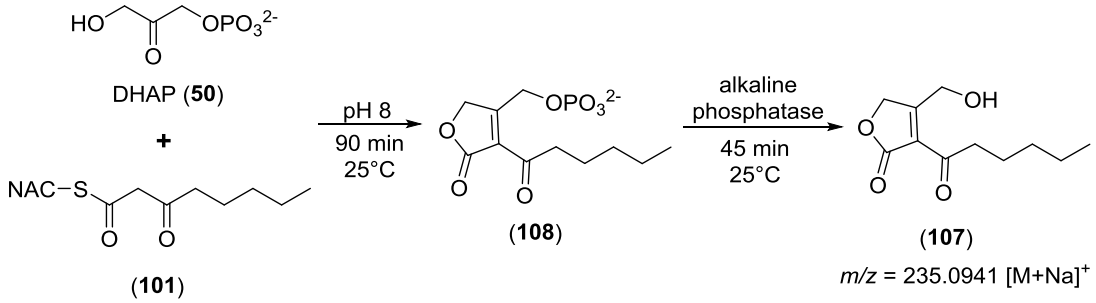


Figure 2.4. Measured (top) and deconvoluted (bottom) mass spectra of His_6-MmfL (calculated mass = 41442.0 Da).

2.5 Characterisation of dephosphorylated product of MmfL-catalysed reaction

Based on previous studies, MmfL-catalysed condensation of DHAP and the NAC β-ketothioester **101,** which is a mimic of the proposed β-ketoacyl-ACP thioester precursor of AHFCA5, was followed by treatment with alkaline phosphatase (Scheme 2.9). In this study, it was also found that a good temperature for MmfL-catalysed reaction and subsequent dephosphorylation is 25 ℃, instead of 30 ℃ used by Nicolas Malet.



Scheme 2.9. MmfL-catalysed condensation of DHAP and NAC β-ketothioester **101**, followed by dephosphorylation to form butenolide **107**.

Reactions were quenched with methanol and the supernatants were analysed by LC-HRMS. LC-HRMS analyses in positive ion mode of the condensation and subsequent dephosphorylation reactions revealed a product that gave rise to an ion with *m/z* 235.0936, corresponding to $[M+Na]^+$ of butenolide **107**. This product was absent from negative controls, confirming that it results from MmfL-catalysed condensation of DHAP and NAC β-ketothioester **101** (Figure 2.5).

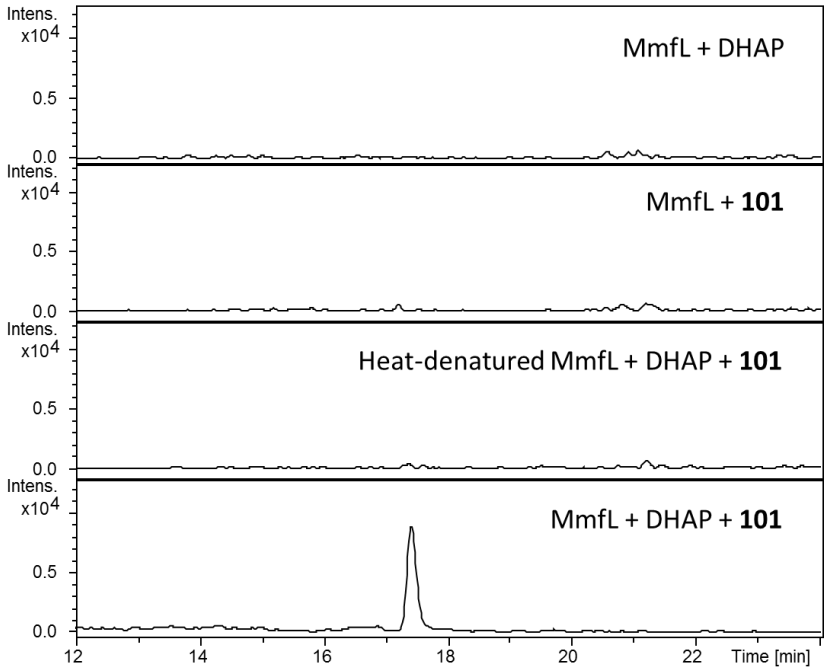


Figure 2.5. EICs for *m/z* = 235.0941 (corresponding to $[M+Na]^+$ for butenolide **107**) from LC-MS analyses of the dephosphorylated product of the MmfL-catalysed reaction of DHAP with NAC β-ketothioester **101**.

The molecular formula of the dephosphorylated product of the MmfL-catalysed condensation of DHAP and NAC β-ketothioester **101** are identical to the molecular formula of AHFCA5 **44**. In order to distinguish between the two species, LC-MS-guided comparisons of the dephosphorylated product of the MmfL-catalysed condensation with synthetic AHFCA5, which was obtained from Nicolas Malet, were carried out. The results from LC-MS shown in Figure 2.6 confirmed that they are different.

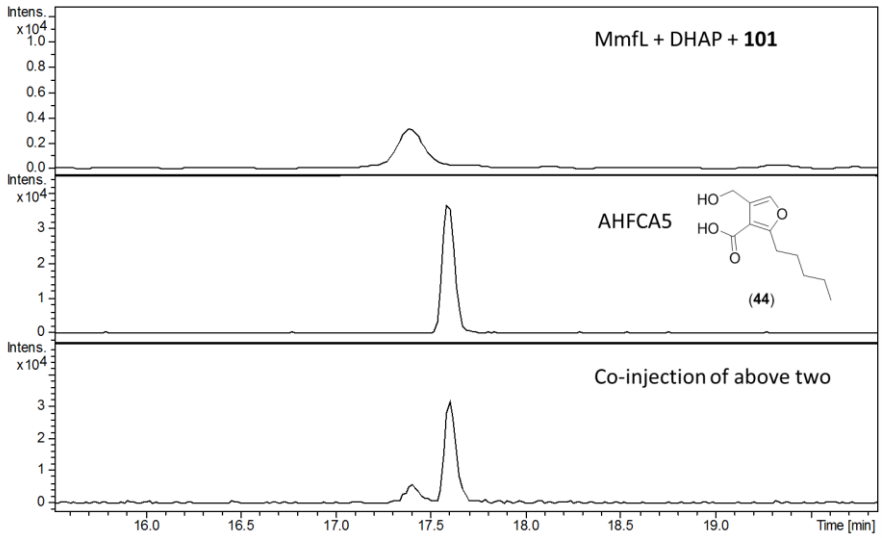


Figure 2.6. EICs for *m/z* = 235.0941 from LC-MS analyses of the dephosphorylated product of the MmfL-catalysed reaction of DHAP with NAC β-ketothioester **101** and synthetic AHFCA5 **44**, as well as a co-injection of them.

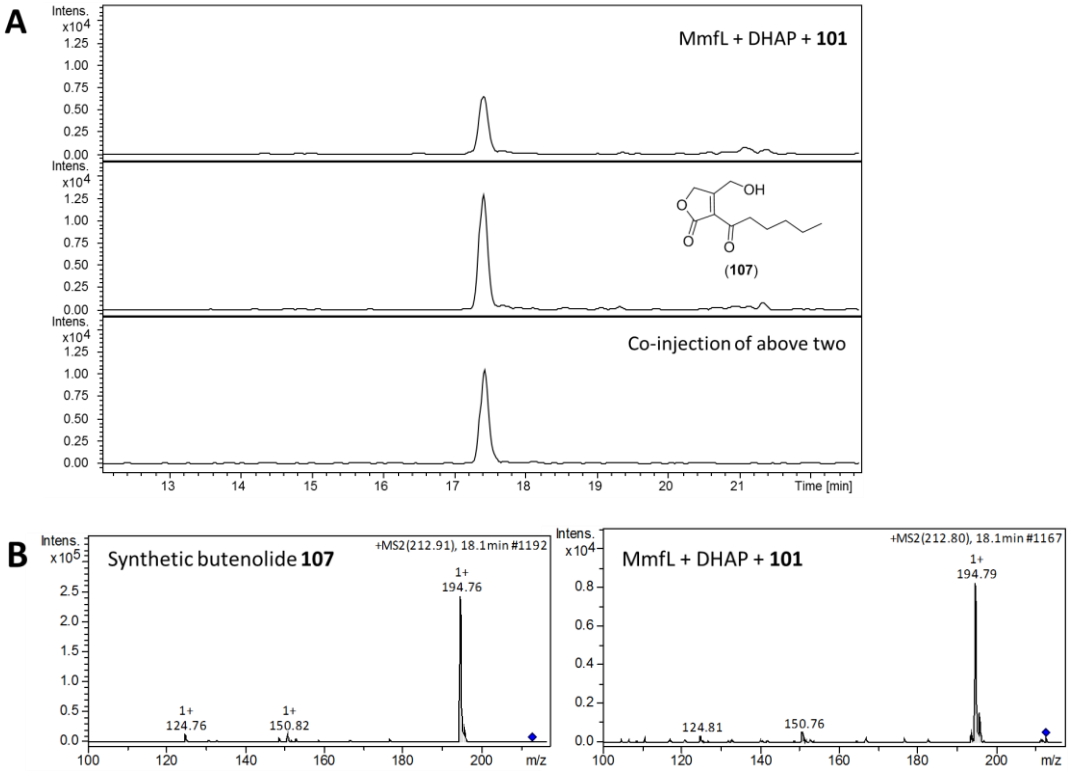
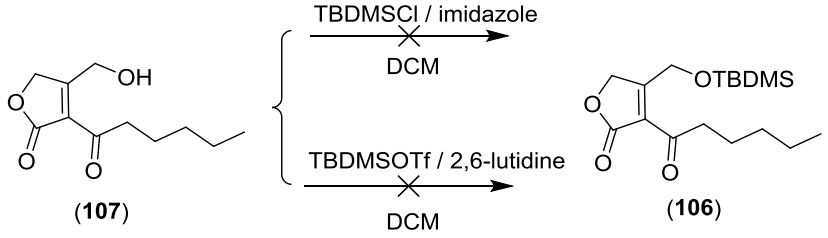


Figure 2.7. (A) EICs for *m/z* = 235.0941 from LC-MS analyses of the dephosphorylated product of the MmfL-catalysed reaction of DHAP with NAC β-ketothioester **101** and synthetic butenolide **107**, as well as a co-injection of them in equal volumes. (B) Comparison of the MS/MS fragmentation spectra of the peaks with *m/z* = 213 (corresponding to $[M+H]^+$) obtained from synthetic butenolide **107** and the enzymatic reaction product. Precursor ions are indicated with diamonds.

Further LC-MS comparisons of the dephosphorylated product of the MmfL-catalysed condensation of DHAP and NAC β-ketothioester **101** were carried out with synthetic butenolide **107** as an authentic standard. The product from enzymatic reactions and synthetic butenolide **107** had the same retention time and the same MS/MS fragmentation pattern as shown in Figure 2.7. This strongly suggests that the dephosphorylated product of the MmfL-catalysed reaction is actually the presumed butenolide.

The condensation/dephosphorylation reaction cascade was modestly scaled up in an attempt to chemically protect the hydroxyl group of the butenolide product **107** with a silyl group. The enzymatic product was extracted with dichloromethane and dried over $MgSO_4$ before trying the silyl protection reaction with TBDMSCl and imidazole. However, the TBDMS-protected butenolide **106** was not detected from LC-MS analysis of the sample. Attempts with a very strong silylating agent, *tert*-butyldimethylsilyl triflate (TBDMSOTf), and 2,6-lutidine also failed, presumably due to the instability of the butenolide **107** under basic conditions of the silylating reactions (Scheme 2.10).



(**107**) (**106**)

Scheme 2.10. Attempted synthesis of the TBDMS-protected butenolide **106** from the dephosphorylated product of the MmfL-catalysed condensation of DHAP and NAC β-ketothioester **101**.

2.6 Summary

MmfL was proposed to catalyse an analogous reaction to AfsA, involving condensation of ACP-bound β-ketothioester with DHAP followed by spontaneous intramolecular aldol condensation and dehydration to form a butenolide intermediate. Previous *in vitro* studies with recombinant His_6-MmfL protein failed to provide enough evidence for structure elucidation of the butenolide.

The synthesis of a butenolide **107** with its side chain corresponding to that of AHFCA5 was achieved in this study. It involves a tandem reaction sequence of an esterification and intramolecular Knoevenagel condensation from acyl Meldrum's acid **100** and TBDMS mono-protected dihydroxyacetone **102**. After preparation of the TBDMS-protected butenolide **106**, a mild deprotection condition using hydrogen fluoride-pyridine complex was adopted to give the desired butenolide **107**, as evidenced by its ESI-MS/MS spectrum. Moreover, LC-MS analysis provided evidence that the TBDMS-protected butenolide compounds are not very stable, leading to a mixture with its deprotected butenolide form.

The dephosphorylated product of the MmfL-catalysed reaction was confirmed to be butenolide from LC-MS/MS comparison with synthetic butenolide **107** as an authentic standard. However, TBDMS protection of the dephosphorylated product of the MmfL-catalysed reaction proved to be difficult, possibly due to the instability of the butenolide structure under basic conditions required for silylating reactions.

Based on current studies, it has been demonstrated that MmfL catalyses the reaction between ACP-bound β-ketothioester and DHAP to form a phosphorylated butenolide intermediate. This process is analogous to that catalysed by AfsA during the γ-butyrolactone signalling molecule A-factor biosynthesis in *S. griseus*.

3. Interactions between DNA, ligand and ArpA-like repressors

3.1 Introduction

A widely used method for studying DNA-protein interactions is electrophoretic mobility shift assays (EMSAs), which are also known as gel shift assays or gel retardation assays.[170-171] Because free DNA and DNA-protein complexes migrate differently during gel electrophoresis, they can be separated and detected on polyacrylamide or agarose gels.

Initial EMSAs were carried out for studying the interactions between MmfR and MmyR, DNA sequences and different AHFCA compounds (work by P. Harrison). Although it is common to use labelled DNA that can be visualised *via* interaction with a corresponding antibody, this approach led to high levels of background. It was found that simply staining with GelRed and visualisation with UV light gave a clearer gel, although more DNA was required due to lower sensitivity of this method.

Some details for running EMSAs were concluded from previous work. Firstly, the non-denaturing polyacrylamide gels are better made fresh and pre-run before loading samples. Pre-running the gels distributes ions equally within the gel and removes un-polymerised acrylamide, salts and ammonium persulfate from the gel, which can interfere with the progression of the sample through the gel; secondly, it is better to prepare the protein in gel filtration buffer which does not contain imidazole. Imidazole salts contained in the protein buffer can cause streaking of the DNA through the gel, possibly due to formation of salt-DNA complexes; finally, ligands should be dissolved in DMSO as AHFCAs solubilised in methanol produced inconsistent results.

A schematic representation of the EMSAs used in this study is shown in Figure 3.1. DNA-protein binding can be identified by running parallel samples with and without protein. If the protein binds to the DNA, there will be retardation of the DNA migration through the gel, resulting in a band shift. Upon addition of an active ligand, the receptor protein interacts with the ligand, leading to a conformational change. This results in a significant decrease of the DNA-binding affinity for the protein. Therefore, the DNA will be released from the protein and its migration through the gel will no longer be retarded.

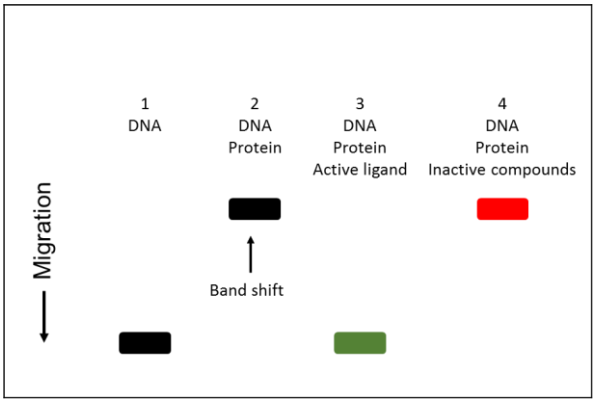


Figure 3.1. Schematic representation of EMSAs in this study.

EMSAs previously performed for MmfR/MmyR included two DNA fragments, one of which is the entire *mmfR-mmfL* intergenic region (194 bp) containing MARE1 and the other is a shorter *mmfR-mmfL* intergenic sequence (44 bp) with an additional 13 bp nucleotides flanking either side. An example of the previous EMSAs is shown in Figure 3.2. It has been confirmed that MmfR protein can bind to both the DNA fragment of the entire *mmfR-mmfL* intergenic region and the shorter 44-bp sequence. This indicates that the MARE1 sequence located between *mmfR* and *mmfL* is one of the binding sites for MmfR. However, due to solubility issues with MmyR, its binding to either of these two DNA fragments was not observed in EMSAs.

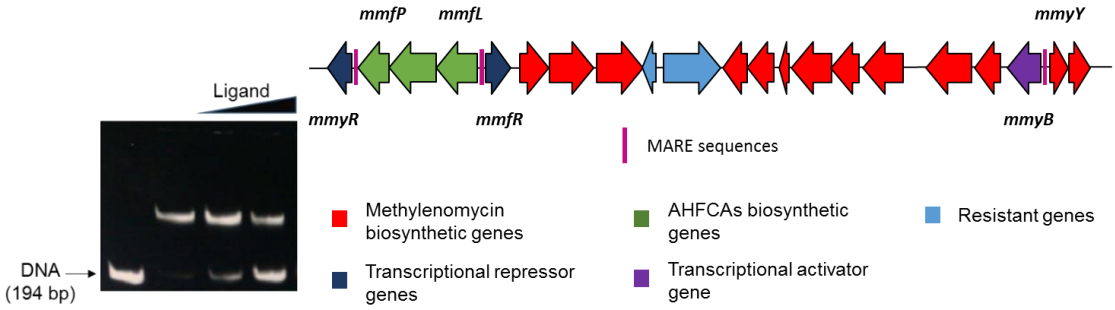


Figure 3.2. An example of the EMSAs with DNA fragment of the entire *mmfR-mmfL* intergenic region (194 bp), MmfR and AHFCA1 by P. Harrison (left), and the *S. coelicolor* methylenomycin biosynthetic gene cluster with location of three proposed MAREs highlighted in magenta (right).

Additionally, EMSAs were used to probe the ligand specificity of MmfR for different AHFCA compounds. These *in vitro* data, combined with the results from the *in vivo* experiments for induction of methylenomycin production (work by N. Malet), revealed some important structural features for AHFCAs to interact with MmfR and cause its

release from the DNA.[77] It has been showed that AHFCAs with alkyl chains between 3 and 5 carbons long are most effective at attenuating the ability of MmfR to bind the DNA fragment.

Preliminary steps have been taken to understand the function of the transcriptional regulator MmfR in regulation of methylenomycin biosynthesis in *S. coelicolor* A3(2). However, further study of the interactions between MmfR, the MARE sequences and signalling molecules is required. In order to study the DNA binding of MmfR more specifically, Section 3.2 will focus on interactions between MmfR and a smaller DNA fragment containing MARE1, as well as shorter oligonucleotides containing the other two MARE sequences. Section 3.3 will introduce a comparison of the interactions of MmfR with the five natural AHFCAs 1-5 in the presence of the DNA sequence containing MARE1. As mentioned in Section 1.3.3, the X-ray crystal structure of MmfR with AHFCA2 bound has been solved by the Corre and Challis groups. The EMSA results with various MmfR mutants based on site-directed mutagenesis will be described in Section 3.4. Section 3.5 will focus on the interaction of DNA and MmfR in the presence of the synthetic SCB1 compared with the presence of AHFCA1. In addition, the interactions between SgnR, a close homologue of MmfR in *S. venezuelae*, and its proposed ARE sequences, as well as their interaction with a cognate ligand, will be summarised in Section 3.6.

3.2 Interactions of MmfR with three MARE sequences

3.2.1. Confirmation of MmfR binding to MARE sequences

The pET151/D-TOPO plasmid containing *mmfR* was provided by Dr. Christophe Corre. Chemically-competent *E. coli* BL21 star (DE3) cells were transformed with the pET151-*mmfR* plasmid and the recombinant His_6-MmfR protein was overproduced and purified by nickel affinity chromatography as described in Section 6.3.16. SDS-PAGE analysis of the purified His_6-MmfR protein (~27.8 kDa) confirmed it had the expected molecular weight (Figure 3.3).

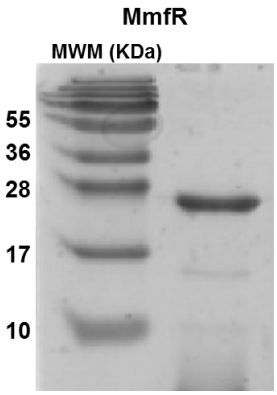


Figure 3.3. SDS-PAGE analysis of His_6-MmfR (~27.8 kDa) after purification. (MWM: molecular weight marker)

The molecular weight of His_6-MmfR was further confirmed by LC-HRMS (Figure 3.4).

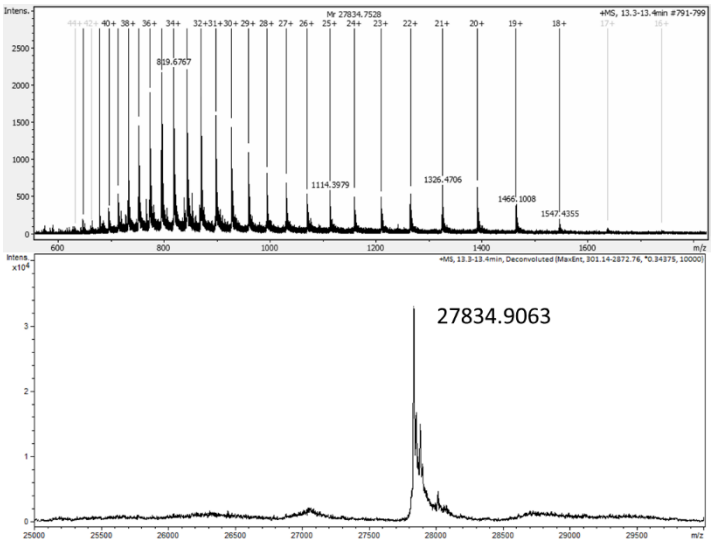


Figure 3.4. Measured (top) and deconvoluted (bottom) mass spectra of His_6-MmfR (calculated mass = 27835.5 Da).

As discussed, the DNA fragments used previously are both considerably longer than the 18 bp MARE sequences shown in Table 1.1. Thus the EMSAs were repeated with shorter DNA sequences (both single- or double-stranded). The nearly palindromic character of the MARE sequences intrigued us. To check whether the palindrome is required for MmfR recognition and binding, we first used a single-stranded DNA fragment, the 24 bp perfectly palindromic sequence, 5'-AACATACCTTCGCGAAGGTATGTT-3', in the EMSA with MmfR. Attempts with 6% non-denaturing polyacrylamide gels and the same amount of DNA (0.1 pmol) previously used all failed. This is likely due to improper pore size of the 6% gel, which is possibly too big for such small single-stranded oligonucleotides. Also, the short DNA may be lower than the detection limit of the DNA

staining and visualisation method used in our EMSAs. The free DNA band could only be observed after the acrylamide concentration of the gel was increased to 20% with the use of 10 times more DNA fragment (1 pmol). However, no band shift was observed with this single-stranded oligonucleotide or the annealed double-stranded palindrome when MmfR was added, even after addition of 100 times more than DNA.

In these attempts, two slightly longer palindromic sequences, 5'-CAACATACCTTCGCGAAGGTATGTTG-3' (26 bp) and 5'-GCAACATACCTTCGCGAAGGTATGTTGC-3' (28 bp), also failed to show interaction with MmfR in both the single-stranded and annealed form. It seems that this perfectly palindromic sequence cannot be bound by MmfR. However, since the double-stranded palindromic sequence was annealed from the single-stranded oligonucleotides without purification, the conclusion needs to be taken with caution.

To improve the EMSAs with small DNA fragments, we then elected to use self-annealing DNA fragments that can form hairpin sequences (Figure 3.5). The 24 bp oligonucleotide in the middle of the hairpin structure is considerably flexible,[172] allowing us to locate the three 18-bp MARE sequences with an additional 3 nucleotides flanking either side. The variable region is surrounded by GC clamps on both the 5' and 3' ends. The sequences were ordered as single-stranded oligonucleotides that annealed to form the hairpin structure (see Section 6.3.18).

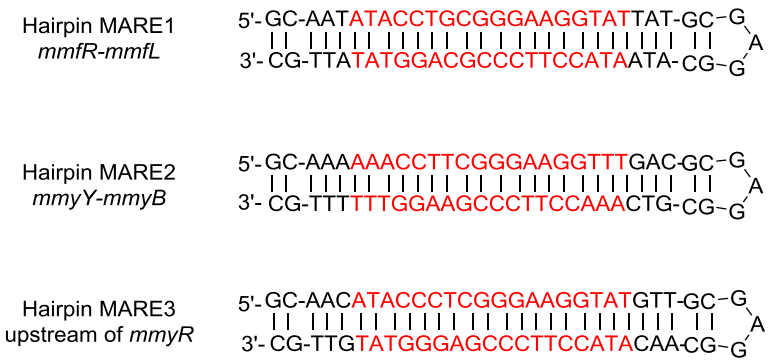


Figure 3.5. Hairpin DNA containing three MAREs (18 bp, coloured in red) for EMSAs.

In view of the still small size of these DNA fragments, attempts were first carried out to optimise the amount of DNA and MmfR used in each reaction and the acrylamide concentration of the gel in order to obtain EMSA images with clear band shifts. A higher amount of DNA fragments, 0.8 pmol, was determined to be appropriate in conjunction with accordingly increased amount of protein, 4.0 pmol. Meanwhile, 10% non-

denaturing polyacrylamide gels showed the best images, although the time of gel electrophoresis had to be increased to more than two hours.

These modifications meant the EMSAs using the hairpin DNA fragment containing the MARE1 sequence worked much better than previously, as shown in Figure 3.6. Moreover, DNA release from MmfR after addition of its natural ligand AHFCA1 was successfully observed. The release profile of DNA from protein was consistent with that observed with the 194-bp DNA of the entire *mmfR-mmfL* intergenic region.

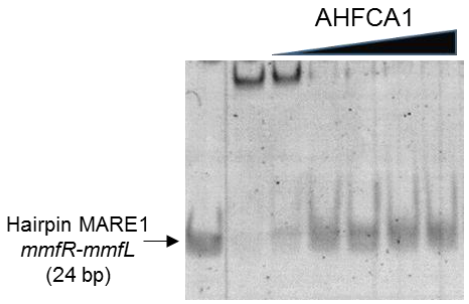


Figure 3.6. Interaction of MmfR with a self-annealing DNA fragment forming a hairpin sequence containing MARE1 located in the *mmfL-mmfR* intergenic region in response to increasing amounts of AHFCA1. Lane 1: DNA fragment only. Lanes 2 to 7: AHFCA1 at 0, 0.8, 8, 20, 40 and 100 nmol respectively. Amounts of DNA and proteins were kept constant (0.8 pmol and 4.0 pmol respectively).

The same conditions were then used for EMSAs with hairpin DNA containing the other two MARE sequences, as shown in Figure 3.7. Though the progression of samples was distorted, leading to skewed images, there are obvious band shifts due to formation of the DNA-MmfR complex. This strongly indicated the ability of MmfR to bind the other two specific MARE sequences. In the presence of the ligand AHFCA1, increasing release of DNA fragments can also be observed. Thus, the three proposed MARE sequences have been experimentally confirmed to be the binding sites for MmfR and could be released upon the interaction of MmfR with AHFCA ligands.

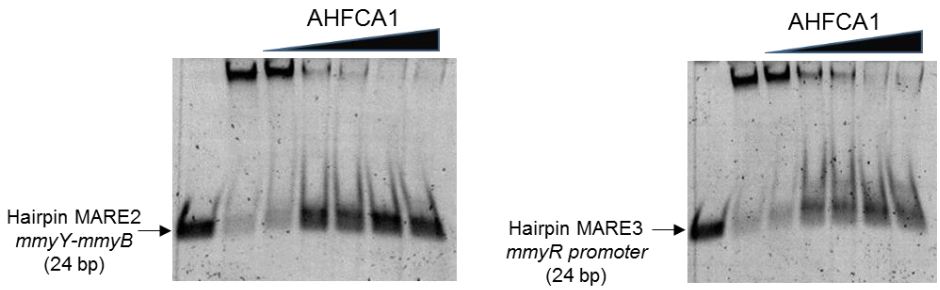


Figure 3.7. Interaction of MmfR with self-annealing DNA fragments forming hairpin sequences containing MARE2 located in the *mmyY-mmyB* intergenic region (left) or MARE3 located upstream of *mmyR* (right) in response to increasing amounts of AHFCA1. Lane 1: DNA fragment only. Lanes 2 to 7: AHFCA1 at 0, 0.8, 8, 20, 40 and 100 nmol, respectively. Amounts of DNA and proteins were kept constant (0.8 pmol and 4.0 pmol respectively).

3.2.2. Comparison of the binding affinity of MmfR to three MAREs

In order to further understand the role of MmfR in the regulation of methylenomycin biosynthesis, determination of the relative binding affinities of MmfR to the three operator regions was investigated. We thereby undertook competition assays, in which two different MARE sequences were present in the same reaction, with increasing concentrations of MmfR. In order to differentiate between the different sequences on a gel, both long and short versions were prepared as shown in Table 3.1.

Table 3.1. Longer and shorter DNA sequences containing each of MAREs.

ARE No.	Location of DNA	Size of large (bp)	Size of short (bp)
MARE1	*mmfR-mmfL*	194	100
MARE2	*mmyY-mmyB*	230	98
MARE3	upstream of *mmyR*	150	100

DNA fragments were amplified for each region by PCR using cosmid C73-787 as template and primers listed in Table 6.1. For MARE1, located in the *mmfR-mmfL* intergenic region, the entire intergenic region (194 bp) and a partial fragment (100 bp) were amplified. For MARE2, located in the *mmyY-mmyB* intergenic region, the entire intergenic region (230 bp) and a partial fragment (98 bp) were amplified. For MARE3 located upstream of *mmyR*, the entire *mmfP-mmyR* intergenic region (150 bp) and one shorter fragment (100 bp) were prepared.

All the fragments still contain their specific MARE sequences. Agarose gel electrophoresis was used to separate the desired PCR products from unused reactants and template DNA (Figure 3.8). The PCR products were excised from the gel and subsequently purified and sequenced.

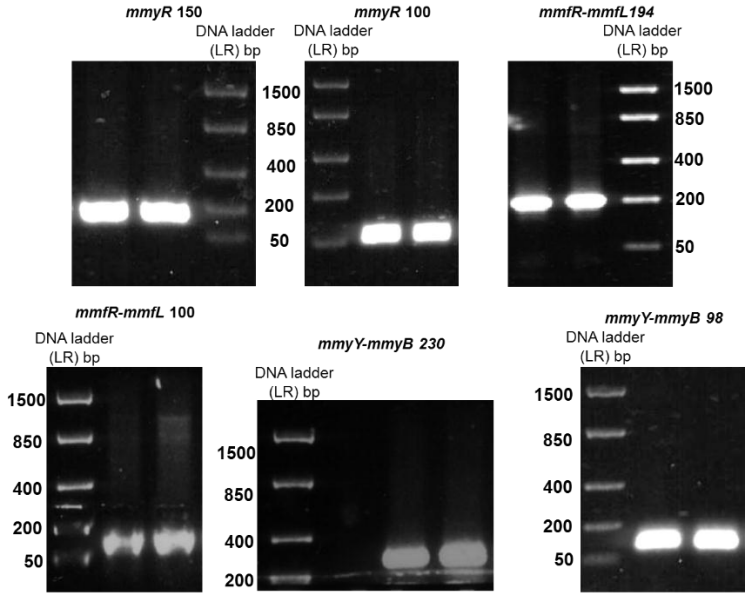


Figure 3.8. Agarose gel electrophoresis analyses of the PCR products for different length of DNA fragments containing MAREs. (LR: low range DNA ladder)

The two operators, one located between *mmfR* and *mmfL* and the other between *mmyY* and *mmyB*, were compared in EMSAs on 6% non-denaturing polyacrylamide gels as shown in Figure 3.9.

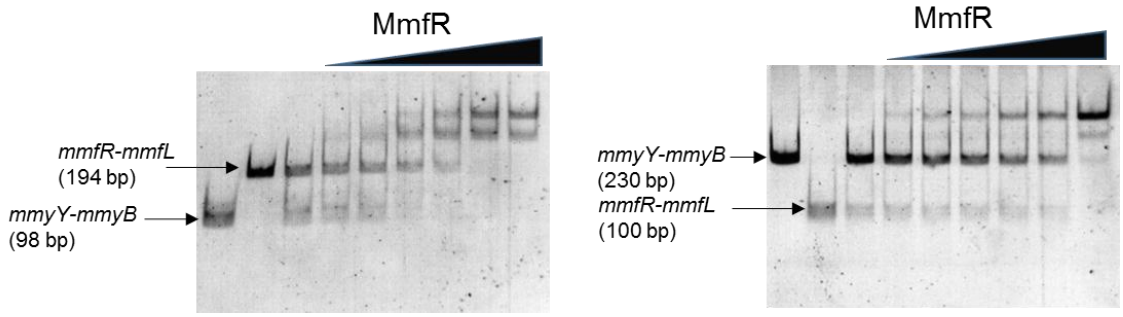


Figure 3.9. Competitive binding of MmfR with two DNA fragments corresponding to the *mmfR-mmfL* and *mmyY-mmyB* intergenic regions. Lanes 1 to 3: DNA fragments only. Lanes 4 to 9: MmfR at 0.1, 0.2, 0.3, 0.4, 0.5 and 0.6 pmol, respectively. The amount of DNA was kept constant (0.05 pmol).

As shown in the left gel in Figure 4.9, the *mmfR-mmfL* fragment was larger than *mmyY-mmyB* while in the right gel, the *mmfR-mmfL* fragment was shorter than *mmyY-mmyB*.

Figure 3.9 (left gel) shows the band shift corresponding to the DNA fragment between *mmyY-mmyB* occurs at lower protein concentration than the band shift for the fragment between *mmfR* and *mmfL*. To check this is not just a product of the different DNA sizes used, the experiment was repeated with a larger *mmyY-mmyB* DNA fragment and a smaller *mmfR-mmfL* DNA fragment (Figure 3.9, right gel). Since the same result was observed, this strongly indicates that MmfR binds the operator between *mmyY* and *mmyB* more strongly than that between *mmfR* and *mmfL*.

Using the same technique, binding of MmfR to the operator regions, located between *mmfR* and *mmfL* and upstream of *mmyR*, was compared (Figure 3.10). The right gel with a longer *mmfR-mmfL* intergenic DNA and shorter sequence upstream of *mmyR*, did not give useful information, without obvious band shifts until the maximum MmfR concentration. In the left gel, the shorter DNA fragment corresponding to the operator between *mmfR* and *mmfL*, exhibited band shifts with lower concentrations of MmfR, suggesting that the affinity of MmfR to this operator is higher than to the one upstream of *mmyR*. However, it was noted that there may be some crosslinking of the DNA fragments when the highest amount of MmfR (0.6 pmol) used, as shown in the rightmost lane of the gel.

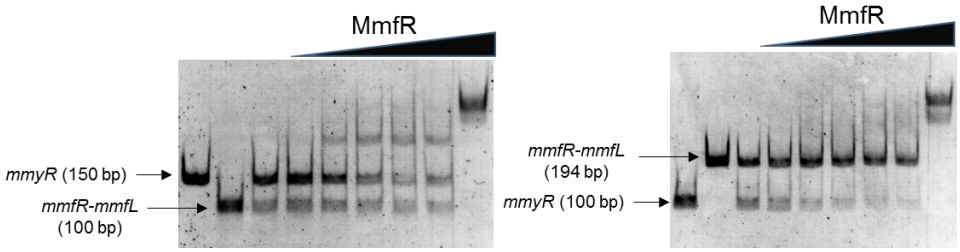


Figure 3.10. Competitive binding of MmfR with two DNA fragments corresponding to the *mmfR-mmfL* intergenic region and upstream of *mmyR*. Lanes 1 to 3: DNA fragments only. Lanes 4 to 9: MmfR at 0.1, 0.2, 0.3, 0.4, 0.5 and 0.6 pmol, respectively. The amount of DNA was kept constant (0.05 pmol).

The above EMSAs provided an affinity order of MmfR to the three MAREs as followed: the operator between *mmyY* and *mmyB* > the operator between *mmfR* and *mmfL* > the operator upstream of *mmyR.*

In order to further confirm this result, the affinity comparison of MmfR to the operator located between *mmyY* and *mmyB* and the operator upstream of *mmyR* was also performed. However, the EMSAs did not work well and the band shifts seemed unclear.

From the gels shown in Figure 3.11, it can be presumably concluded from the band shifts that the operator between *mmyY* and *mmyB* has higher affinity to MmfR compared to the operator upstream of *mmyR*. This is consistent with the affinity order obtained from the above two competitive binding assays. These results would provide more evidence for understanding the role of MmfR in the regulation of methylenomycin biosynthesis in *S. coelicolor* A3(2).

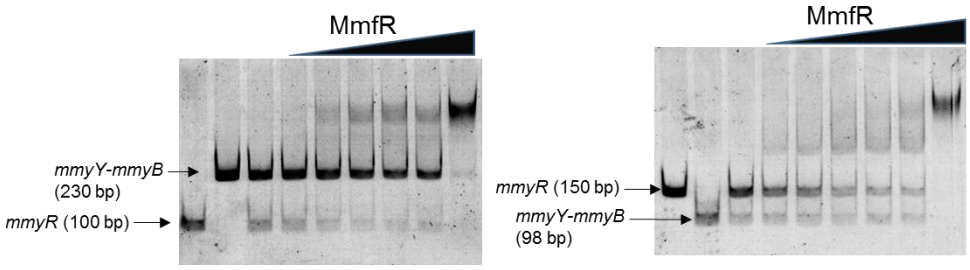


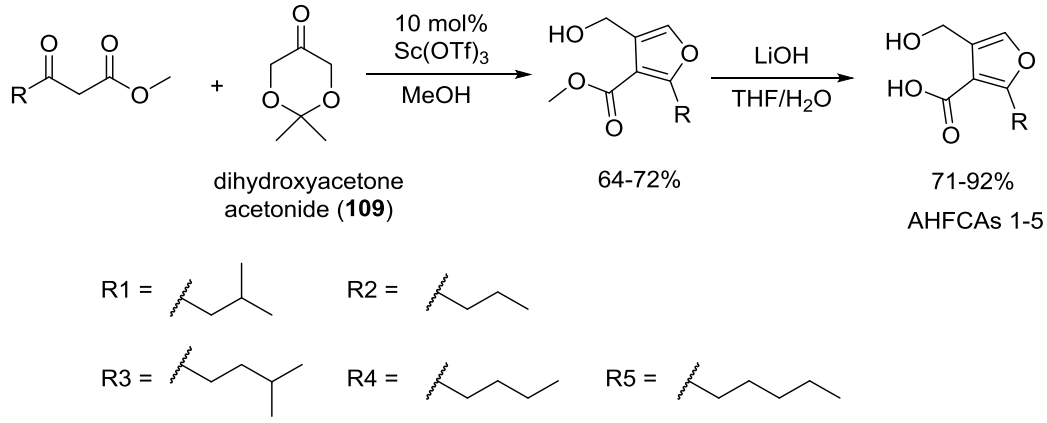
Figure 3.11. Competitive binding of MmfR with two DNA fragments corresponding to the *mmyY-mmyB* intergenic region and upstream of *mmyR*. Lanes 1 to 3: DNA fragments only. Lanes 4 to 9: MmfR at 0.1, 0.2, 0.3, 0.4, 0.5 and 0.6 pmol, respectively. The amount of DNA was kept constant (0.05 pmol).

3.3 Interactions of MmfR with natural AHFCAs

As discussed in Section 1.2.1, there are five natural AHFCAs that act as autoinducers for methylenomycin production in *S. coelicolor*. In order to assess their relative ability to bind MmfR and release the protein from cognate operators, EMSAs were carried out on 6% non-denaturing polyacrylamide gels for each AHFCA compound using DNA fragments of the entire *mmfR-mmfL* or the entire *mmyY-mmyB* intergenic region. These fragments were chosen because they have the highest affinity to MmfR and their long size provides better EMSA images after DNA staining and visualisation under UV.

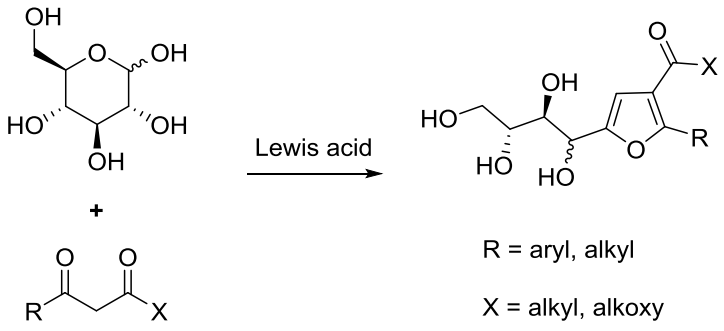
3.3.1 Synthesis of AHFCA1 and AHFCA3

The ligands, AHFCA2, AHFCA4 and AHFCA5, used in the EMSAs were obtained from Nicolas Malet. AHFCA1 and AHFCA3 were chemically synthesised in this study. The synthesis of AHFCAs has been well established in the Challis group based on a reported synthetic route by Sello and co-workers.[76-77]



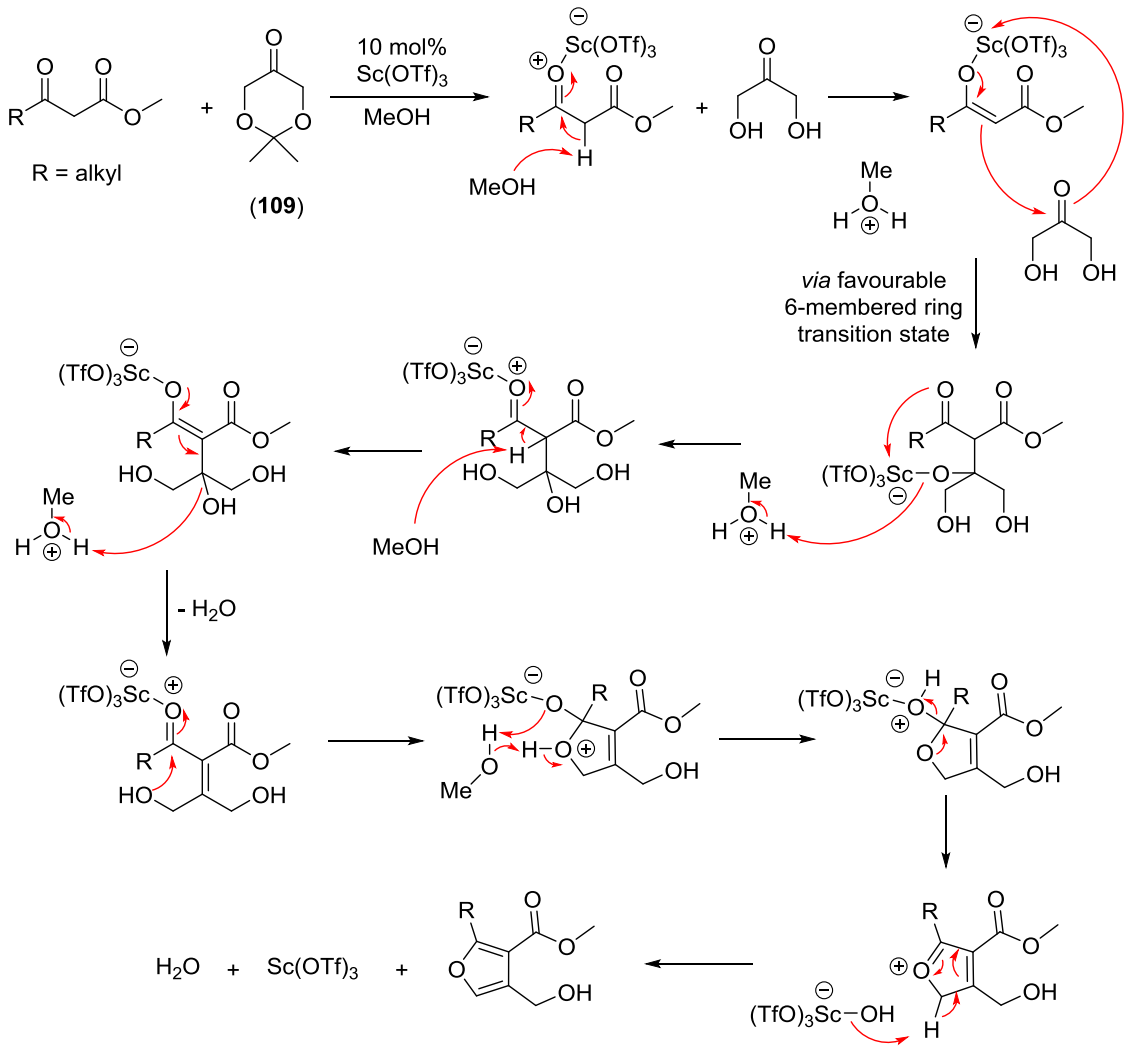
Scheme 3.1. Synthesis of AHFCAs reported by Sello and co-workers.[76]

The reported synthesis of AHFCAs utilised scandium triflate as a Lewis acid catalyst for the condensation of dihydroxyacetone acetonide **109** with a β-ketoester, affording methyl 2-alkyl-4-hydroxymethylfuran-3-carboxylates. Subsequent lithium hydroxide-mediated hydrolysis of the methyl ester gave the desired AHFCAs (Scheme 3.1). This synthesis was identified accidentally during their attempts to synthesise γ-butyrolactone signalling molecules *via* butenolide intermediates.[76] The scandium triflate-catalysed reaction has been noted as a variant of the Garcia Gonzalez reaction (Scheme 3.2).



Scheme 3.2. General concept of the Garcia Gonzalez reaction.

A possible mechanism for the condensation of **109** and the β-ketoester has been proposed involving scandium triflate-mediated *in situ* deprotection of **109** and enolisation of the β-ketoester. Knoevenagel condensation between dihydroxyacetone and the enol, followed by addition of the hydroxyl group to the ketone yielding a ketal intermediate that undergoes dehydration and aromatisation to give the furan products (Scheme 3.3).[77]

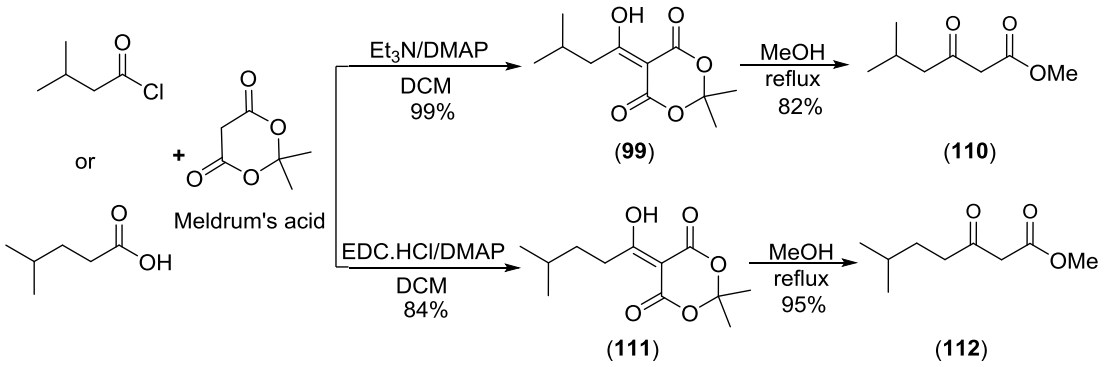


Scheme 3.3. Proposed mechanism for Knoevenagel condensation and subsequent furan formation catalysed by $Sc(OTf)_3$.[77]

The synthetic approach is characterised by its low cost of catalyst, utilising only 10 mol% of scandium triflate, and mild reaction conditions at room temperature in good yield. The reactions are neither air- nor moisture-sensitive which makes it very easy to handle. This route has previously been confirmed as a very reproducible way for synthesis of AHFCAs in the Challis group. Thus it was used in this work for synthesis of AHFCA1 and AHFCA3, as well as AHFCA6 and AHFCA7 in Section 3.6.3.

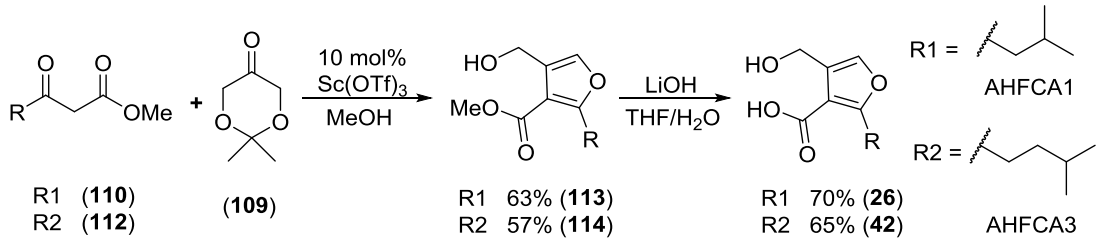
The synthesis of the β-ketoester starting materials was accomplished *via* acyl Meldrum's acid intermediates, which were prepared from the corresponding acid or acid chloride.[77,167] Both 3-methylbutyryl chloride and 4-methylpentanoic acid are available in our lab. The mechanism for the coupling of the carboxylic acid and Meldrum's acid,

induced by EDC and DMAP is described in Section 2.2. Hydrolysis of the acyl Meldrum's acid in methanol gives the corresponding β-ketoester (Scheme 3.4).



Scheme 3.4. Synthesis of the β-ketoesters **110** and **112** for AHFAC1 and AHFCA3, respectively.

Scandium triflate-catalysed condensation of dihydroxyacetone acetonide **109** with the β-keto esters **110** and **112** afforded the corresponding furans **113** and **114** in good yield after purification by flash chromatography. The ester group within these furans was hydrolysed with lithium hydroxide, yielding the AHFCA1 **26** and AHFCA3 **42** as waxy solids (Scheme 3.5). The overall yields for this two-step process are 44% for AHFCA1 and 37% for AHFCA3.



Scheme 3.5. Synthesis of AHFCA1 and AHFCA3.

3.3.2 Comparison of the binding affinity of the different intergenic regions to MmfR with natural AHFCAs 1-5

The EMSAs showing the effect of increasing concentration of different AHFCAs on binding of MmfR to the 194-bp DNA fragment of the entire intergenic region between *mmfR* and *mmfL* is shown in Figure 3.12. The more DNA released from MmfR, the more active the AHFCA is. Due to the semi-quantitative nature of EMSAs and the complicated interactions in such a three-component system involving DNA, protein and ligand,

quantification of the DNA band released by addition of different AHFCAs is difficult. However, some general conclusions can be made.

Undoubtedly, AHFCA1 was shown to be the strongest ligand among the five to interact with MmfR and cause release of the DNA, while AHFCA2 obviously showed the lowest amount of DNA release. AHFCA5 exhibited similar activity to AHFCA4 and the potency of AHFCA3 seems higher than both of them but lower than AHFCA1. Thus, based on the EMSAs using the 194-bp DNA fragment of the entire *mmfR-mmfL* intergenic region, the order of the activity of the five natural AHFCAs was concluded to be as follows: AHFCA1 > AHFCA3 > AHFCA5 ≈ AHFCA4 > AHFCA2.

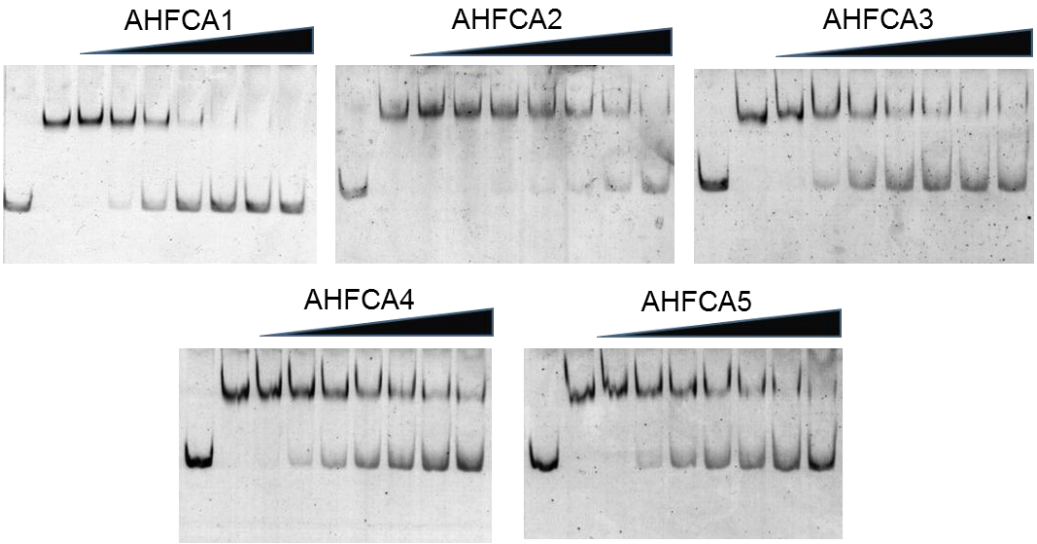


Figure 3.12. Interaction of MmfR with the 194-bp DNA fragment of the entire *mmfR-mmfL* intergenic region in response to increasing amounts of natural AHFCAs 1-5. Lane 1: DNA fragment only. Lanes 2 to 9: AHFCAs at 0, 0.8, 4, 8, 14, 20, 40 and 100 nmol, respectively. Amounts of DNA and proteins were kept constant (0.1 pmol and 1.8 pmol respectively).

The results with the 230-bp DNA fragment of the entire intergenic region between *mmyY* and *mmyB* are shown in Figure 3.13. Compared to Figure 3.12, it can be clearly seen that the ability of AHFCAs 1-5 to release the *mmyY-mmyB* intergenic DNA from the DNA-MmfR complex is lower than the *mmfR-mmfL* intergenic DNA sequence. This indicates that the binding affinity of MmfR to MARE2 located between *mmyY* and *mmyB* is stronger than to MARE1 between *mmfR* and *mmfL*, further supporting the affinity order confirmed in Section 3.2. Generally, AHFCA1 and AHFCA3 still showed higher efficacy than the other three AHFCAs.

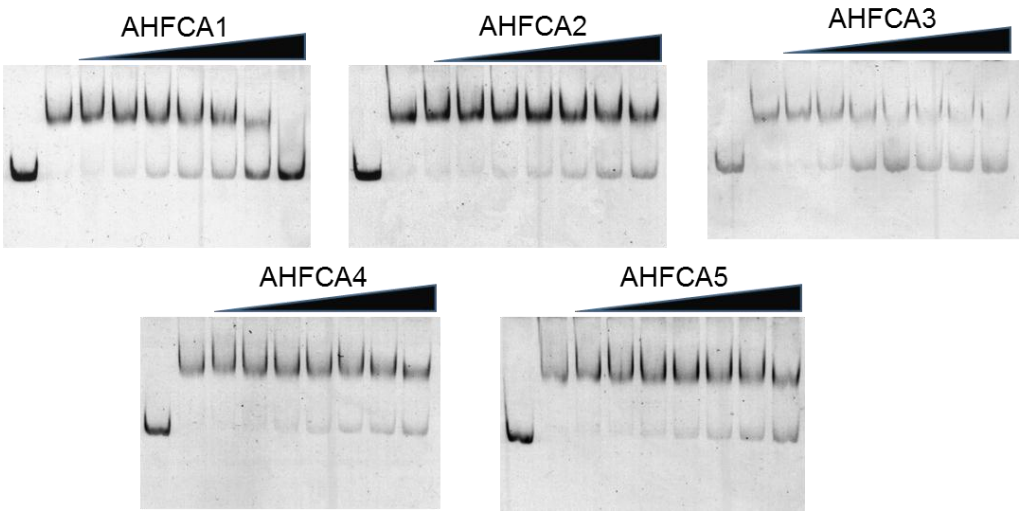


Figure 3.13. Interaction of MmfR with the 230-bp DNA fragment of the entire *mmyY-mmyB* intergenic region in response to increasing amounts of natural AHFCAs 1-5. Lane 1: DNA fragment only. Lanes 2 to 9: AHFCAs at 0, 0.8, 4, 8, 14, 20, 40 and 100 nmol, respectively. Amounts of DNA and proteins were kept constant (0.1 pmol and 1.8 pmol respectively).

3.4 Key residues in the ligand binding of AHFCAs by MmfR

3.4.1 Analysis of the crystal structure

As discussed in Section 1.3, the TetR family transcriptional regulators are highly conserved in their *N*-terminal DNA-binding domain while the rest of the sequence is less conserved.[78] This, to some extent, explains why these repressor proteins bind to ARE consensus sequences. It is possible that minor alterations in the sequence of the HTH motif formed by the *N*-terminal residues dictate their specificity for specific DNA binding sites. Moreover, it is likely that the less conserved part in these receptor proteins allows for variability in their ligand binding specificity. Thus, they are responsive to certain structural types of ligands.

The MmfR structure with AHFCA2 bound is the first example of an ArpA-like repressor structure with its ligand bound. This allows us to identify the ligand binding site in MmfR as shown in Figure 3.14. From a detailed analysis of the structure, it appears that the hydroxyl group on the phenyl ring of Tyr 85 participates directly in binding AHFCAs through hydrogen bond formation with their hydrophilic part. Additional residues may be involved in stabilising the interaction, such as Gln 130. The amide group at the end of

its side chain is in close proximity to the hydrophilic groups of AHFCA2. Although the side chain of Tyr 144 is pointing away from the ligand, its backbone amide seems to be involved in the interaction with the carboxylic acid group of the ligand. Additionally, there is a pocket for the hydrocarbon chain of AHFCA2, in which a residue, Leu 110, is situated at the bottom. The presence of Leu 110 at the bottom of the hydrophobic pocket may determine the ligand specificity of MmfR to AHFCAs with alkyl chains of modest length. It is tempting to speculate that a smaller residue, such as glycine, may contribute to the formation of a deeper hydrophobic pocket for the alkyl chain of AHFCAs. As shown in Figure 1.5, longer side chains are often seen in GBLs, such as SCBs.

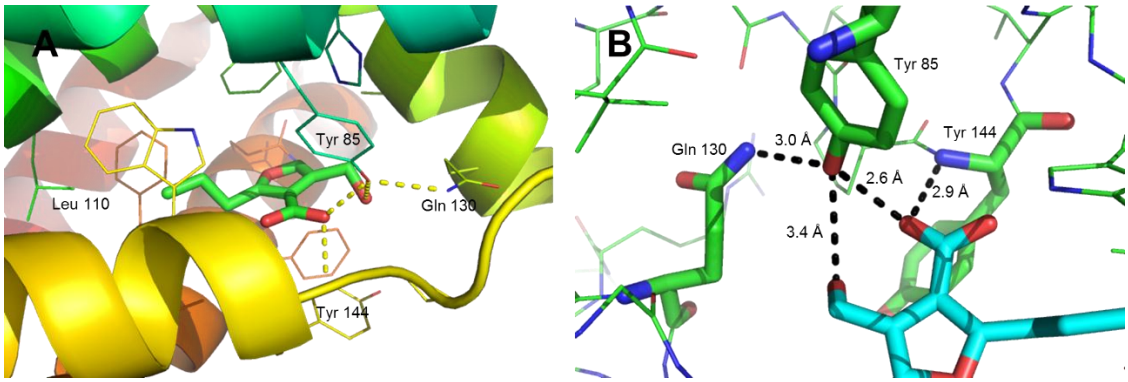


Figure 3.14. (A) Proposed key residues of MmfR interacting with AHFCA2 from solved crystal structure. (B) Key hydrogen bonding interactions in the hydrophilic part of AHFCA2.

In order to probe the above hypothesis, site-directed mutagenesis studies, followed by EMSAs, were carried out. Mutations were focused on the proposed key residues in binding AHFCAs, including Tyr 85, Gln 130, Tyr 144 and Leu 110.

3.4.2 Preparation of MmfR mutants

Since the hydroxyl group of Tyr 85 was proposed to form a hydrogen bond with the carboxylic acid group and the hydroxyl group of AHFCA ligand, Tyr 85 was mutated to a phenylalanine, as well as an alanine. As the amide group on the side chain of Gln 130 was proposed to participate in the hydrogen bond formation with the hydrophilic part of AHFCAs, Gln 130 was mutated to a glutamic acid, which does not contain an amide group at the end of its side chain. From the crystal structure at the position of Tyr 144, it was proposed that the backbone amide may be involved in the interaction with the carboxylic acid group of AHFCAs. In order to confirm this, Tyr 144 was mutated to a

phenylalanine, as well as an alanine. The plasmids of MmfR mutants with alanine were provided by Kathryn Styles, a PhD student in the Corre group.

The pET151/D-TOPO plasmid containing *mmfR* was used as the template for site-directed mutagenesis using Q5 site-directed mutagenesis kit with primers and conditions listed in Table 6.2. The constructs were sequence verified. *E. coli* BL21 star (DE3) competent cells were then transformed with these recombinant plasmids. It should be noted that *E. coli* BL21 star (DE3)pLysS competent cells were transformed with the recombinant plasmid of mutant Y144F. This is due to a slow growth of *E. coli* cells, around 6 hours to reach OD_{600} > 0.6, and low expression level of the protein when the mutant Y144F was expressed in *E. coli* BL21 star (DE3) cells. *E. coli* BL21 star (DE3)pLysS cells are more suitable for expression of potentially growth-inhibiting recombinant proteins due to its expression of T7 lysozyme, a T7 RNA polymerase inhibitor that prevents leaky expression in un-induced cells.

MmfR mutants were overproduced and purified as *N*-terminal His_6-tagged proteins. SDS-PAGE analysis confirmed they had the expected molecular weight (~27.8 kDa) (Figure 3.15).

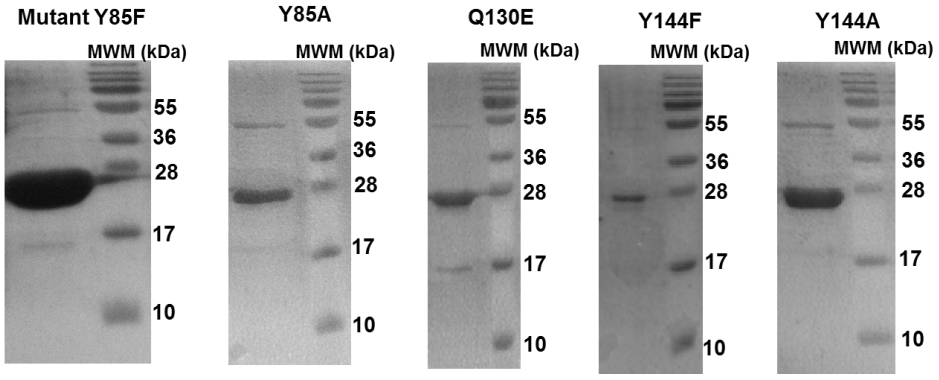


Figure 3.15. SDS-PAGE analysis of His_6-MmfR mutants, Y85F, Y85A, Q130E, Y144F and Y144A, after purification. (MWM: molecular weight marker)

The identity of these MmfR mutants, Y85F (27819.5 Da), Y85A (27743.5 Da), Q130E (27836.5 Da), Y144F (27819.5 Da) and Y144A (27743.5 Da), was further confirmed by LC-HRMS (Figure 3.16).

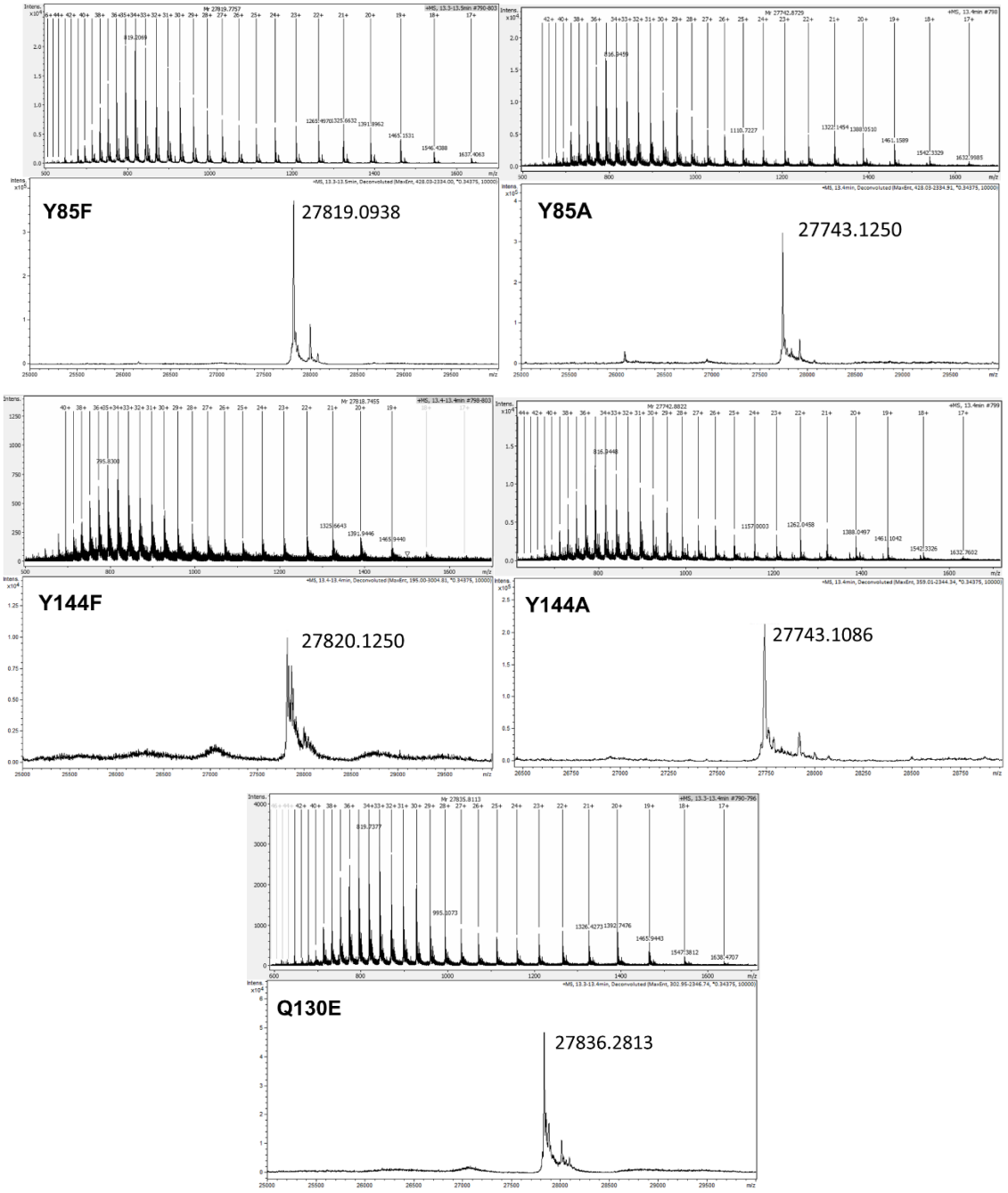


Figure 3.16. Measured (top) and deconvoluted (bottom) mass spectra of His_6-MmfR mutants, Y85F (calculated mass = 27819.5 Da), Y85A (calculated mass = 27743.5 Da), Y144F (calculated mass = 27819.5 Da), Y144A (calculated mass = 27743.5 Da) and Q130E (calculated mass = 27836.5 Da).

3.4.3 Attempts to mutate Leu 110

Considerable efforts were made to mutate Leu 110 to spatially smaller residues, such as glycine, alanine and valine, thus leading to a deeper binding pocket probably suitable for AHFCA ligands with longer alkyl chains.

The first mutation tried was the L110G mutant, which failed due to low expression of the protein and difficult purification. Expression of the MmfR mutant L110A in *E. coli* BL21 star (DE3) cells encountered problem of growth inhibition and very low level or even no expression of the protein, similarly to the mutant Y144F, as well as purification problems. The problem was not solved even when the mutant L110A was expressed in *E. coli* BL21 star (DE3)pLysS cells. More surprisingly, the L110V mutant was found to be present in the insoluble fraction as shown in Figure 3.17. The above results indicate that probably Leu110 is vital for proper folding and stability of MmfR.

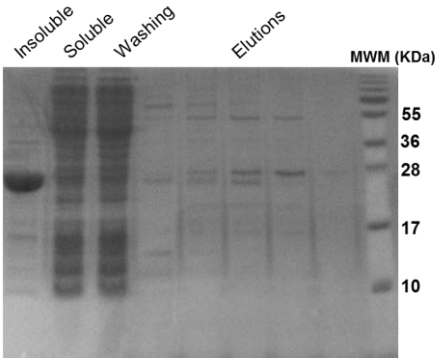


Figure 3.17. SDS-PAGE analysis in the attempt to purify the MmfR mutant L110V.

3.4.4 Effect of mutations on ligand binding

The activity of MmfR mutants, Y85F, Y85A, Q130E, Y144F and Y144A, was assessed using EMSAs. The DNA fragment used was the 194-bp sequence of the entire *mmfR-mmfL* intergenic region. The ligand used was AHFCA1, which showed the strongest interaction with MmfR as described in Section 3.3.2.

To assure the mutants were still able to bind DNA, a pre-EMSA was performed for each MmfR mutant in comparison with wild-type (WT) MmfR. With increasing concentrations of protein, it showed that all the mutations of MmfR in this study do not lead to altered DNA binding affinity for MmfR.

3.4.4.1. Tyr 85

The ability of MmfR mutants Y85F and Y85A and WT MmfR to bind DNA in the presence of AHFCA1 were then compared (Figure 3.18).

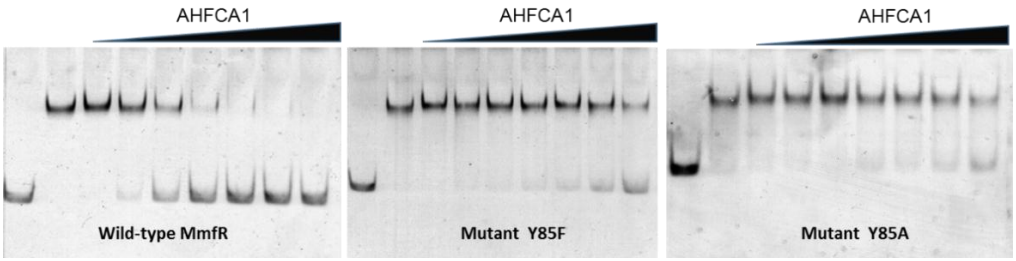


Figure 3.18. Comparison of the effect of AHFCA1 on protein:DNA (*mmfL-mmfR* fragment, 194 bp) dissociation when WT MmfR, MmfR Y85F and Y85A were used. Lane 1: DNA fragment only. Lanes 2 to 9: AHFCA1 at 0, 0.8, 4, 8, 14, 20, 40 and 100 nmol, respectively. Amount of DNA and proteins were kept constant (0.1 pmol and 1.8 pmol respectively).

Both of the MmfR mutants at Tyr 85, Y85A and Y85F, showed significantly reduced DNA release in the presence of AHFCA1 in comparison to WT MmfR. Interestingly, the Y85F mutant, which only lacks the hydroxyl group on the side chain, gave a similar result to the Y85A mutant. This highlights the importance of the hydroxyl group in the interaction of MmfR with its ligand. The result is highly consistent with the information about Tyr 85 obtained from the crystal structure.

3.4.4.2. Gln 130

The ability of MmfR mutant Q130E and WT MmfR to bind DNA in the presence of AHFCA1 were compared (Figure 3.19).

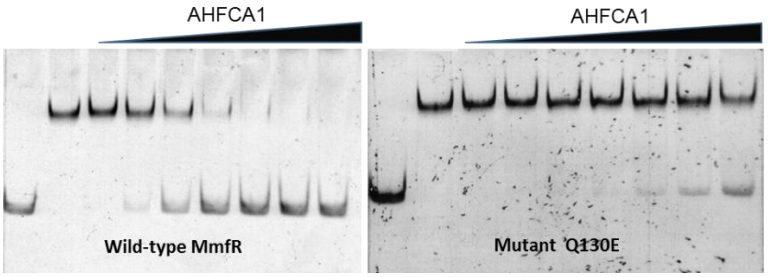


Figure 3.19. Comparison of the effect of AHFCA1 on protein:DNA (*mmfL-mmfR* fragment, 194 bp) dissociation when WT MmfR and MmfR Q130E mutant were used. Lane 1: DNA fragment only. Lanes 2 to 9: AHFCA1 at 0, 0.8, 4, 8, 14, 20, 40 and 100 nmol, respectively. Amount of DNA and proteins were kept constant (0.1 pmol and 1.8 pmol respectively).

The Q130E mutant showed dramatically reduced ability to respond to AHFCA1 and cause DNA release. Therefore, loss of the amide group of Gln 130 in this mutant

attenuates the ligand recognition and binding of AHFCAs by MmfR, which correlates well with the proposed role for Gln 130 from the crystal structure.

3.4.4.3. Tyr 144

The ability of MmfR mutants Y144F and Y144A and WT MmfR to bind DNA in the presence of AHFCA1 were compared (Figure 3.20).

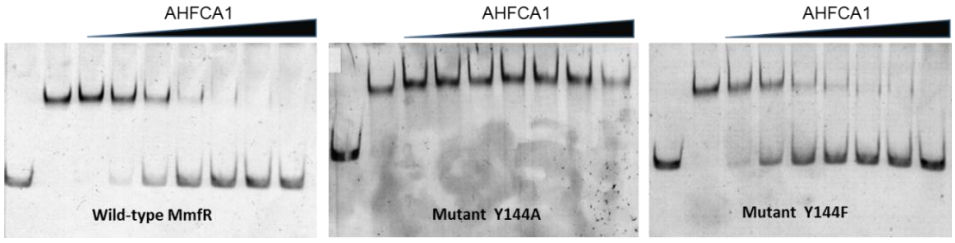


Figure 3.20. Comparison of the effect of AHFCA1 on protein:DNA (*mmfL-mmfR* fragment, 194 bp) dissociation when WT MmfR, MmfR Y144A and Y144F were used. Lane 1: DNA fragment only. Lanes 2 to 9: AHFCA1 at 0, 0.8, 4, 8, 14, 20, 40 and 100 nmol, respectively. Amount of DNA and proteins were kept constant (0.1 pmol and 1.8 pmol respectively).

Figure 3.20 (right image) shows that AHFCA1 retained its ability to release DNA from the mutant Y144F protein at a level that was comparable to, or possibly even more effective than, WT MmfR (left image). This strongly indicates that it is not the side chain of the residue Tyr 144 involved in the interaction with the AHFCA ligand. It is very likely that the backbone amide at that position participates in ligand recognition and binding with the hydrophilic portion of the AHFCAs, as proposed from the crystal structure. However, no AHFCA1-induced release of the DNA was observed for the mutant Y144A, probably due to alteration of the ligand binding pocket when Tyr 144 was replaced with a comparatively small residue.

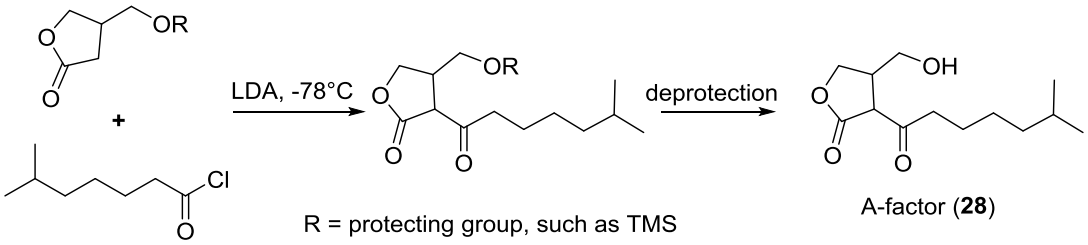
3.5 Ligand binding specificity of MmfR

As discussed in Section 1.2, there are two structural types of signalling molecules present in *S. coelicolor*, one of which is the γ-butyrolactone type SCBs, while the other is the furan type AHFCAs for the regulation of methylenomycin biosynthesis. In order to determine whether the regulation systems in *S. coelicolor* A3(2) are pleiotropic or pathway-specific, we investigated the ability of MmfR to interact with SCB1. Thus, we

chemically synthesised the γ-butyrolactone signalling molecule SCB1 and assessed the responsiveness of MmfR to it by EMSAs.

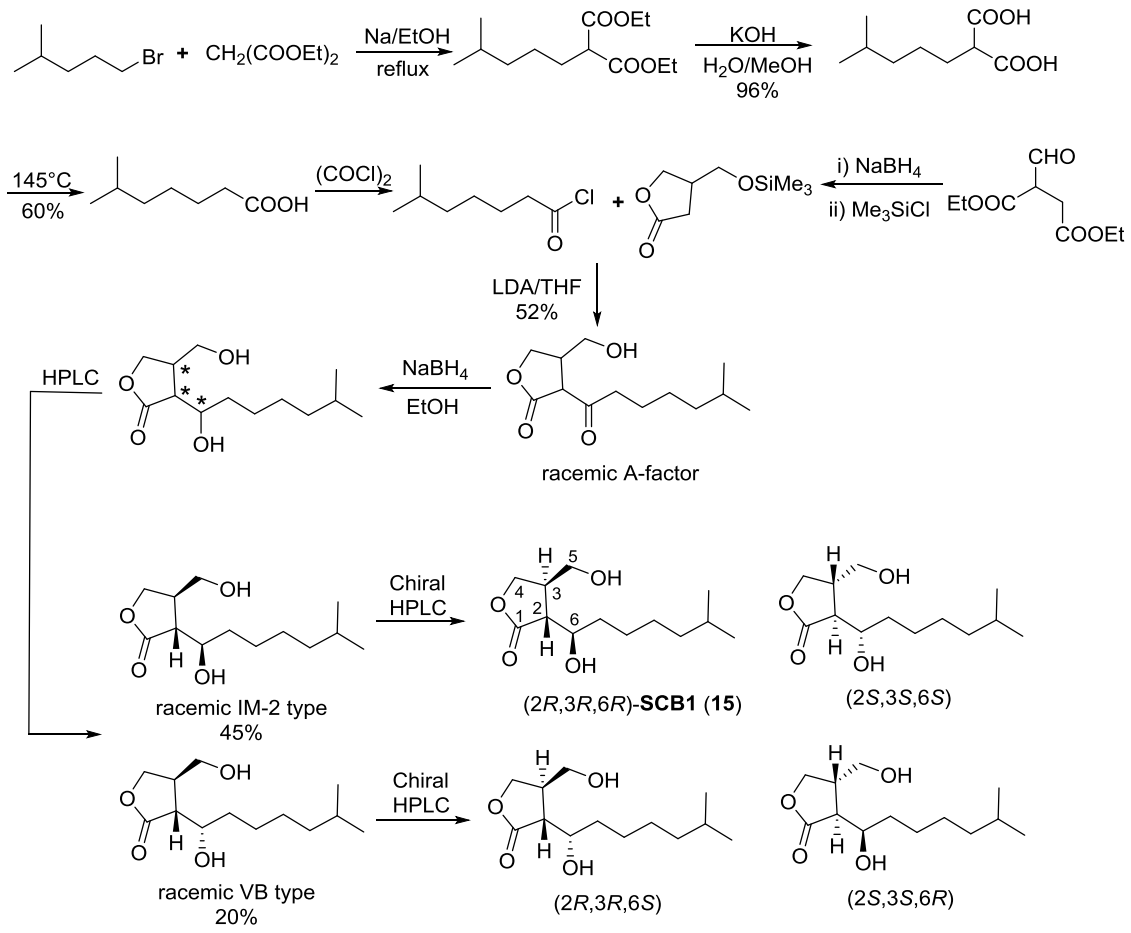
3.5.1 Synthesis of SCB1

The total synthesis of GBLs has been reported many times.[168,173-178] As A-factor was the first identified γ-butyrolactone signalling compound, most efforts have been made on its synthesis. A most common strategy involves acylation of a protected β-hydroxymethyl-γ-butyrolactone to introduce an α acyl group (Scheme 3.6). However, lots of evidence suggested that preparation of the protected β-hydroxymethyl-γ-butyrolactone is not easy and would be the key step in the route.



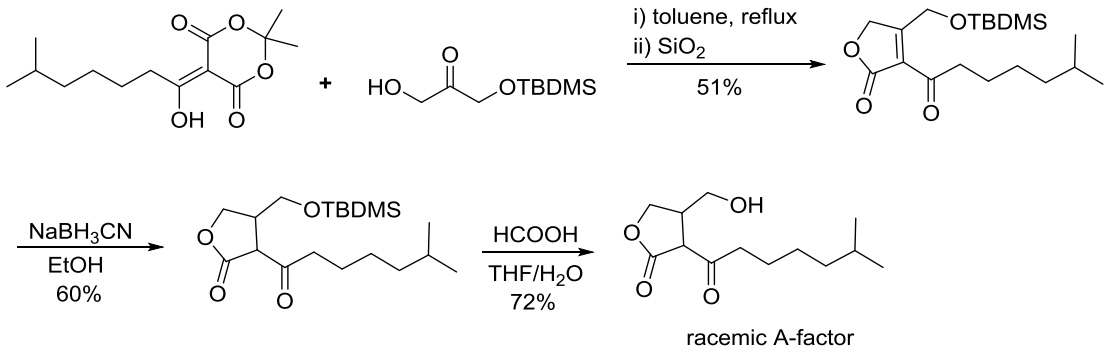
Scheme 3.6. Key steps in most reported synthesis of A-factor.

The structure of SCB1 **15**, (2*R*,3*R*,1'*R*)-2-(1'-hydroxy-6-methylheptyl)-3-hydroxymethyl butanolide, was first determined by Takano and co-workers in 2000 by comparing the purified metabolite with the chemically synthesised SCB1 and its stereoisomers.[57] Their synthetic route is shown in Scheme 3.7 by using the aforementioned strategy for constructing the α-position acylated β-hydroxymethyl-γ-butyrolactone in the presence of lithium diisopropylamide (LDA). Reduction of A-factor with $NaBH_4$ and separation by C_{18} HPLC gave the IM-2-type enantiomers (2*R*,3*R*,6*R* and 2*S*,3*S*,6*S*) and VB-type enantiomers (2*R*,3*R*,6*S* and 2*S*,3*S*,6*R*) of SCB1. They further separated each racemic mixture by chiral HPLC to yield optically pure compounds in order to characterise the absolute configuration of natural SCB1, which has been determined as 2*R*,3*R*,6*R*.



Scheme 3.7. Synthesis of SCB1 reported by Takano and co-workers.[57]

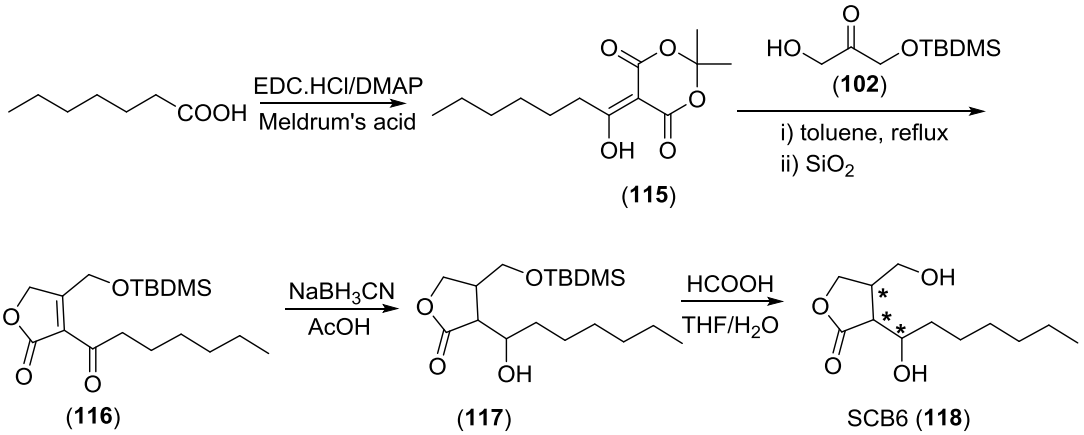
More recently, an alternative synthetic methodology was reported by Sello and co-workers for the synthesis of A-factor (Scheme 3.8).[168] Compared to previous syntheses, this approach provides a more convenient access to the core structure of GBLs using inexpensive and easily handled starting materials and reagents. It was also used in preparation of the protected butenolide for investigation of MmfL catalytic activity and discussed in Section 2.3. The alkene within the resulting butenolide is reduced using sodium cyanoborohydride and subsequent deprotection of the TBDMS group with acid affords racemic A-factor. Thus, an efficient stereoselective synthesis of A-factor and related GBLs has yet to be achieved.



Scheme 3.8. Synthesis of A-factor reported by Sello and co-workers.[168]

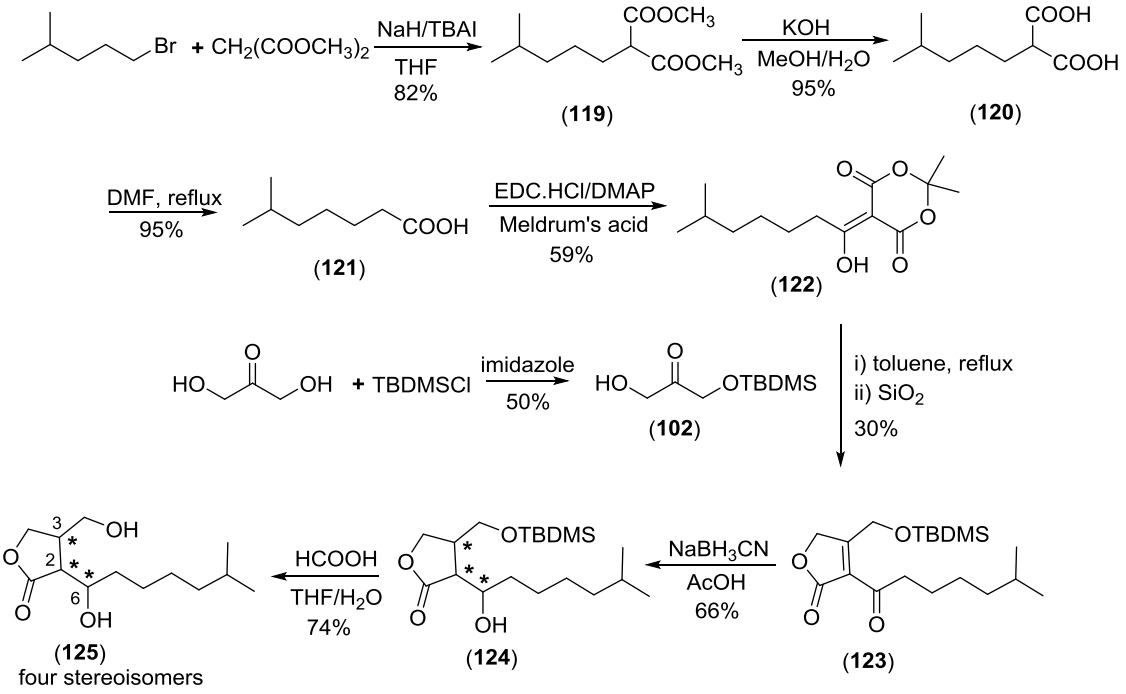
The structure of SCB1 is very close to A-factor except that SCBs possess a hydroxyl group instead of a ketone in their alkyl chains. This feature results in the requirement of an additional reduction step in their synthesis compared with A-factor. Thus the reported route by Sello *et al.* was slightly modified to adapt for our specific target molecule. For reduction of the alkene on the butenolide ring, Sello *et al.* used sodium cyanoborohydride in ethanol as the reducing condition, with strict control of the reaction time to selectively reduce the alkene without over-reduction of the ketone group on the side chain. In contrast to their synthesis, reduction of the alkene and ketone was required for SCB1 and so a more potent reducing agent was required. Selectivity for the ester was still required. So sodium cyanoborohydride and acetic acid was used and provided the desired alcohol after 8 hours in a yield of 66%.

For the purposes of optimisation, we first used heptanoic acid which is commercially available instead of 6-methylheptanoic acid, therefore resulting in the preparation of a SCB1 analogue **118**, namely SCB6 (Scheme 3.9).[60]



Scheme 3.9. Synthesis of SCB6 for optimisation purpose.

6-Methylheptanoic acid was prepared using a synthetic route similar to that used by Takano and co-workers. To simplify the preparation for acyl Meldrum's acid **122**, the combination of EDC and DMAP was used for coupling of 6-methylheptanoic acid with Meldrum's acid. This is same as the synthesis of acyl Meldrum's acid for NAC β-ketothioesters described in Section 2.2 and preparation of AHFCAs described in Section 3.3.1. Our synthetic route for SCB1 is summarised in Scheme 3.10.



Scheme 3.10. Synthesis of SCB1 and its stereoisomers in this study.

There are three stereogenic centers formed after the reduction. Similar to that reported by Takano and co-workers using $NaBH_4$, it resulted in formation of two pairs of enantiomers. The production of four stereoisomers instead of the possible eight is due to the steric hindrance of the molecule, which only results in products with the two groups at C-2 and C-3 position of the butyrolactone ring on opposite side. The products involve the IM-2 type enantiomers (2*R*,3*R*,6*R* and 2*S*,3*S*,6*S*) and VB type enantiomers (2*R*,3*R*,6*S* and 2*S*,3*S*,6*R*) of SCB1.

The separation of the two pairs of enantiomers were performed on column chromatography with silica gel after deprotection of the TBDMS group to yield the desired racemic SCB1. The 400 MHz ^{1}H NMR spectra of racemic SCB1 and the other VB type compounds are identical to the reported data, confirming that SCB1 was

successfully prepared. The difference between racemic SCB1 and their diastereoisomers in ^{1}H NMR is obvious (Figure 3.21). In addition, the spectra of SCB6 **118** also showed the same key features as SCB1. Given the qualitative purpose of our subsequent bioassays, racemic SCB1 was not further separated and directly applied to the following experiments.

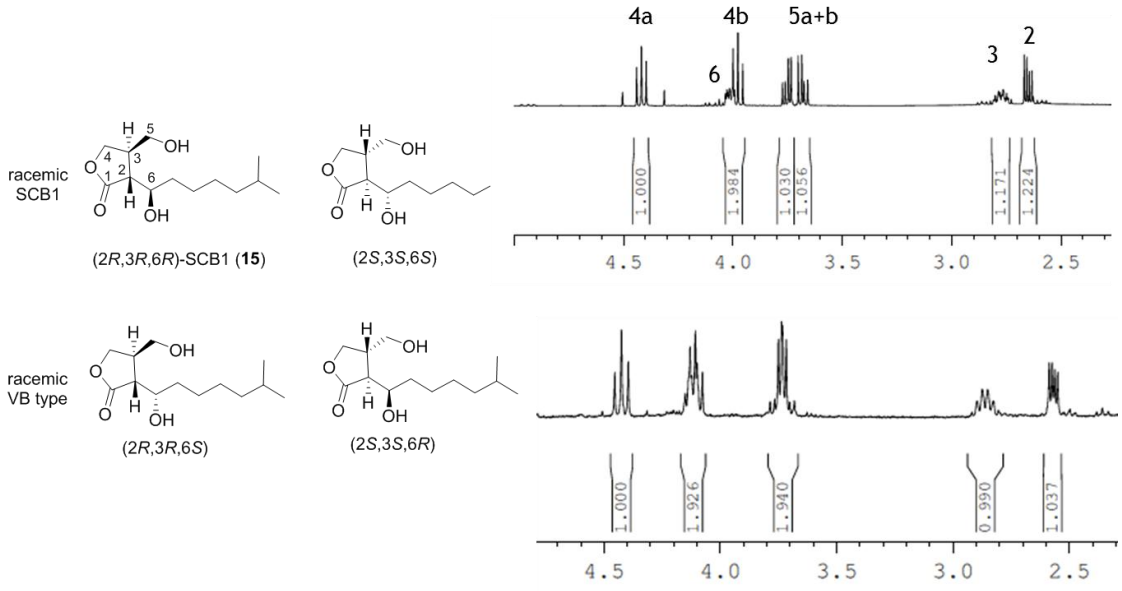


Figure 3.21. Comparison of the 400 MHz ^{1}H NMR spectra of racemic SCB1 and their diastereoisomers.

3.5.2 Interaction of MmfR with SCB1

In order to assess whether MmfR is responsive to SCB1, EMSAs were performed with the 194-bp DNA fragment of the entire intergenic region between *mmfR* and *mmfL*. The EMSA result with increasing concentrations of racemic SCB1 was compared to that with AHFCA1 (Figure 3.22).

In order to obtain reliable data, the concentration of racemic SCB1 was increased to 20 mM, which is four times higher than that previously used for other EMSAs. The DNA and protein were prepared from the same batch and their concentrations were kept constant.

The difference in ability to release MmfR from its operator is clear between SCB1 and AHFCA1. Addition of AHFCA1 resulted in a band shift caused by the DNA release from MmfR-DNA complex at a concentration of 0.4 mM or higher. However, the EMSA with same amounts of SCB1 did not lead to a band shift. In fact, no DNA band shift can be

seen even when the concentration of SCB1 was increased to 20 mM. This indicates that the γ-butyrolactone signalling molecule SCB1 could not bind to MmfR and initiate its release from the operator. The unresponsiveness of MmfR to γ-butyrolactone implies that its ligand binding is highly specific, supporting the hypothesis that GBLs and AHFCAs mediated regulation is pathway-specific in *S. coelicolor* A3(2).

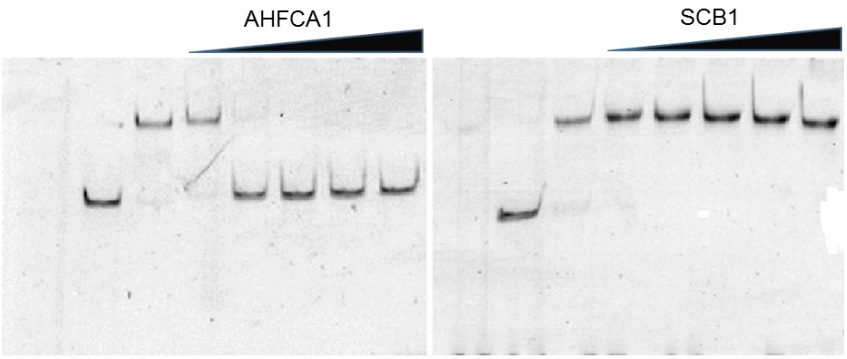


Figure 3.22. Interaction of MmfR with the 194-bp DNA fragment of the entire *mmfR-mmfL* intergenic region in response to increasing amounts of AHFCA1 and racemic SCB1. Lane 1: DNA fragment only. Lanes 2 to 7: AHFCA1 or racemic SCB1 at 0, 0.8, 8, 20, 40 and 400 nmol, respectively. Amounts of DNA and proteins were kept constant (0.1 pmol and 1.6 pmol respectively).

3.6 Investigation of a MmfR homologue, SgnR

As discussed in Section 1.3, SgnR has been identified from the *gbn* cluster in *S. venezuelae* as a homologue of MmfR with high sequence similarity. ARE sequences, which are situated in the region upstream of *gbnA*, the intergenic regions of *sgnR-sgnL* and *gbnR-sgnH*, are nearly identical to those identified in the *S. coelicolor mmy* gene cluster (Figure 1.15 and Table 1.1). Moreover, AHFCA-like compounds, including AHFCA6 and AHFCA7, were identified as the metabolites from the culture and proposed to be the autoinducers for gaburedin production from the *gbn* cluster.[60]

The high similarity in the regulation system between the *S. venezuelae gbn* cluster and the *S. coelicolor mmy* cluster intrigued us. In order to probe whether the regulatory system involving SgnR in the *gbn* cluster shares the same features as MmfR in methylenomycin production, we used the same methodology described previously to study the interactions between SgnR, the AREs and the identified AHFCA compounds, such as AHFCA6. The DNA-binding ability of SgnR for each of the AREs as well as their interaction with AHFCA6 was investigated *in vitro* by EMSAs.

3.6.1 Cloning of *sgnR*

The *sgnR* gene sequence was amplified by PCR from the SV4 F01 cosmid using primers designed to ensure directional cloning of the PCR products into pET151/D-TOPO vector as described in Section 6.3.13. Agarose gel electrophoresis was used to separate the desired PCR products from unused reactants and template DNA (Figure 3.23).

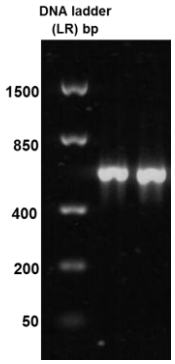


Figure 3.23. Agarose gel electrophoresis analysis of the *sgnR* PCR products (604 bp). (LR: low range DNA ladder)

The PCR product was excised from the gel and subsequently purified. The purified PCR product was then cloned into pET151/D-TOPO and *E. coli* TOP10 cells (IBA StarGate) were transformed with the new plasmid. Positive colonies containing the new plasmid pET151-*sgnR* were identified by selection on ampicillin plates, followed by restriction digest with *PstI* to give two fragment bands at 5760 and 1585 bp (Figure 3.24, left image). Insertion of the correct DNA sequence into the plasmid was also checked by PCR amplification of the gene of interest using the same primer pairs (Figure 3.24, right image).

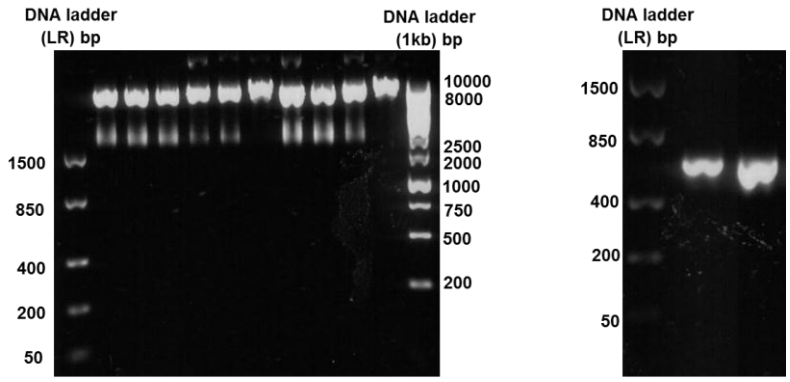


Figure 3.24. Agarose gel electrophoresis analysis of restriction digests (left) using *PstI* to confirm the constructed pET151-*sgnR* plasmid, and PCRs (right) using the constructed pET151-*sgnR* plasmid as template (604 bp). (LR: low range DNA ladder)

Plasmids that gave both PCR products of the expected size and restriction digest patterns consistent with the insertion of the genes of interest were sent for sequencing, with amplification using the T7 primer pair. Sequencing data confirmed that the plasmid contained the intact sequence of *sgnR*.

3.6.2 Overproduction and purification of recombinant His_6-SgnR protein

Chemically-competent *E. coli* BL21 star (DE3) cells were transformed with the pET151-*sgnR* plasmid. Total and soluble protein samples from the IPTG-induced expression of each gene were analysed by SDS-PAGE. The SgnR protein was overproduced and a clear band at the expected size (~26.2 kDa) was present in the total protein sample and the soluble fraction (Figure 3.25, left image). His_6-SgnR was purified by nickel affinity chromatography as described in Section 6.3.16. SDS-PAGE analysis of the purified His_6-SgnR protein confirmed it had the expected molecular weight (Figure 3.25, right image).

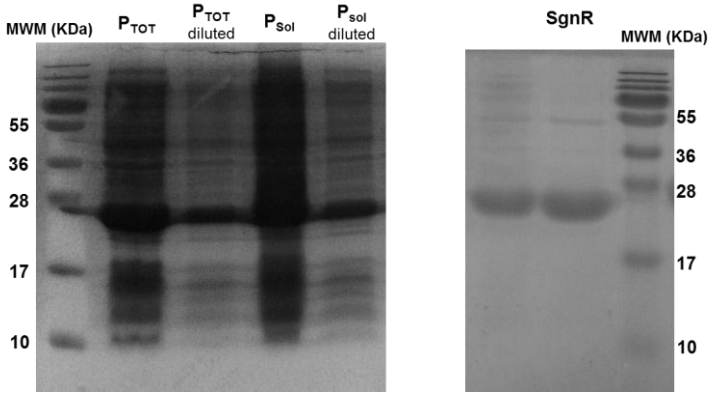


Figure 3.25. SDS-PAGE analysis of total (P_{TOT}) and soluble (P_{Sol}) protein fractions resulting from overproduction of His_6-SgnR in *E. coli* BL21star (DE3) (left), and His_6-SgnR after purification (right). (MWM: molecular weight marker)

The identity of the purified recombinant SgnR was further confirmed by LC-HRMS. The determined molecular weight is consistent with the expected for His_6-SgnR (26159.5 Da). (Figure 3.26).

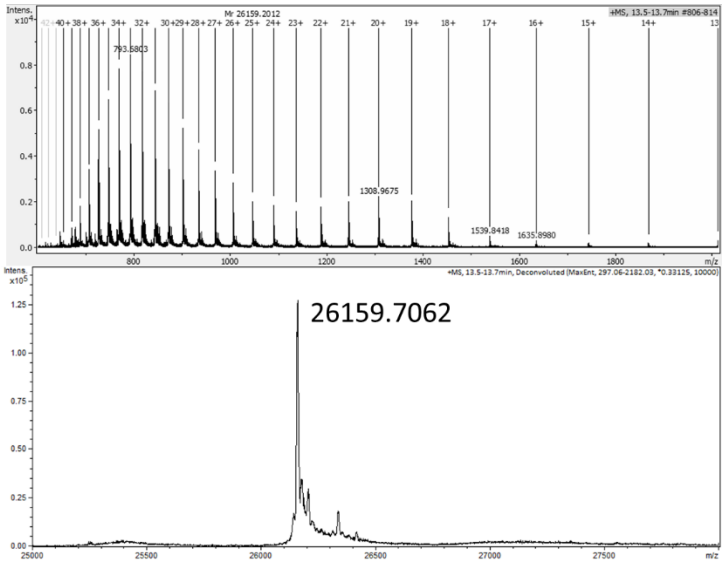
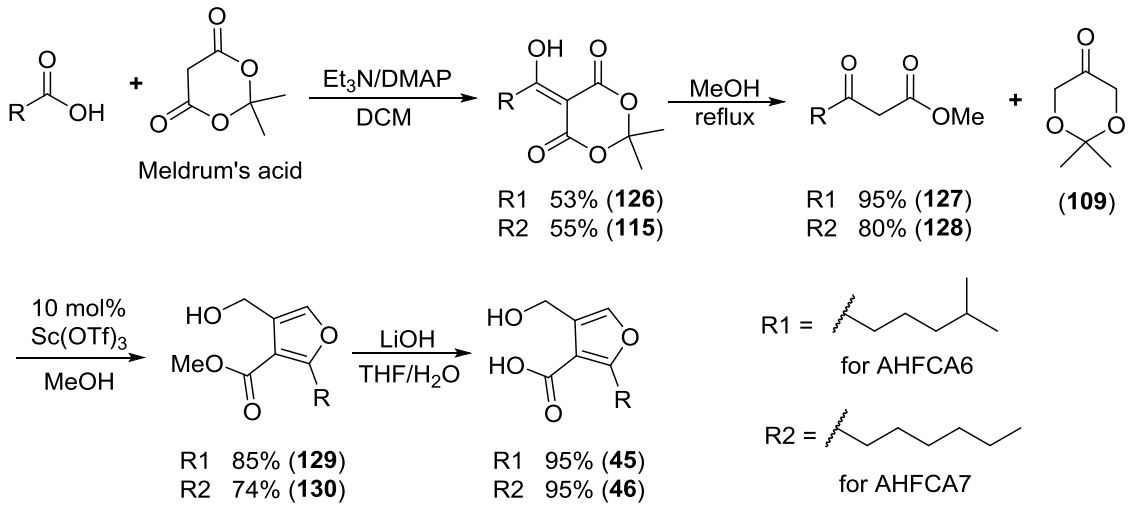


Figure 3.26. Measured (top) and deconvoluted (bottom) mass spectra of His_6-SgnR (calculated mass = 26159.5 Da).

3.6.3 Synthesis of AHFCA6 and AHFCA7

AHFCA6 and AHFCA7 were chemically synthesised using the same synthetic methodology as for AHFCA1 and AHFCA3 described in Section 3.3.1. As the alkyl chains in AHFCA6 and AHFCA7 were an isohexyl and a hexyl group, respectively, the starting acylation material for the acyl Meldrum's acid was 5-methylhexanoic acid for AHFCA6 and heptanoic acid for AHFCA7.

The combination of EDC and DMAP was used for the coupling between the carboxylic acid with Meldrum's acid. After formation of the β-ketoesters **127** and **128**, they were separately condensed with dihydroxyacetone acetonide **109** in the presence of scandium triflate to afford the corresponding furans **129** and **130** after purification by flash chromatography. The ester group within these furans was hydrolysed with lithium hydroxide, yielding the AHFCA6 **45** and AHFCA7 **46** as waxy solids. A summary of the synthesis of AHFCA6 and AHFCA7 is described in Scheme 3.11.



Scheme 3.11. Synthesis of AHFCA6 and AHFCA7.

The prepared AHFCA6 and AHFCA7 were used as synthetic standards in LC-MS comparison with the *S. venezuelae* metabolites for their structure elucidation (work by J. Sidda).[60]

3.6.4 EMSAs with SgnR

Due to the high sequence similarity of AREs present in the *gbn* cluster with those in the methylenomycin biosynthetic gene cluster (Table 1.1), the DNA-binding affinity of SgnR was first tested in EMSAs with the 194-bp DNA fragment of the entire *mmfR-mmfL* intergenic region on 6% non-denaturing polyacrylamide gel. A complete band shift was observed when the concentration of SgnR was increased to 1.8 pmol, suggesting it is capable of binding to the ARE sequence (Figure 3.27).

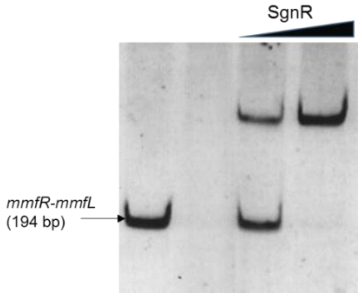


Figure 3.27. Interaction of SgnR with the 194-bp DNA fragment of the entire *mmfR-mmfL* intergenic region. Lane 1: DNA fragment only. Lanes 2 and 3: SgnR at 0.9 and 1.8 pmol, respectively. Amount of DNA was kept constant (0.1 pmol).

The entire intergenic sequence between *sgnR* and *sgnL* (121 bp) was then amplified by PCR from the cosmid SV4 F01 using primers listed in Table 6.1. The DNA fragments

containing the identified ARE1 sequence were analysed, purified from agarose gel (Figure 3.28) and sent for sequencing. The correct DNA fragments confirmed by sequencing were used in the EMSAs with SgnR on 6% non-denaturing polyacrylamide gel.

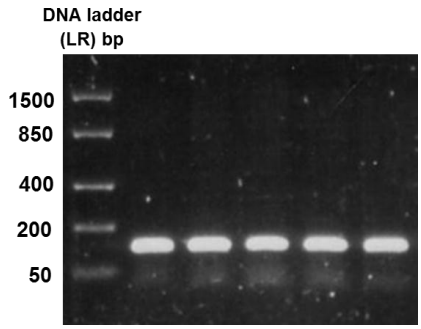


Figure 3.28. Agarose gel electrophoresis analysis of the PCR products of the entire *sgnR-sgnL* intergenic region (121 bp). (LR: low range DNA ladder)

The EMSA using the entire *sgnR-sgnL* intergenic DNA fragment containing the ARE1 sequence showed band shifts with addition of same concentrations of DNA and SgnR for the *mmfR-mmfL* intergenic region (0.1 and 1.8 pmol, respectively) (Figure 3.29). Moreover, the DNA fragments were released from the SgnR-DNA complexes in the presence of increasing amount of AHFCA6. This strongly indicates the ability of SgnR to bind the proposed operator between *sgnR* and *sgnL* and respond to AHFCA compounds.

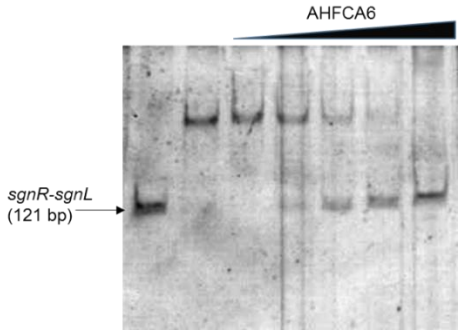


Figure 3.29. Interaction of SgnR with the 121-bp DNA fragment of the entire *sgnR-sgnL* intergenic region in response to increasing amounts of AHFCA6. Lane 1: DNA fragment only. Lanes 2 to 7: AHFCA6 at 0, 0.8, 8, 20, 40 and 400 nmol, respectively. Amounts of DNA and proteins were kept constant (0.1 pmol and 1.8 pmol, respectively).

The *in vitro* interaction of SgnR with the other two operators, one located upstream of *gbnA* and another between *gbnR* and *sgnH*, were also assessed with EMSAs with 10% non-denaturing polyacrylamide gels (Figure 3.30). In order to test the DNA binding

ability of SgnR more specifically, the DNA fragments used were short annealed double-stranded oligonucleotides (28 bp) containing the corresponding 18-bp ARE sequences. The specific oligonucleotides are listed in Table 6.4 in Chapter 6. It can be seen that SgnR could also bind to the other two operators and release the DNA in the presence of AHFCA6, providing increased understanding of the regulatory mechanism involving SgnR in the *gbn* cluter.

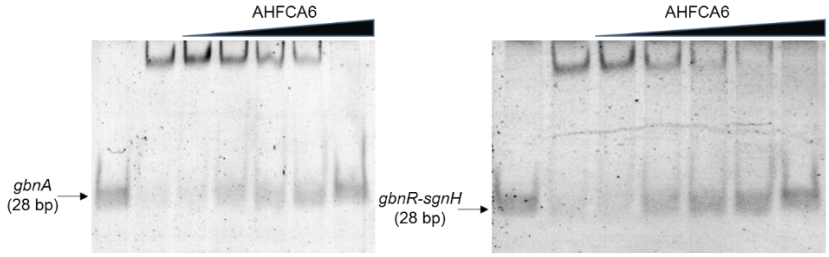


Figure 3.30. Interaction of SgnR with the 28-bp annealed DNA fragments containing ARE2 located upstream of *gbnA* (left) or ARE3 located in the *gbnR-sgnH* intergenic region (right) in response to increasing amounts of AHFCA6. Lane 1: DNA fragment only. Lanes 2 to 7: AHFCA6 at 0, 0.8, 8, 20, 40 and 400 nmol, respectively. Amounts of DNA and proteins were kept constant (0.8 pmol and 4.0 pmol respectively).

3.7 Key notes for studying *in vitro* DNA-protein-ligand interaction using EMSAs

From their extensive use in this study, EMSAs have proved to be an efficient way to study the *in vitro* interaction not only between DNA and protein but also their interaction with cognate ligands. Some general rules can be drawn from these studies as follows.

Firstly, the appropriate acrylamide concentration of the gel is important for running an effective EMSA, which mainly depends on the size of oligonucleotides. It can be varied, generally in the range of 5% to 25%. Lower percentage gels are better for bigger size DNA fragments, while higher percentages are needed for smaller fragments, but resulting in longer running time for a good image. We have learnt that 6% gel is enough for EMSAs with DNA fragments longer than 90 bp and 10% gel better for oligonucleotides less than 30 bp, such as the hairpin sequences and annealed 28-bp oligonucleotides used in this study.

Moreover, the ratio of DNA to protein should be optimised in assays. This could be easily determined by increasing the concentration of protein used for binding a certain amount of DNA fragments. It is necessary to determine the amount of protein that completely

binds the amount of DNA used whilst not having too great excess. This is especially important if a ligand compound will be added to test its ability to interact with the DNA-protein complex and release DNA. Presence of "excessive" protein results in consumption of more ligand to release the DNA. This is vital for EMSAs aiming at comparison of ligand affinity and thus amounts of DNA and protein should be kept constant in assays with each ligand. In an extreme case using too much protein, much more ligand would be needed, thereby leading to precipitation of the protein. This results in a "pseudo"-positive result that the compound used is effective in releasing DNA due to no retardation by inactive protein, which would be misleading. However, this can be easily distinguished from other samples during EMSAs due to its transformation to a cloudy mixture after addition of the ligand as show in Figure 3.31.

Figure 3.31. An example of protein-DNA-ligand reaction with precipitation of the protein in presence of the compound at a final concentration of 200 mM.

In addition, the DNA concentration should be added at the minimum level that allows clear visualisation of the DNA band under UV after GelRed staining. Higher amount of DNA leads to the use of more protein, thus requiring more ligand for DNA release. In the case of DNA fragments longer than 90 bp, it has been established that the amounts of DNA and protein for a good image are 0.1 pmol and 1.8 pmol, respectively. In the case of DNA fragments shorter than 30 bp, 0.8 pmol of DNA and 4.0 pmol of protein seems better. For DNA in the range of 30 to 90 bp, the amount and ratio of DNA and protein still needs to be determined due to lack of relevant experience in this study.

3.8 Summary

From *in vitro* EMSA assays, it has been confirmed that three conserved MARE sequences within the *mmy* gene cluster are the binding sites for MmfR, including the operator upstream of *mmyR* and the operators located in the *mmyY-mmyB, mmfR-mmfL* intergenic regions. The affinity order of MmfR to the three operators was demonstrated from the competitive binding assays as followed: the operator between *mmyY* and *mmyB* > the operator between *mmfR* and *mmfL* > the operator upstream of *mmyR*. This provides more information about the regulation mechanism involving MmfR in methylenomycin biosynthesis. Meanwhile, comparison of the ability of the five natural AHFCAs 1-5 on the dissociation of the MmfR-DNA complexes showed their order as: AHFCA1 > AHFCA3 > AHFCA5 ≈ AHFCA4 > AHFCA2.

Based on the X-ray crystal structure of MmfR-AHFCA2 complex, MmfR mutants at specific residues, Tyr85Ala, Tyr85Phe, Tyr144Ala, Tyr144Phe and Gln130Glu, were prepared using site-directed mutagenesis. *In vitro* interactions with EMSAs between DNA, AHFCA1 and these mutants have revealed the molecular basis for AHFCA recognition by MmfR. The hydroxyl group of Tyr85 and the amide group on the side chain of Gln130 were confirmed to play important roles in the interaction of MmfR with the ligand. Loss of these groups attenuates the ability of ligand to release DNA from MmfR. For Tyr144, it is likely that its backbone amide participates in ligand recognition and binding while its side chain may also be involved in the formation of the ligand binding pocket. Leu110 is probably vital for proper folding and stability of MmfR.

SCB1, a γ-butyrolactone signaling molecule from *S. coelicolor*, has been successfully prepared in this study. Current *in vitro* data obtained from EMSAs showed that SCB1 was unable to release MmfR from its operator. This demonstrates that ligand binding for MmfR is highly specific, supporting the hypothesis that regulation of signalling system with GBLs or AHFCAs is pathway-specific in *S. coelicolor* A3(2).

From the *in vitro* experiments with SgnR, it can be indicated that this MmfR homologue is able to bind both the MARE sequence from *S. coelicolor* and the three ARE sequences from *S. venezuelae gbn* gene cluster, including the operator upstream of *gbnA* and the operators located in the *sgnR-sgnL, gbnR-sgnH* intergenic regions. The SgnR-DNA complexes can be dissociated by addition of AHFCA6 to release the DNA on EMSAs,

indicating a similar regulation system as MmfR involved in the *mmy* gene cluster. Some key notes for investigating the DNA-protein-ligand interactions on EMSAs have also been summarised.

4. Watasemycin biosynthesis in *Streptomyces venezuelae* ATCC10712

4.1 Introduction to previous work

4.1.1 Upregulation of 'pyochelin-like' gene cluster

BldM is an orphan atypical response regulator that plays a crucial role in morphological differentiation.[5] It was noted that transcription of a cluster containing fifteen genes, *sven0503-sven0517*, is co-ordinately up-regulated in the *S. venezuelae bldM* mutant compared to the wild type strain based on DNA microarray analyses (Figure 4.1).[136] Additionally, Bibb *et al.* reported that transcription of the chloramphenicol biosynthetic gene cluster is significantly increased in the same *bldM* mutant.[179]

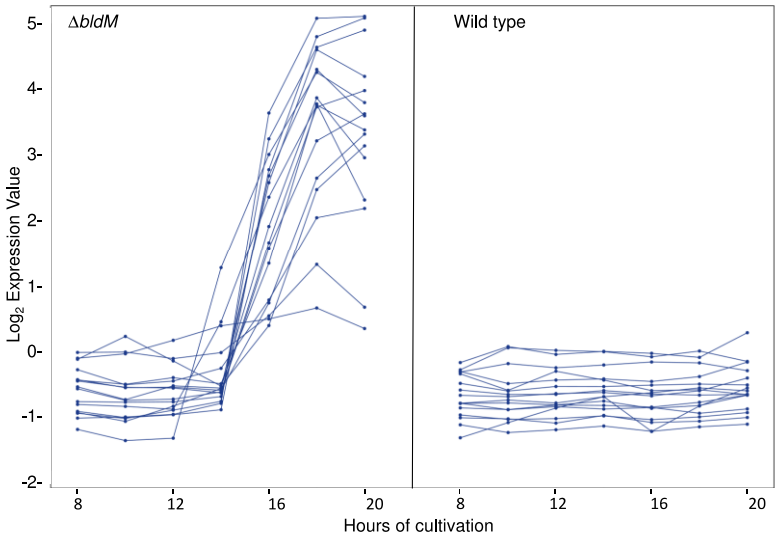


Figure 4.1. Microarray expression profiles of the cluster of genes, *sven0503-sven0517*, in *S. venezuelae*. The y axis shows normalised transcript abundance.[136]

Sequence analysis of the *sven0503-sven0517* gene cluster revealed that many genes in the cluster encode close homologues of the pyochelin biosynthetic enzymes in *S. scabies* (Figure 4.2 and Table 4.1). A comparison of pyochelin biosynthesis in the Gram-negative *P. aeruginosa* and the Gram-positive *S. scabies* can be found in Section 1.5.2.

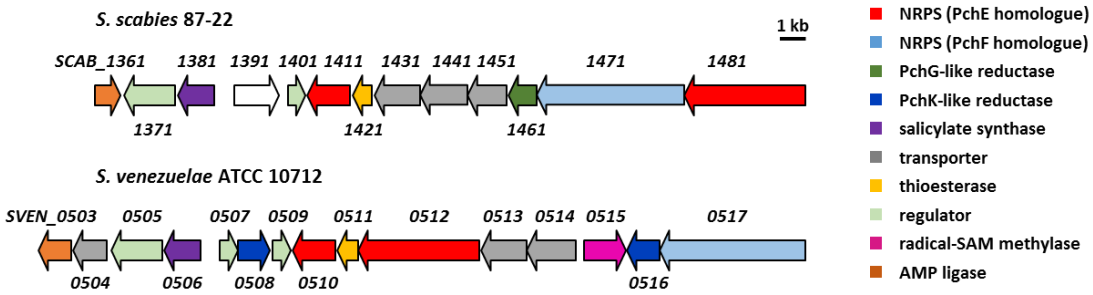


Figure 4.2. Organisation of the pyochelin and watasemycin biosynthetic gene clusters in *S. scabies* 87-22 and *S. venezuelae* ATCC10712, respectively. Genes that encode functionally analogous proteins are annotated in the same colour.

Sven0506 exhibits 82% sequence similarity to SCAB1381 (PchA/B) and is thus proposed to be a salicylate synthase. Sven0510 is highly homologous to SCAB1411 (PchD), which is the A domain responsible for salicylate activation in *S. scabies*, and Sven0511 is predicted to be a proofreading type II TE due to its sequence similarity to SCAB1421 (PchC). Sven0512 and Sven0517 are homologues of the NRPSs SCAB1481 (PchE) and SCAB1471 (PchF), respectively. It is noteworthy that Sven0512 appears to lack an MT-like domain, which is believed to catalyse the epimerisation of the stereocenter in the first thiazoline ring during pyochelin biosynthesis in *P. aeruginosa* and *S. scabies*. Both Sven0508 and Sven0516 show sequence similarity to PchK, which has been proposed to catalyse reduction of the second thiazoline ring in the *P. protegens* enantio-pyochelin biosynthetic pathway. There is one homologue of the AfsR-family regulator SCAB1371 (Sven0505) and two homologues of the TetR-family regulator SCAB1401 (Sven0507 and Sven0509). Finally, *sven0515* putatively encodes a class B RS methylase. As discussed in Section 1.6.1, these types of RS methylases are able to catalyse methylation reactions on un-activated carbon centres. This indicates that a C-methylation step may be involved in the biosynthesis of a pyochelin-like metabolite from the *sven0503-sven0517* gene cluster.

Table 4.1. Proposed functions of proteins encoded by the watasemycin biosynthetic gene cluster.[136]

Gene (accession number)	Size (bp)	Homologue (accession number); origin	Identity/ similarity (%)	Proposed function
sven0503 (CCA53790)	1250	SCAB1361 (CBG67360); *S. scabies* 87.22	44/51	AMP-ligase
sven0504 (CCA53791)	1247	Na^+/H^+ antiporter (WP_032767693); *Streptomyces* sp. CNS654	88/91	Na^+/H^+ antiporter
sven0505 (CCA53792)	1947	SCAB1371 (CBG67361); *S. scabies* 87.22	59/80	AfsR-family regulator
sven0506 (CCA53793)	1365	SCAB1381 (CBG67362); *S. scabies* 87.22	61/82	Salicylate synthase
sven0507 (CCA53794)	690	SCAB1401 (CBG67364); *S. scabies* 87.22	57/74	TetR-family regulator
sven0508 (CCA53795)	1212	*Trans*-acting enoyl reductase (AEV87025); *Actinoplanes* sp. SE50/110	36/46	Reductase
sven0509 (CCA53796)	687	SCAB1401 (CBG67364); *S. scabies* 87.22	70/88	TetR-family regulator
sven0510 (CCA53797)	1602	SCAB1411 (CBG67365); *S. scabies* 87.22	71/79	Salicyl-AMP-ligase
sven0511 (CCA53798)	801	SCAB1421 (CBG67366); *S. scabies* 87.22	55/63	Thioesterase
sven0512 (CCA53799)	4545	SCAB1481 (CBG67372); *S. scabies* 87.22	59/68	NRPS
sven0513 (CCA53800)	1710	SCAB1431 (CBG67367); *S. scabies* 87.22	69/82	ABC transporter permease/ATPase
sven0514 (CCA53801)	1842	SCAB1441 (CBG67368); *S. scabies* 87.22	68/79	ABC transporter permease/ATPase
sven0515 (CCA53802)	1605	Radical SAM domain protein (AEV87031); *Actinoplanes* sp. SE50/110	65/77	Radical SAM methyltransferase
sven0516 (CCA53803)	1227	*Trans*-acting enoyl reductase (AEV87025); *Actinoplanes* sp. SE50/110	48/56	Reductase
sven0517 (CCA53804)	5481	SCAB1471 (CBG67371); *S. scabies* 87.22	47/59	NRPS

4.1.2 Isolation and characterisation of watasemycin and thiazostatin from *S. venezuelae*

Dr. Yuki Inahashi, a Postdoc in the Challis group, successfully expressed the *sven0503-sven0517* gene cluster in a *Streptomyces* superhost *S. coelicolor* M1152 and carried out further LC-MS analyses. *Streptomyces* M1152 has been engineered to remove the actinorhodin (*act*), prodiginine (*red*), calcium-dependent antibiotic (*cda*) and coelimycin (*cpk*) gene clusters and contains a point mutation in *rpoB*, which encodes the RNA polymerase β-subunit, to pleiotropically increase the production of specialised metabolites.[42]

LC-MS analyses of the culture supernatant of the *S. coelicolor* M1152/SV-2_E03::*Ssp*I strain, which contains the *sven0503-sven0517* gene cluster, identified metabolites giving rise to ions with *m/z* corresponding to $[M+H]^+$ for the known 2-hydroxyphenylthiazolines thiazostatin **82**, watasemycin **83**, pyochelin **74**, aerugine **75** and pulicatins A/B **77**/**78**, respectively, in comparison to wild type *S. coelicolor* M1152 (Figure 4.3). This was further confirmed by LC-HRMS analysis carried out in this study (Figure 4.4).

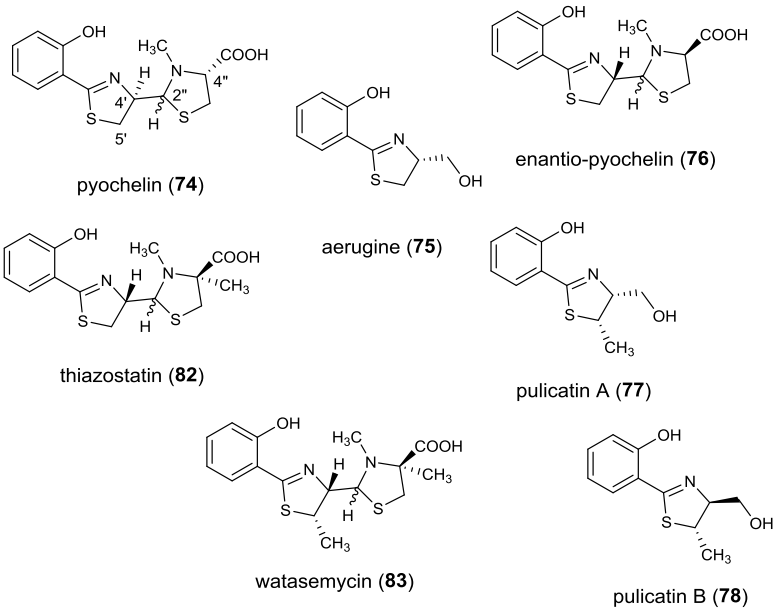


Figure 4.3. Proposed structures of 2-hydroxyphenylthiazolines produced from the *sven0503-sven0517* gene cluster in *S. venezuelae* ATCC10712 based on high resolution LC-MS analyses.

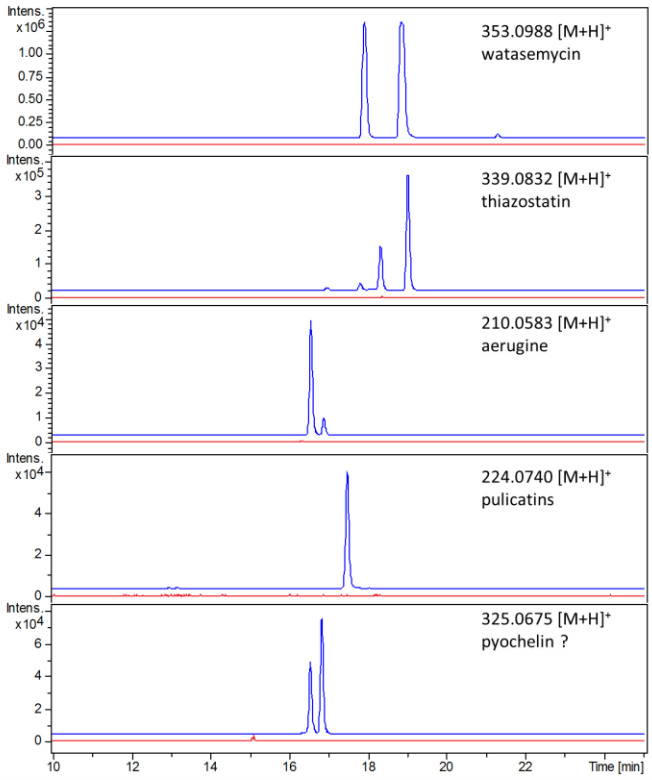


Figure 4.4. Extracted ion chromatograms (EICs) at *m/z* = 353.0988, 339.0832, 210.0583, 224.0740 and 325.0675, corresponding to $[M+H]^+$ for thiazostatin, watasemycin, aerugine, pulicatins A/B and pyochelin, respectively, from LC-HRMS analyses of the ethanol extract of *S. coelicolor* M1152/SV-2_E03::*Ssp*I culture supernatant (blue traces), in comparison to wild type *S. coelicolor* M1152 (red traces). Further work in this study would give more information about the peak corresponding to $[M+H]^+$ for pyochelin (see Section 4.2).

The main metabolic products with molecular formulae corresponding to thiazostatin and watasemycin were extracted and isolated from the culture supernatant of *S. coelicolor* M1152/SV-2_E03::*Ssp*I strain by Dr. Yuki Inahashi. NMR spectroscopic analysis confirmed their identity by comparison to the reported data for thiazostatins A and B **82**, and watasemycins A and B **83**, isolated from *S. tolurosus* and *Streptomyces* sp. TP-A0597, respectively.[116-117] Both thiazostatin and watasemycin are close analogues of pyochelin **74** and enantio-pyochelin **76**. Thiazostatin differs only by the addition of a methyl group at C-4" of the thiazolidine ring while watasemycin, besides this, has an additional methyl group at C-5' on the thiazoline ring. Other 2-hydroxyphenylthiazolines putatively identified by LC-MS are pyochelin, aerugine and pulicatins A/B.

In order to investigate the biosynthesis of thiazostatin and watasemycin, characterisation of the metabolite with molecular formula identical to pyochelin was pursued. As the structure and biosynthesis of pyochelin **74** and enantio-pyochelin **76** have been well

studied and characterised and are potential metabolites in biosynthesis, they would serve as a model for investigating these pyochelin-like compounds. Section 4.2 will focus on synthesis and characterisation of the metabolite with same molecular formula as pyochelin.

As mentioned above, two putative thiazoline reductases, Sven0508 and Sven0516, are present in the gene cluster, which is different from the *P. aeruginosa* or *S. scabies* pyochelin biosynthetic pathway. Specific roles of these two enzymes in watasemycin biosynthesis will be described in Section 4.3. Moreover, the absolute stereochemistry of thiazostatin and watasemycin has not been previously characterised, with only relative stereochemistry assigned. Section 4.4 will focus on reassignment of the relative and absolute configuration of these compounds based on the obtained experimental results and analysis. Section 4.5 will focus on studying the function of Sven0515, a putative class B RS methylase, and the stereochemical course of the methylation reaction it catalyses. A proposed biosynthetic pathway for watasemycin will be described in Section 4.6.

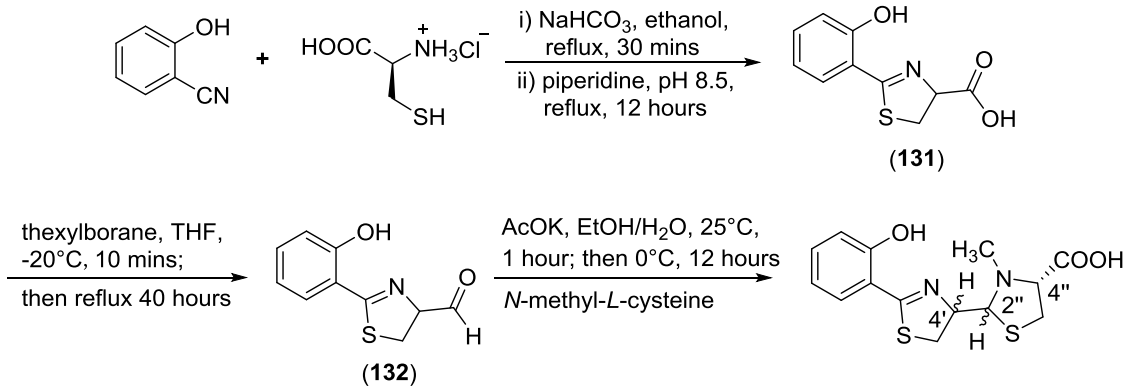
4.2 Identification and characterisation of isopyochelin

4.2.1 Synthesis of pyochelin

Assuming that the metabolite with molecular formula identical to pyochelin was actually pyochelin, we aimed to chemically synthesise pyochelin and its stereoisomer neopyochelin based on literature procedures.[180-182] Then these compounds would be used to compare to the pyochelin-like metabolite from *S. coelicolor* M1152/SV-2_E03::*Ssp*I culture supernatant using LC-MS.

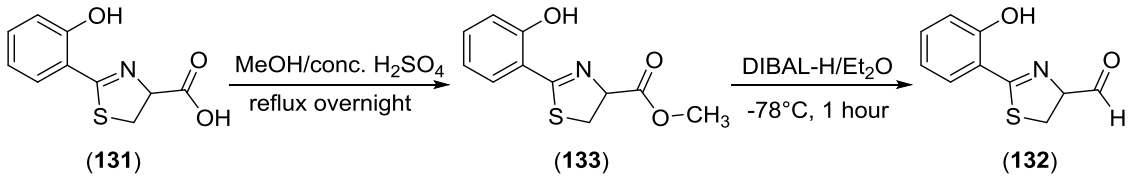
The first total synthesis of pyochelin was achieved in a three-step procedure by Ankenbauer *et al.* in 1988 with a total yield of around 8% as shown in Scheme 4.1.[180] Carboxylic acid **131** was prepared from 2-hydroxybenzonitrile and L-cysteine hydrochloride in high yield. However, in their attempts to synthesise aldehyde **132**, a variety of reagents did not work except reduction of the carboxylic acid **131** by thexylborane. Although the aldehyde **132** was produced, the yield of this step was very poor (15%). The low yield was attributed to the major product being the alcohol resulting from over-reduction of the aldehyde. Moreover, the aldehyde **132** was found to be

exceedingly labile. Condensation of the aldehyde **132** and *N*-methyl-L-cysteine hydrochloride was the final step to provide three stereoisomers of pyochelin: pyochelin I, II and neopyochelin, with another stereoisomer not mentioned at that time.



Scheme 4.1. First total synthesis of pyochelin reported by Ankenbauer *et al*.[180]

The synthetic route was later modified by Rinehart *et al.* with successful assignment of the absolute stereochemistry of the pyochelins.[181] The reduction step to yield the aldehyde **132** was reported to be significantly improved by using diisobutylaluminium hydride (DIBAL-H) for reduction of the methyl ester **133** at -78 ℃ for 1 hour (Scheme 4.2). Again, the aldehyde **132** proved too unstable to provide acceptable analytical data. The other steps were all same as that reported in the first total synthesis. The aldehyde was immediately used for the condensation with *N*-methyl-L-cysteine to give four of the eight possible pyochelin stereoisomers: pyochelins I and II **74**, and neopyochelins I and II (Figure 4.5). It was reported that the diastereomeric pyochelins cannot be separated in pure form due to their rapid interconversion. Thus, assignment of their stereochemistry was mainly based on spectral data of pyochelin methyl esters, which are comparably more stable and can be purified, and the X-ray structure of a mutasynthetic pyochelin analogue methyl ester.



Scheme 4.2. Improved synthesis of the aldehyde **132** by Rinehart *et al*.[181]

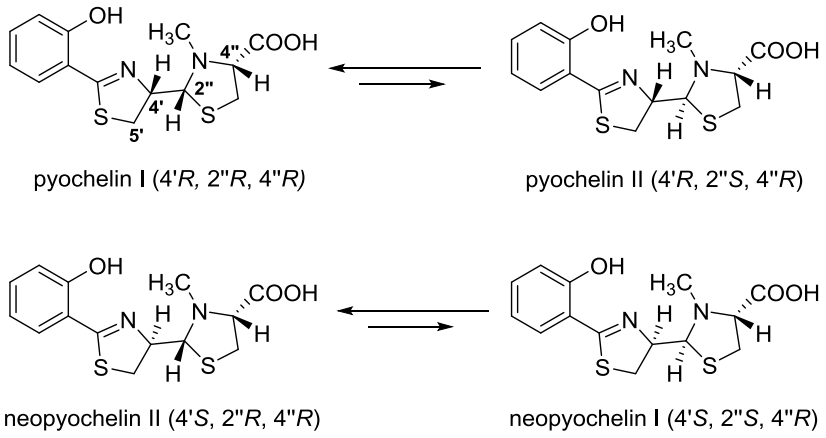
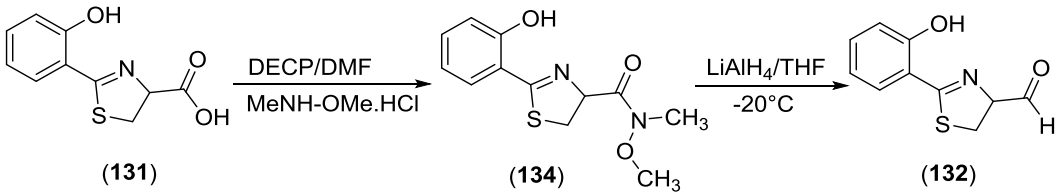


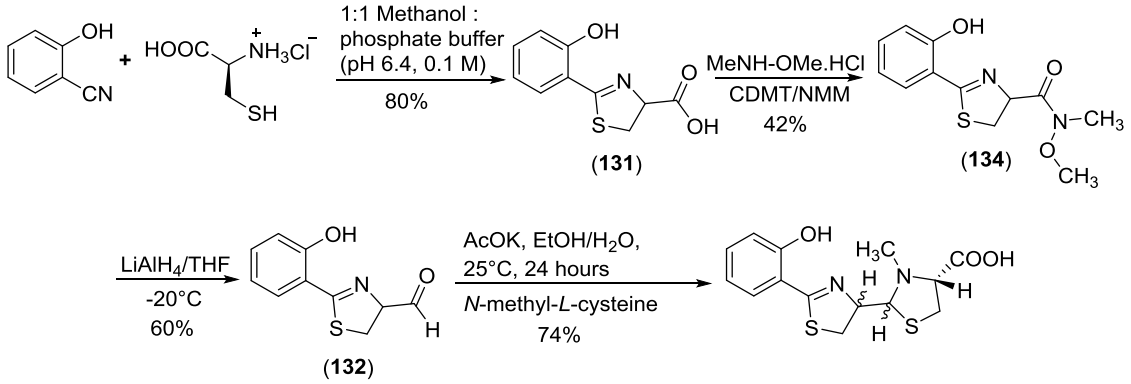
Figure 4.5. The structure of synthetic pyochelin and neopyochelin.

To further improve the synthesis of pyochelin, Zamri and co-workers elected to use alternative reaction conditions for the first step, i.e. condensation of 2-hydroxybenzonitrile with L-cysteine.[182] The reaction was performed in a 1:1 mixture of phosphate buffer 0.1 M, pH 6.4 and methanol at 50 ℃ for 4 days, in which epimerisation of the carboxylic acid **131** at C-4' due to the very basic pH conditions caused by piperidine could be avoided. They also used lithium aluminium hydride for reduction of *N*-methoxy-*N*-methyl hydroxamate **134** to prepare aldehyde **132** (Scheme 4.3). The hydroxamate **134** was prepared using diethyl cyanophosphonate (DECP) as the coupling agent for acid **131** and *N,O*-dimethylhydroxylamine in a yield of more than 90%. This synthetic approach represents the most efficient synthesis of pyochelin to date, and has been used as the basis for the synthesis of pyochelin and analogues in our study. However, the coupling agent DECP they used for preparation of hydroxamate **134** is known to be highly toxic and moisture-sensitive and therefore required careful handling.



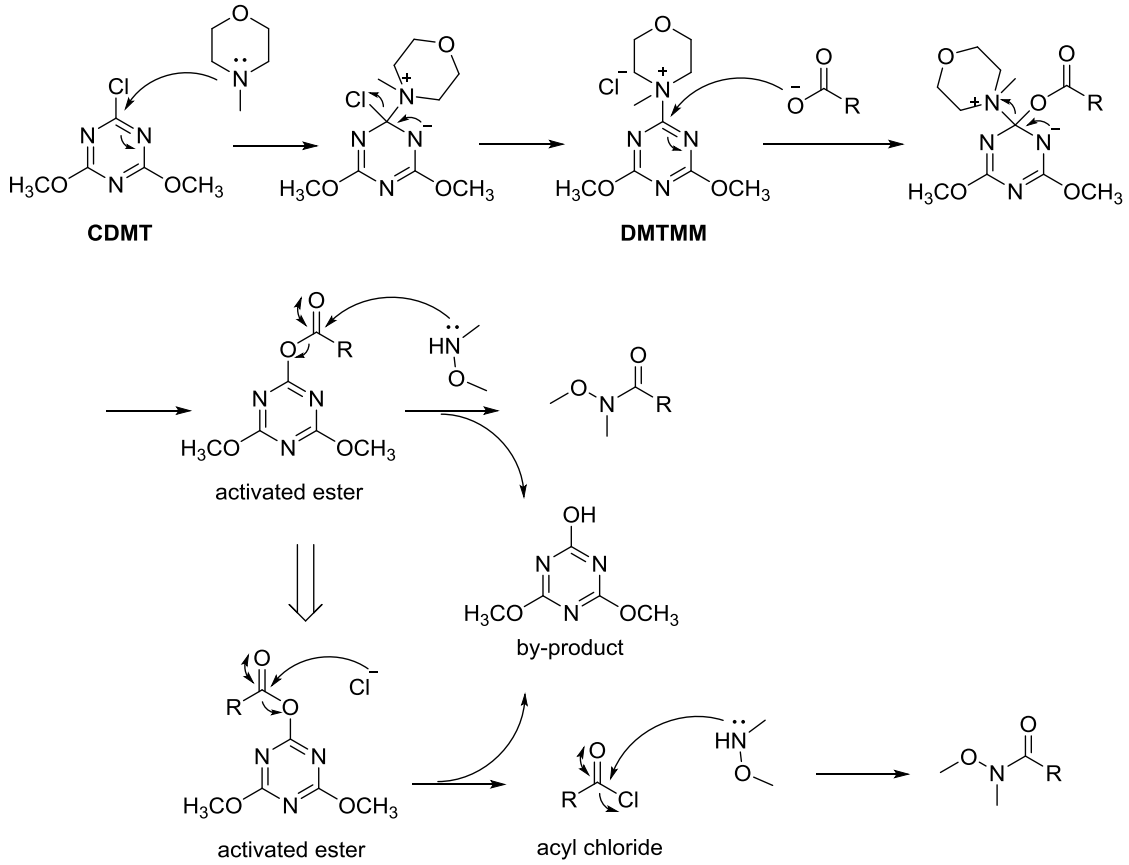
Scheme 4.3. Improved synthesis of aldehyde **132** by Zamri and co-workers.[182]

Our synthetic route for pyochelin was slightly modified from that reported by Zamri *et al.* using 2-chloro-4,6-dimethoxy-1,3,5-triazine (CDMT) to couple acid **131** and *N,O*-dimethylhydroxylamine, in the presence of *N*-methylmorpholine (NMM), as shown in Scheme 4.4.



Scheme 4.4. Synthetic route to pyochelin used in this study.

CDMT is a commercially available triazine-based coupling reagent, which works by converting carboxylic acids to activated triazine esters. It requires the use of a tertiary amine base, such as NMM, for acid activation, and the coupling can be carried out in solvents such as THF, DMF, and EtOAc. The mechanism of its participation in coupling reactions remains unclear. However, activation of the carboxylic acid *via* activated ester or *via* acyl chloride have been proposed (Scheme 4.5).[183] *In situ* NMM and CDMT form the intermediate 4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholinium chloride (DMTMM), which can also be isolated and used as a coupling reagent independently.[184] In each cycle of catalysis, 1-hydroxy-3,5-dimethoxytriazine is generated as a by-product. CDMT coupling afforded *N*-methoxy-*N*-methyl hydroxamate **134** in a yield of 42%. The modest yield seems to be due to generation of several by-products which made the product purification by column chromatography difficult.



Scheme 4.5. The proposed mechanism for CDMT-mediated coupling of acid **131** and *N,O*-dimethylhydroxylamine.

Other steps used in the synthesis were same as those reported by Zamri *et al*.[182] The present synthesis gives only four stereoisomers instead of the possible eight allowed by the three chiral centres, C-4', C-2" and C-4", in pyochelin. Since there is no report of isomerisation at C-4" during thiazolidine syntheses under the present conditions, the fixed centre, C-4", is assigned as *R* configuration, since it derives from *N*-methyl-L-cysteine. The observed four diastereoisomers are thus due to isomerism at C-4' and C-2" under the reaction conditions. The fact that the configuration at C-4" in pyochelin and the α-carbon of *N*-methyl-L-cysteine used in the condensation are identical not only benefits assignment of the pyochelin stereochemistry, as described by Rinehart *et al*., but also aids in the study of other compounds of similar structural types.

4.2.2 LC-MS analysis of synthetic pyochelin

Due to the rapid interconversion of the pyochelin acids, synthetic product was used to compare to the natural metabolites by LC-MS without further purification. Two of the synthetic stereoisomers, pyochelins I and II, correspond to the two naturally occurring

pyochelins **74** found in *S. scabies* and *P. aeruginosa*. However, none of these four diastereomeric compounds had the same retention time as the metabolite with molecular formula corresponding to pyochelin (Figure 4.6). The other four diastereoisomers of the possible eight which are *S* configuration at C-4" would be enantiomers of these four diastereoisomers we obtained, thus they would have the same retention times. Therefore, it can be concluded that the metabolite is neither pyochelin nor any of its stereoisomers.

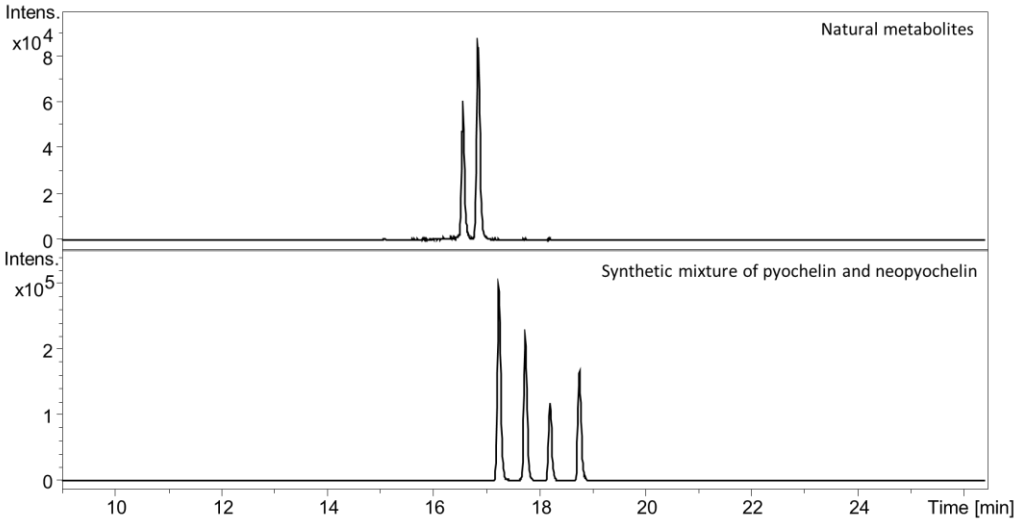
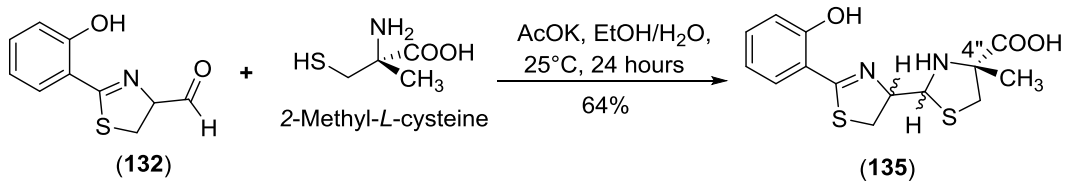


Figure 4.6. EICs at *m/z* = 325.0675, corresponding to $[M+H]^+$ for the metabolite (top trace) from LC-HRMS analyses of the ethanol extract of *S. coelicolor* M1152/SV-2_E03::*Ssp*I culture supernatant, and synthetic mixture of pyochelins and neopyochelins (bottom trace), including two stereoisomers identical to natural pyochelin from *S. scabies* 87-22 and *P. aeruginosa*.

4.2.3 Synthesis and characterisation of isopyochelin

Considering both thiazostatin and watasemycin have a methyl group at C-4", we hypothesised that the metabolite was a structural isomer of pyochelin **74** in which C-4" is methylated instead of the nitrogen atom of the thiazolidine. To test this hypothesis, an authentic standard of the C-4" methylated isomer of pyochelin, named as isopyochelin, was synthesised from condensation of the aforementioned aldehyde **132** with 2-methyl-L-cysteine, which cannot epimerise during the synthesis (Scheme 4.6).



Scheme 4.6. Synthesis of a mixture of isopyochelin stereoisomers **135** with *R* configuration at C-4".

A mixture of four synthetic diastereomeric compounds **135** with 4"*R* configuration were compared with the metabolites with *m/z* of $[M+H]^+$ = 325.0675 in the culture extract. Two of the synthetic stereoisomers had the same retention time as the metabolites in the extract (Figure 4.7), thus showing isopyochelin is a metabolic product of the gene cluster.

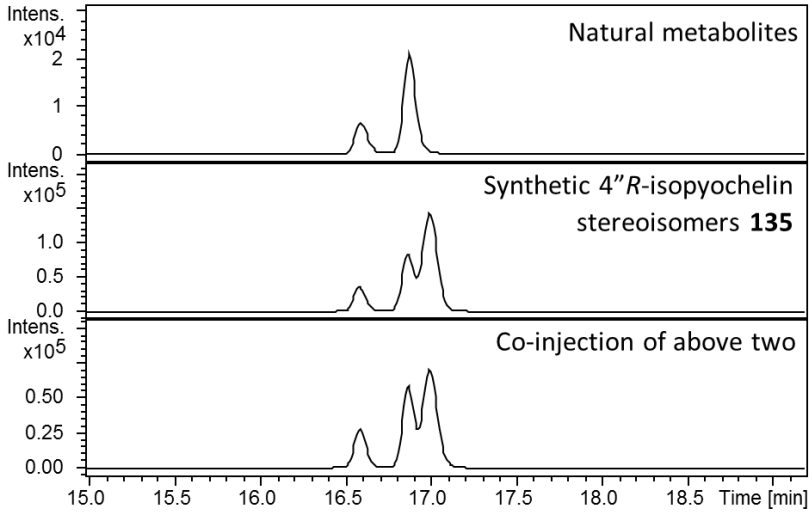
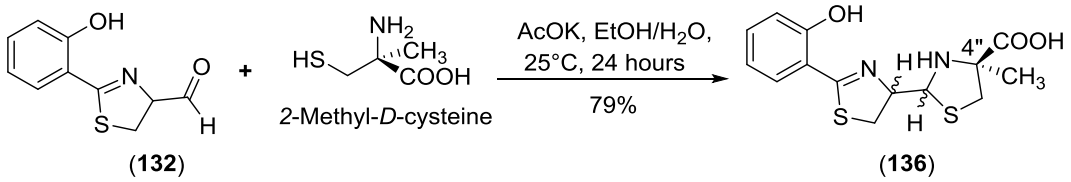


Figure 4.7. EICs for *m/z* = 325.0675, corresponding to $[M+H]^+$ for isopyochelin, from LC-HRMS analyses of the ethanol extract of *S. coelicolor* M1152/SV-2_E03::*Ssp*I culture supernatant (top trace), the authentic synthetic standard of the isopyochelin stereoisomers from 2-methyl-L-cysteine (middle trace) and the extract to which an approximately equimolar quantity of the synthetic standard has been added (bottom trace).

In order to assign the absolute configuration of isopyochelin, we further synthesised the four stereoisomers resulting from condensation of aldehyde **132** with 2-methyl-D-cysteine. This yielded isopyochelin stereoisomers **136** with 4"*S* configuration (Scheme 4.7).



Scheme 4.7. Synthesis of a mixture of isopyochelin stereoisomers **136** with *S* configuration at C-4".

The enantiopurity of the two commercially-sourced amino acids, 2-methyl-L-cysteine and 2-methyl-D-cysteine, was confirmed by CD studies prior to commencing the synthesis. Solutions containing an equimolar quantity of each amino acid showed mirror image CD spectra, confirming their opposite absolute stereochemistry (Figure 4.8).

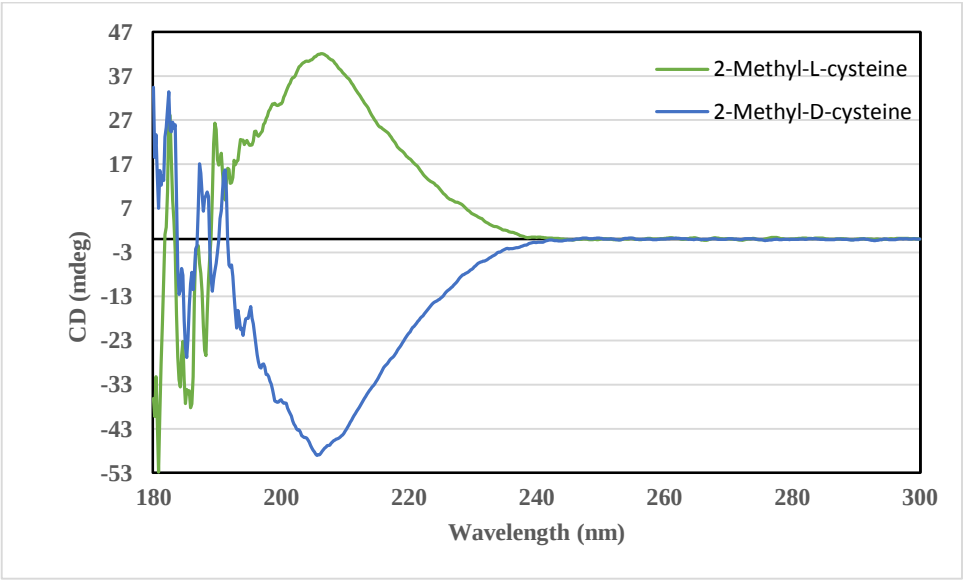


Figure 4.8. CD spectra of commercially available 2-methyl-L-cysteine and 2-methyl-D-cysteine used in the synthesis of isopyochelin.

LC-MS analyses on a homochiral stationary phase of all the synthetic products and the natural metabolites was performed using a ChiralPAK IE analytical column packed with amylose *tris* (3,5-dichlorophenylcarbamate). Both synthetic products were injected separately, followed by co-injection with the ethanol extract of the culture supernatant to determine which included the stereoisomers corresponding to the metabolite (Figure 4.9). Synthetic stereoisomers of isopyochelin with *R* configuration at C-4" can be completely separated on the column. None of these had the same retention time as the natural products. On the other hand, two of the synthetic diastereoisomers with *S* configuration at C-4" had identical retention times to the natural products, although only three peaks

were observed for the synthetic compounds, indicating incomplete separation on the column. Thus, the absolute configuration at C-4" in isopyochelin was established as *S*.

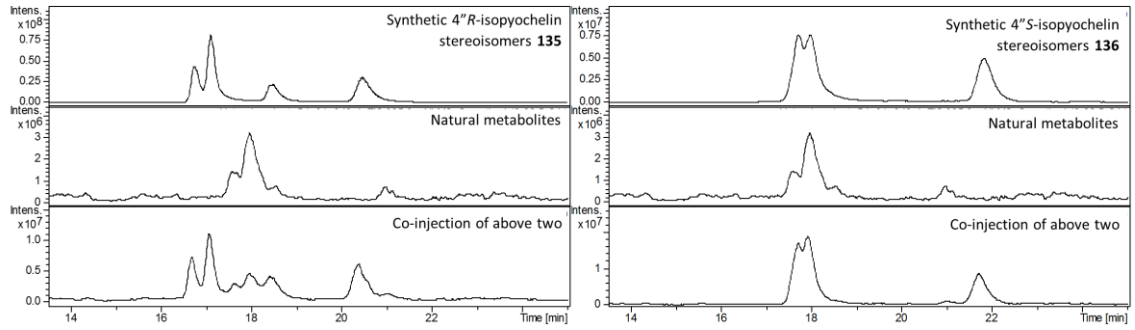


Figure 4.9. EICs at *m/z* = 325.0 (corresponding to $[M+H]^+$ for isopyochelin) from LC-HRMS analyses of the synthetic stereoisomers of isopyochelin and the natural metabolites from the ethanol extract of *S. coelicolor* M1152/SV-2_E03::*Ssp*I culture supernatant on a homochiral stationary phase. Comparison with 4"*R*-isopyochelins **135** (left) and 4"*S*-isopyochelins **136** (right) is shown separately.

4.3 Roles of Sven0508 and Sven0516 in the biosynthetic pathway

As discussed in Section 4.1.1, both *sven0508* and *sven0516* putatively encode PchK-like thiazoline reductases in the *sven0503-sven0517* gene cluster. However, it was proposed that only one thiazoline reductase is necessary for the reduction step in watasemycin and thiazostatin biosynthesis since they have the same scaffold structure as pyochelin and enantio-pyochelin. Therefore, individual in-frame deletions of these genes were constructed by Dr. Yuki Inahashi to investigate their roles in the production of 2-hydroxyphenylthiazolines in *S. venezuelae*.

In order to see whether and how production of the 2-hydroxyphenylthiazoline metabolic products would be affected by gene deletion of *sven0508* or *sven0516*, the *S. coelicolor* M1152/SV-2_E03*::Ssp*I and Δ*sven0508* and Δ*sven0516* mutant strains were grown on both mannitol soya flour medium (SFM) and yeast-extract dextrose (YD) medium. Rich medium SFM is commonly used as solid medium for spore production, with mannitol being the primary carbon source. YD medium is a nutritionally complete medium, in which yeast extract is the source of nitrogen and growth factors while dextrose provides an energy source for the growth of microorganisms. YD medium was used as liquid medium for heterologous expression of specialised metabolites in this study.

4.3.1 Sven0516 is required for 2-hydroxyphenylthiazoline biosynthesis

The ethanol extract of the culture broth from the Δ*sven0516* mutant was analysed by LC-HRMS. None of the 2-hydroxyphenylthiazoline metabolic products could be detected in the extract from the Δ*sven0516* mutant (Figure 4.10). Thus it can be concluded that the PchK homologue, Sven0516, is required for the biosynthesis of all the 2-hydroxyphenylthiazoline metabolic products of the *sven0503-sven0517* gene cluster and probably functions as a thiazoline reductase.

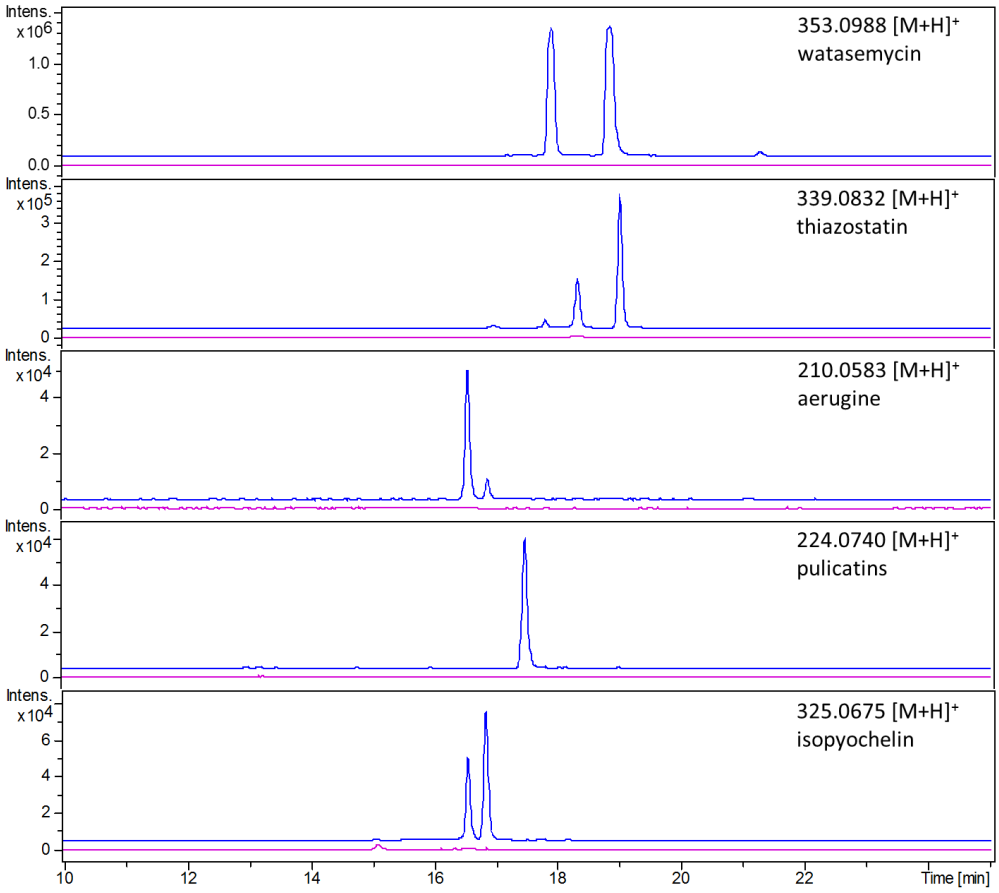
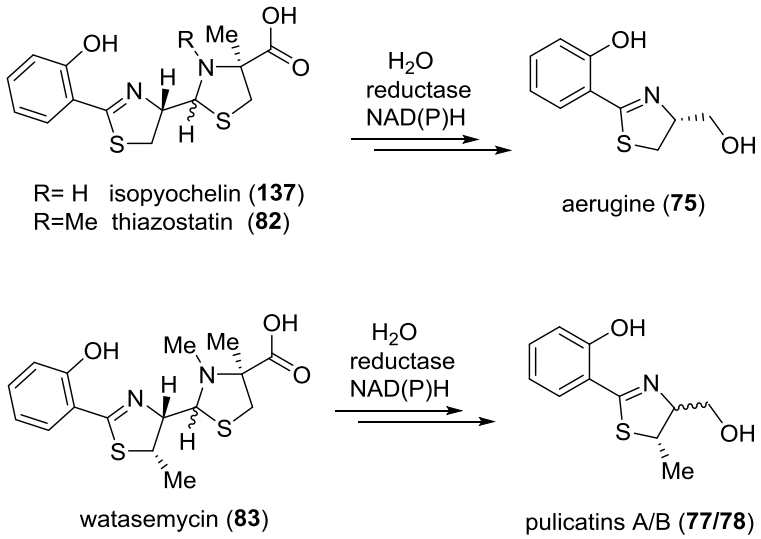


Figure 4.10. EICs at *m/z* = 353.0988, 339.0832, 210.0583, 224.0740 and 325.0675, corresponding to $[M+H]^+$ for watasemycin, thiazostatin, aerugine, pulicatins A/B and isopyochelin, respectively, from LC-HRMS analyses of the ethanol extract of the Δ*sven0516* mutant culture supernatant (purple traces), in comparison to the *S. coelicolor* M1152/SV-2_E03::*Ssp*I strain (blue traces).

Interestingly, deletion of *sven0516* also abolishes the production of aerugine **75** and pulicatins A/B **77**/**78**, the assembly of which ostensibly does not require a thiazoline reductase. Thus we suspect that aerugine and pulicatins A/B are derived from hydrolytic

cleavage and subsequent reduction of thiazostatin/isopyochelin and watasemycin, respectively, rather than being direct products of the pathway (Scheme 4.8).



Scheme 4.8. Proposed formation of aerugine **75** and pulicatins A/B **77**/**78**, from hydrolysis and reduction of isopyochelin/thiazostatin and watasemycin, respectively.

4.3.2 Sven0508 is not involved in biosynthesis of 2-hydroxyphenylthiazolines

The ethanol extract of the culture broth from the Δ*sven0508* mutant was also analysed by LC-HRMS. Watasemycin **83**, thiazostatin **82**, aerugine **75**, pulicatins A/B **77**/**78** and isopyochelin were all still observed in the extract from the Δ*sven0508* mutant, at comparable levels to the wild type strain (Figure 4.11). This implied that the PchK homologue, Sven0508, is not involved in the biosynthesis of the metabolic products of the *sven0503-sven0517* gene cluster.

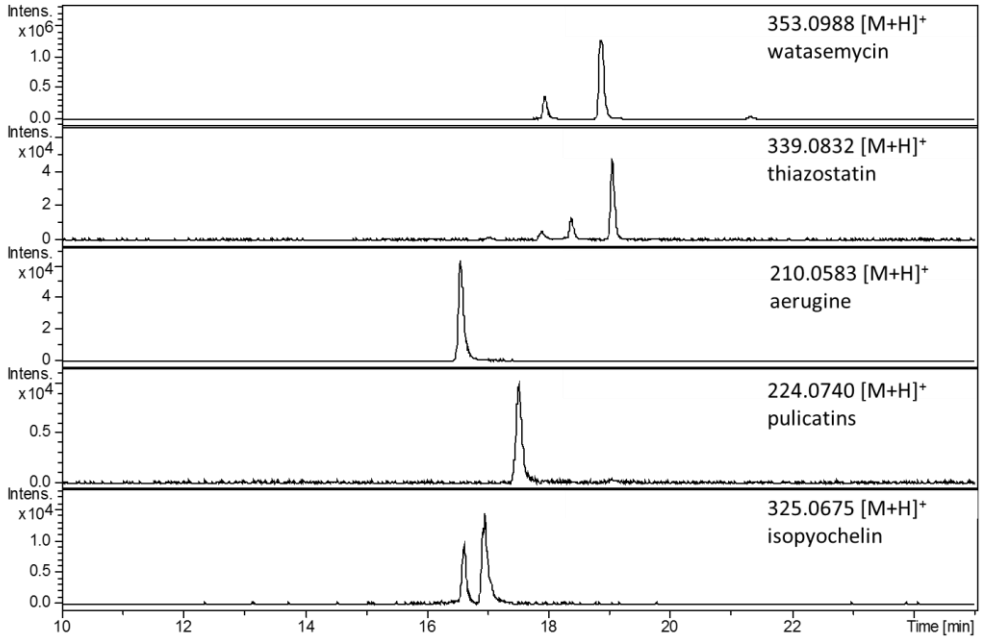


Figure 4.11. EICs at *m/z* = 353.0988, 339.0832, 210.0583, 224.0740 and 325.0675, corresponding to $[M+H]^+$ for watasemycin, thiazostatin, aerugine, pulicatins A/B and isopyochelin, respectively, from LC-HRMS analyses of the ethanol extract of the Δ*sven0508* mutant culture supernatant.

4.4 Reassignment of the absolute stereochemistry of watasemycin and thiazostatin

Watasemycin is structurally closely related to pyochelin and thiazostatin. The relative stereochemistry of thiazostatin and watasemycin was assigned in the literature as shown in Figure 4.12 on the basis of NOE NMR studies.[117] As mentioned previously, PchK is proposed to catalyse reduction of the second thiazoline, resulting in production of enantio-pyochelin **76** which has *S* configuration at both C-4' and C-4". Herein, we have established that the absolute configuration at C-4" in isopyochelin is *S*. Since Sven0516 appears to function as the thiazoline reductase in the biosynthesis of all the 2-hydroxyphenylthiazolines from the *sven0503-sven0517* gene cluster, including thiazostatin and watasemycin, it is reasonable to speculate that they all have same *S* absolute configuration at C-4". Combined with the previously assigned relative stereochemistry for thiazostatin and watasemycin, the absolute configuration can be deduced as 4'*R*, 5'*R*, 4"*S* for watasemycin and 4'*R*, 4"*S* for thiazostatin, as shown in the structure circled in red in Figure 4.12.

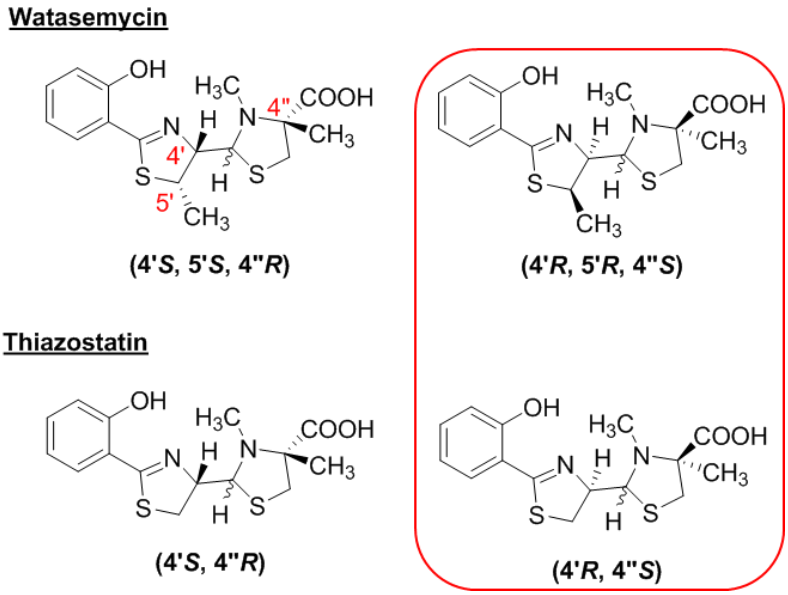


Figure 4.12. Two reported possible absolute configurations of watasemycin and thiazostatin. The absolute stereochemistry with 4"*S* configuration is circled in red.

Besides thiazostatin and watasemycin, aerugine **75** and pulicatins A/B **77/78** are also proposed to be the products of the *sven0503-sven0517* gene cluster based on high resolution LC-MS analyses. Their structures have been fully elucidated as discussed in Section 1.5.3 (Figure 1.17). It was noticed that pulicatins A/B **77/78** have a methyl group at C-5', which is the same as in watasemycin, and the absolute configuration at that stereocenter has been confirmed to be *S*.[115] This is contradictory to the assignment of the absolute stereochemistry of watasemycin based on the relative stereochemistry reported by Sasaki *et al*.[117] Thus we suspect that their assignment of the relative stereochemistry may be incorrect for watasemycin and thiazostatin. Moreover, NOE studies of watasemycin showed that the methyl group at C-5' is proposed to be *anti* to the hydrogen atom at C-4', suggesting their absolute configuration may be either 4'*R*, 5'*R* or 4'*S*, 5'*S*. Taken together, the absolute stereochemistry can be reassigned as 4'*S*, 5'*S*, 4"*S* and 4'*S*, 4"*S* for watasemycin and thiazostatin, respectively, as shown in Figure 4.13.

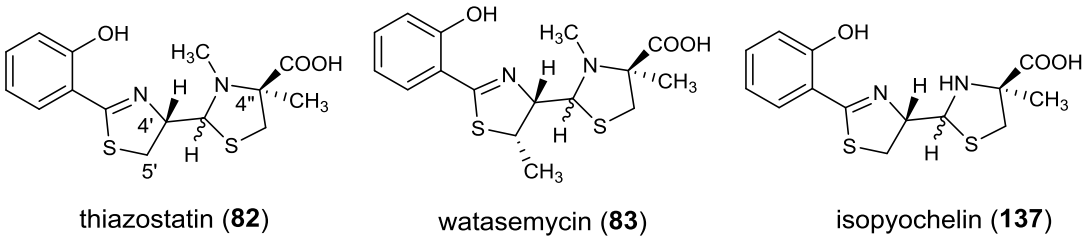


Figure 4.13. Reassigned stereochemistry of thiazostatin, watasemycin and isopyochelin.

An MT-like domain in PchE (or homologue) is believed to catalyse the epimerisation of the stereocenter in the first thiazoline during pyochelin biosynthesis in *P. aeruginosa* and *S. scabies*, resulting in the *R* configuration at C-4'.[114,121] As discussed in Section 4.1.1, bioinformatics analysis showed that this region in Sven0512 seems to be non-functional. Lack of the MT-like domain in the PchE homologue in *P. protegens* leads to the production of enantio-pyochelin **76** with 4'*S* configuration.[114] The seemingly inactive MT-like epimerisation domain in Sven0512 supports our assignment of a 4'*S* configuration in watasemycin and thiazostatin. However, more work is needed to fully elucidate their absolute stereochemistry.

4.5 Role of Sven0515 in watasemycin biosynthesis

4.5.1 Sven0515 methylates C-5' of thiazostatin to produce watasemycin

Bioinformatics analyses showed that the product of the *sven0515* gene is homologous to class B RS methylases, which are known to catalyse methylation on unreactive carbon centres in the biosynthesis of a wide variety of specialised metabolites. Compared to pyochelin, thiazostatin **82** contains one additional methyl group at C-4" and besides this, watasemycin **83** carrys another methyl group at C-5' (Figure 4.13). Based on its similarity to class B RS methylases, Sven0515 is predicted to be responsible for introducing one or both of these. In order to determine the specific role of Sven0515 in the biosynthesis of 2-hydroxyphenylthiazolines, a deletion mutant of *sven0515* was also constructed by Dr. Yuki Inahashi.

LC-MS analysis of the ethanol extract from the culture supernatant of the Δ*sven0515* mutant strain showed that it is able to produce aerugine **75**, thiazostatin **82** and isopyochelin **137**, but not watasemycin **83** or pulicatins A/B **77**/**78**, which all contain a methyl group at C-5' (Figure 4.14). Thus it seems that Sven0515 is required for the methylation of C-5', but not C-4", due to the production of thiazostatin and isopyochelin, both of which contain the methyl group at C-4".

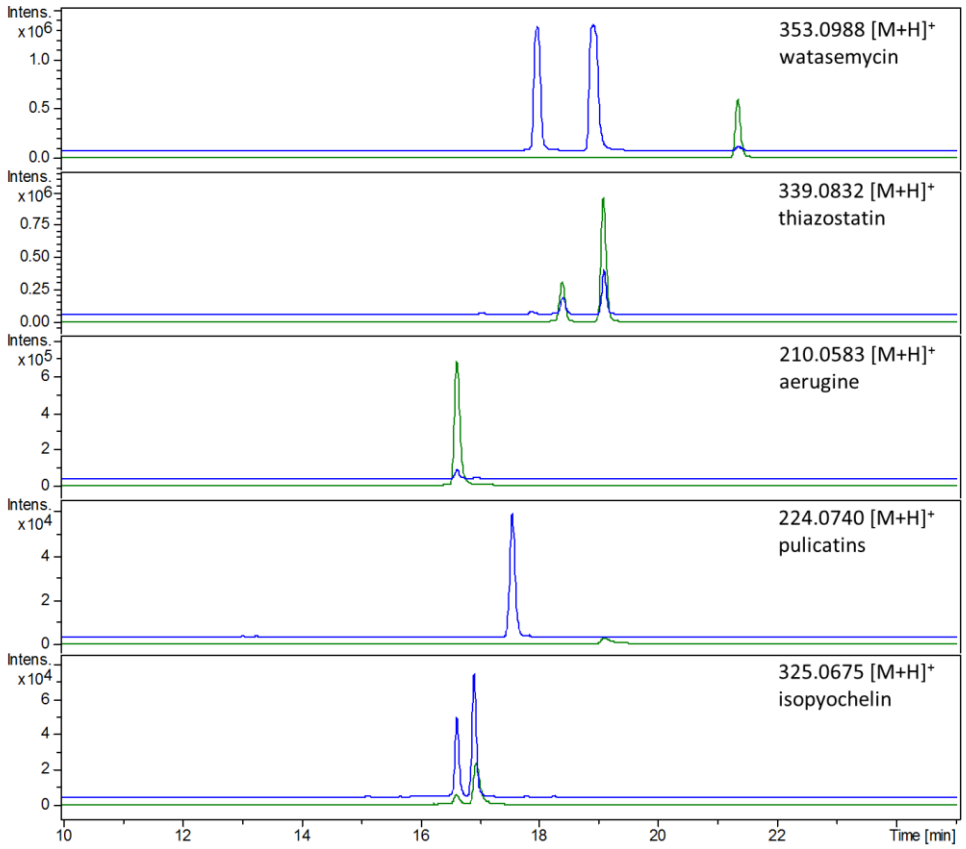


Figure 4.14. EICs at *m/z* = 353.0988, 339.0832, 210.0583, 224.0740 and 325.0675, corresponding to $[M+H]^+$ for watasemycin, thiazostatin, aerugine, pulicatins A/B and isopyochelin, respectively, from LC-MS analyses of the ethanol extract of the Δ*sven0515* culture supernatant (green traces), in comparison to the *S. coelicolor* M1152/SV-2_E03::*Ssp*I strain (blue traces).

To investigate the timing of C-5' methylation, we fed the metabolites in the extract of the Δ*sven0515* mutant to the Δ*sven0516* mutant. The Δ*sven0516* mutant was confirmed not to produce any of the 2-hydroxyphenylthiazolines due to absence of the thiazoline reductase as described in Section 4.3.2. Incubation of the culture for four days after feeding with the extract of the Δ*sven0515* mutant resulted in nearly complete conversion of thiazostatin **82** in the extract to watasemycin **83** (Figure 4.15), indicating that thiazostatin is the substrate of Sven0515. It also suggests that Sven0515-catalysed methylation of thiazostatin is the final step in watasemycin biosynthesis.

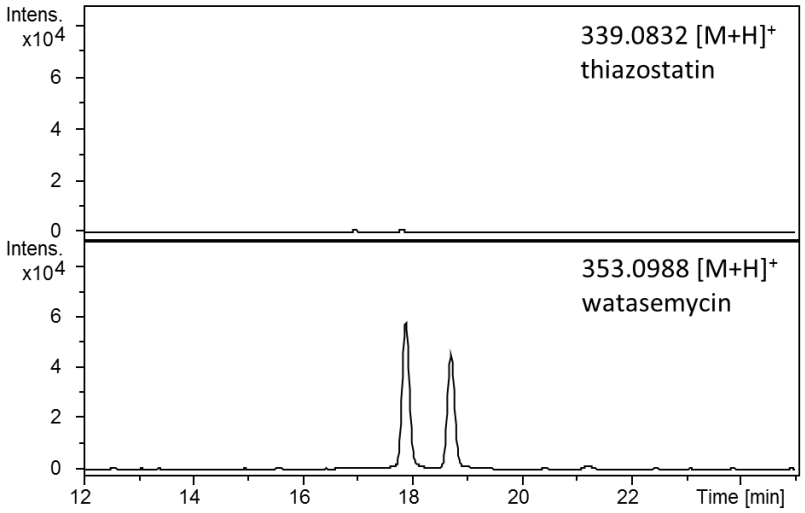
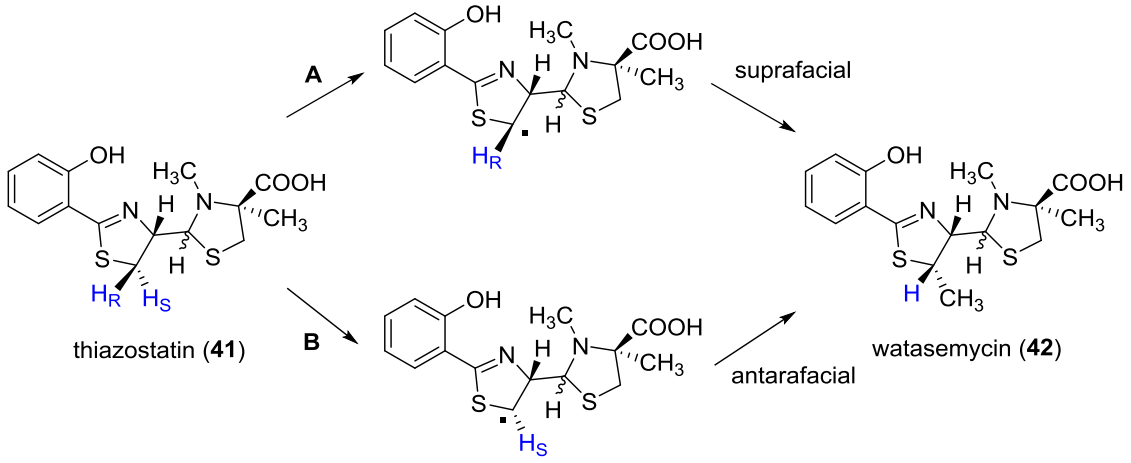


Figure 4.15. EICs from LC-MS analyses of the ethanol extract from the Δ*sven0516* mutant fed with the extract from the Δ*sven0515* mutant.

4.5.2 Synthesis of stereospecifically deuterium-labelled cysteine

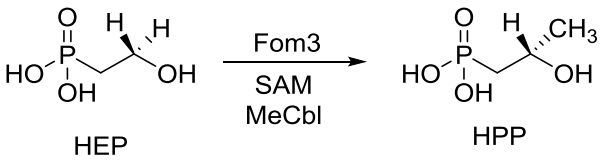
As discussed in Section 1.5.2, the biosynthesis of pyochelin and enantio-pyochelin involves condensation of the salicylate with L-cysteine to afford the first thiazoline ring. The biosynthesis of thiazostatin and watasemycin is predicted to be similar for this step. The C-5' on the thiazoline ring of thiazostatin would then be methylated by Sven0515 as the final step in watasemycin biosynthesis.

Based on the determination of 5'*S* configuration in watasemycin, there are two possible stereochemical courses for this reaction (Scheme 4.9). Firstly, Sven0515 could stereospecifically abstract the *pro-S* hydrogen atom from C-5', followed by appending of a methyl group from MeCbl to the resulting carbon in a suprafacial process with retention of stereochemistry, route A in Scheme 4.9. Alternatively, Sven0515 could stereospecifically abstract the *pro-R* hydrogen atom from C-5', followed by appending a methyl group from MeCbl to the resulting carbon in an antarafacial process with inversion of stereochemistry, route B in Scheme 4.9. A third possibility is non-stereospecific abstraction of either *pro-S* or *pro-R* hydrogen atom from C-5' resulting in methylation *via* a mixture of routes A and B in Scheme 4.9. Base on reported studies about other homologues of class B RS methylases, the methylation is predicted to proceed with inversion of stereochemistry at the reacting carbon.



Scheme 4.9. Possible stereochemical courses for Sven0515-catalysed methylation, which could occur *via* primarily route A or route B, or *via* a mixture of routes A and B.

Fom3, one of the first recognised class B RS methylases, has been characterised *in vitro* to catalyse the stereoselective methylation of 2-hydroxyethylphosphonate (HEP) to afford *S*-2-hydroxypropylphosphonate (HPP),[185-187] the penultimate step in the biosynthesis of the broad spectrum antibiotic fosfomycin (Scheme 4.10).



Scheme 4.10. The stereoselective methylation catalysed by Fom3 in fosfomycin biosynthesis.

Prior to its successful characterisation *in vitro*, in order to investigate the stereochemical course during fosfomycin **138** biosynthesis, Hammerschmidt and co-workers synthesised HEPs deuterium labelled at the C-2 position, [2,2-$^{2}H_2$]-HEP, (*R*)-[2-$^{2}H_1$]-HEP and (*S*)-[2-$^{2}H_1$]-HEP, and fed them to the fosfomycin producing strain *Streptomyces fradiae* (Figure 4.16).[188-189] Deuterium incorporation into fosfomycin was measured based by NMR spectroscopic analysis of (1*R*,2*R*)-(2-amino-l-hydroxypropy1)phosphonic acid, which was produced from fosfomycin by reaction with ammonia. When the two deuterium labelled compound [2,2-$^{2}H_2$]-HEP, or *pro-S* deuterium-labelled compound (*S*)-[2-$^{2}H_1$]-HEP, were fed to *S. fradiae*, 34% and 32% of the fosfomycin produced contained label, respectively. However, feeding with the *pro-R* deuterium-labelled compound, (*R*)-[2-$^{2}H_1$]-HEP, resulted in no retention of the label. Thus, a mechanism was proposed by van der Donk *et al.* that the *pro-R* hydrogen atom

was removed from C-2 of HEP during the methylation catalysed by Fom3.[187] This should be followed by an antarafacial process of the resulting HEP radical reacting with MeCbl to achieve the methyl group with *S*-configuration at the reacting carbon in HPP.

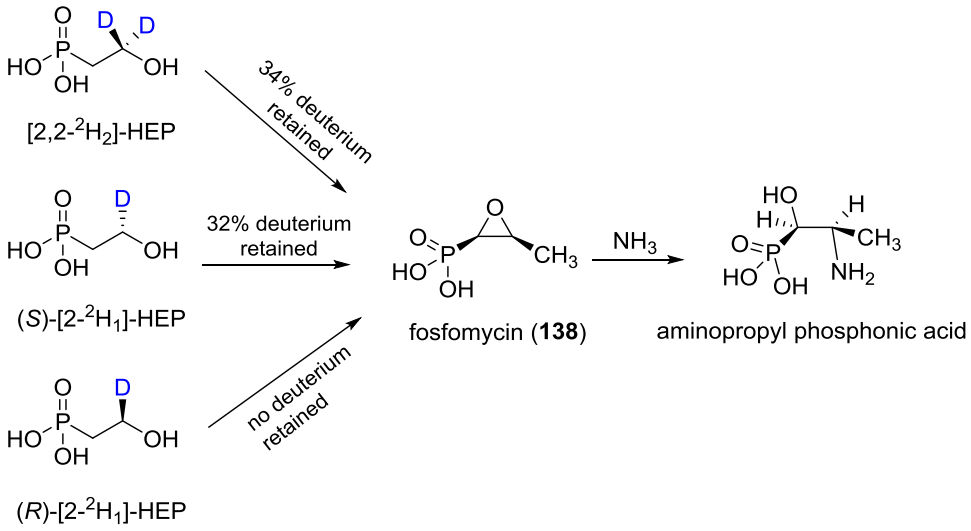


Figure 4.16. Feeding of stereospecifically deuterium-labelled substrates to determine the stereochemical course of the C-methylation reaction in fosfomycin biosynthesis.

A second example of a methylation that occurs with stereochemical inversion is biosynthesis of macrocyclic peptide antibiotic bottromycin, which harbors several non-proteinogenic amino acids. The enzymes responsible for these reactions also belong to the class B family of radical SAM methylases. In order to determine the mechanism of the methylations, Kellenberger *et al.* carried out feeding experiments in *Streptomyces bottropensis* with chemically synthesised isotope-labelled precursors.[190] These precursors included isotopomers of the corresponding amino acids, valine and phenylalanine (Figure 4.17). The valine isotopomer carrys one methyl group labelled by ^{13}C and the other by deuterium, while both of the phenylalanine isotopomers have a stereo-defined deuterium label at the β position. The incorporation results show that the methylations occur with inversion of configuration at the centres undergoing functionalisation.

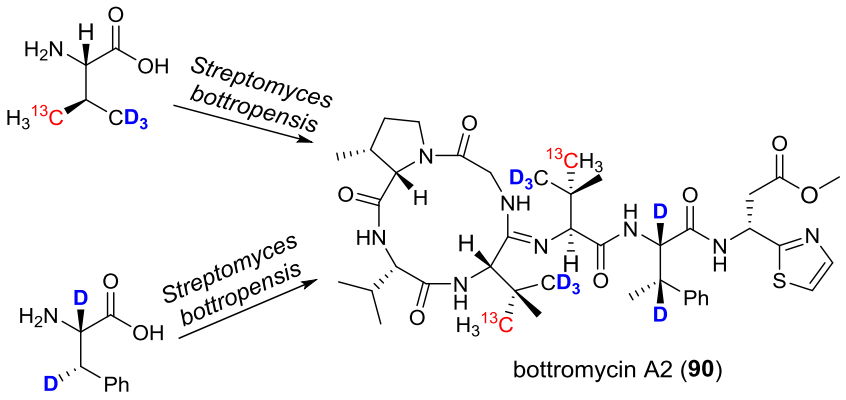
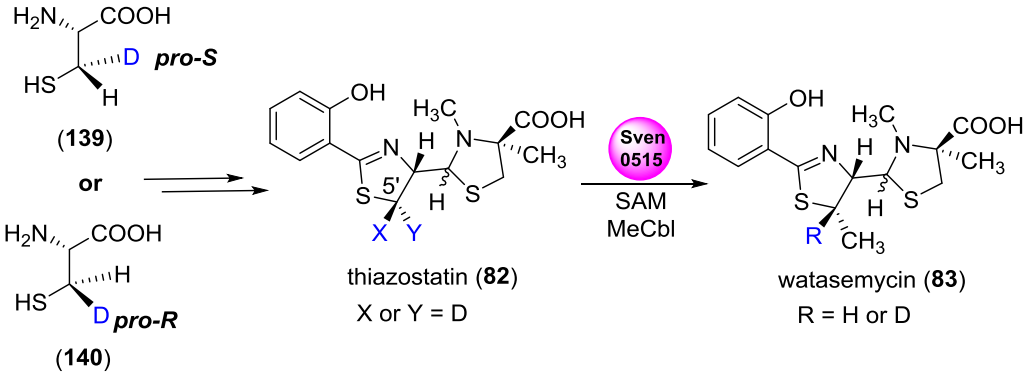


Figure 4.17. Feeding of stereospecifically isotope-labelled amino acids to determine the stereochemical course of the methylations in bottromycin biosynthesis, reported by Kellenberger and Arigoni.[190]

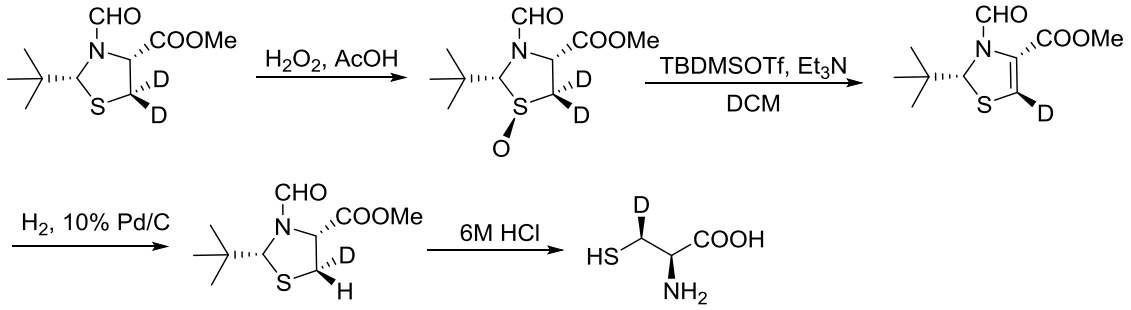
The use of stereospecifically isotope-labelled substrates is an effective method for elucidating the stereochemical course of enzymatic reactions. It was envisaged that a similar approach could be employed to investigate the stereochemical course of the methylation step catalysed by Sven0515 (Scheme 4.11).



Scheme 4.11. Proposed approach for mechanistic studies of Sven0515-mediated methylation using stereospecifically deuterium-labelled L-cysteine.

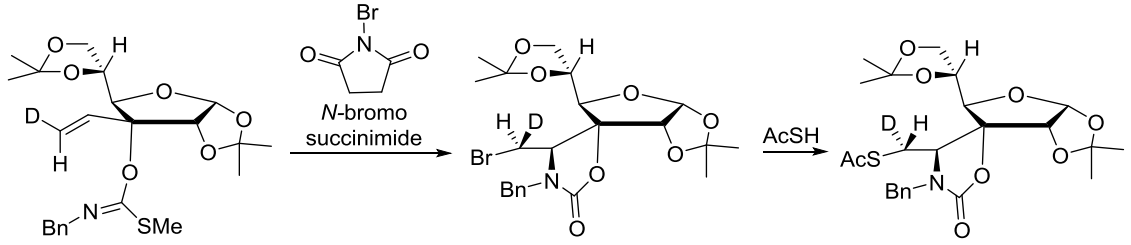
Similar to pyochelin and enantio-pyochelin, two molecules of L-cysteine should be incorporated into the structure of thiazostatin and watasemycin during biosynthesis. Thus, feeding stereospecifically deuterium-labelled L-cysteine at the β-position to the producing strain would lead to thiazostatin or watasemycin with a chiral label at the target methylating position, C-5'and provide evidence to distinguish between the possible scenarios in Scheme 4.9.

Therefore, a synthetic method for the stereoselective incorporation of a deuterium atom into the β-position of L-cysteine was required. Since cysteine is easily oxidised, it is normally prepared as the homodimer form L-cystine, which is biochemically convertible into its monomer L-cysteine. Examination of the literature revealed several methods for the synthesis of cystine or cysteine stereospecifically deuterated or tritiated at the β-position. However, most of the reported methods involve enzymatic and/or resolution processes in order to obtain enantiomerically pure products.[191-198] There are few examples of purely chemical methods for introducing an isotope label in a stereospecific manner.[199-200] One reported method is reductive hydrogenation of the sterically hindered and deuterated thiazoline with palladium followed by hydrolysis to achieve the desired stereospecifically deuterium-labelled cysteine (Scheme 4.12).[199] The modest total yield of this synthetic approach, involving preparation of the thiazoline from corresponding deuterium-labelled cysteine, made it too expensive and time-consuming for our purposes.



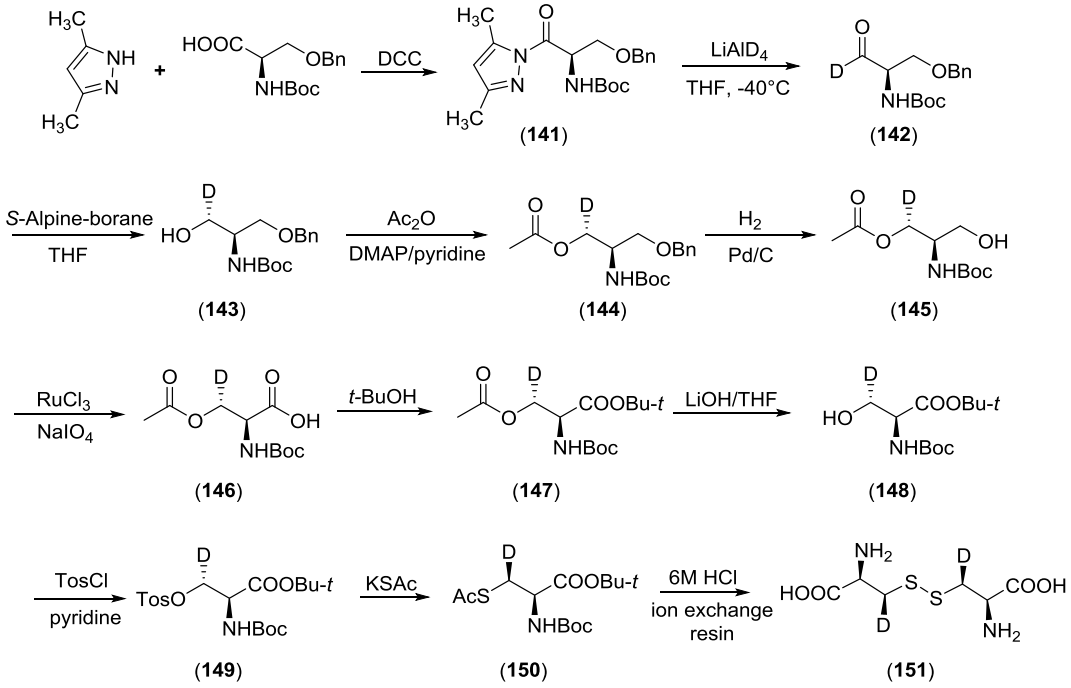
Scheme 4.12. Previous synthesis of stereospecifically deuterium-labelled L-cysteine at the β-position through a deuterated thiazoline.[199]

Another approach was proposed by Maeda *et al.* to use a carbohydrate chirality template with double steric constraint to achieve the chirally β-deuterated L-amino acids (Scheme 4.13).[200]



Scheme 4.13. Proposed method for synthesis of stereospecifically deuterium-labelled L-cysteine at the β-position using a carbohydrate chirality template by Maeda *et al.*[200]

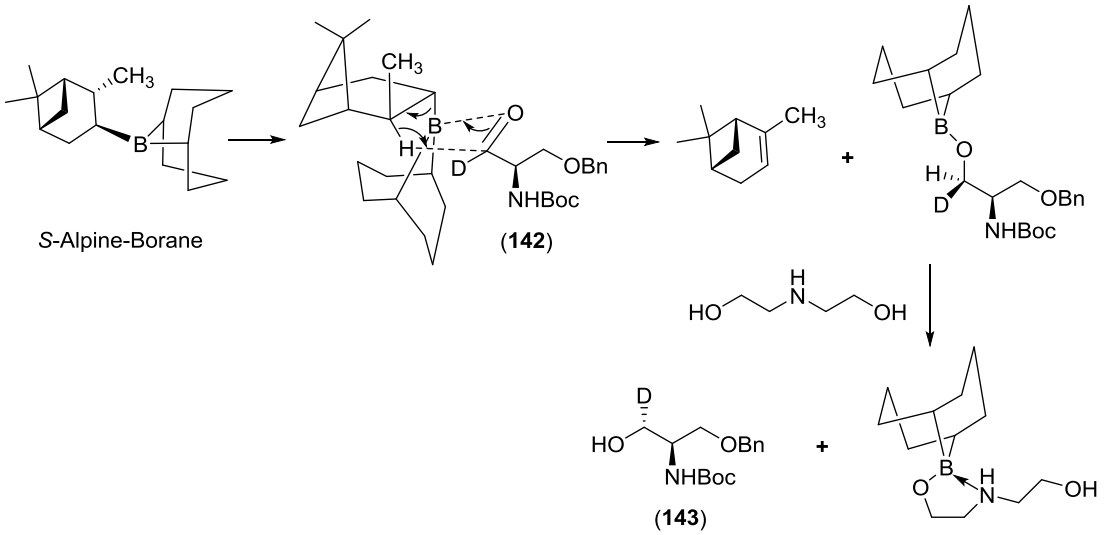
More recently, Oba *et al.* developed another method for introducing a stereospecific deuterium label at the β-position of L-cystine as outlined in Scheme 4.14.[201] The stereospecific deuterium label was obtained by transforming the carboxyl group of D-serine into the stereospecifically deuterium-labelled alcohol *via* asymmetric reduction of 1-deuterio aldehyde. Most steps of this route were reported in excellent yields and the reagents used can be easily handled. Due to the *S*-Alpine-Borane used in the reduction step of the route, the stereochemistry of the deuterium label at the β-position was *S*-configured and the diastereoselectivity of deuterium substitution was reported to be approximately 95% de from integration of the ^{1}H NMR spectrum. Based on these advantages, we adopted this approach for our synthesis of the stereospecifically deuterium-labelled L-cysteine (*S*-3-^{2}H-L-cysteine **139)** (Scheme 4.14).



Scheme 4.14. The synthetic route for (2*R*,2'*R*,3*S*,3'*S*)-[3,3'-^{2}H$_2$]cystine reported by Oba *et al.*[201]

The 11-step synthesis begins with the preparation of [1-^{2}H]serinal derivative **142**. O-Bn and N-Boc-protected D-serine, which is commercially available, was condensed with 3,5-dimethylpyrazole and the pyrazolide **141** obtained was reduced with lithium aluminium deuteride to give the required 1-deuterio aldehyde **142**. To avoid racemisation of this *N*-protected α-amino aldehyde,[202] it was immediately used in the next step without purification. This step requires careful quenching and control of the temperature to avoid over-reduction. Asymmetric reduction of 1-deuterio aldehyde **142** using *S*-Alpine-

Borane yielded stereospecifically deuterium-labelled alcohol **143** in more than 90% yield. The mechanism of the reduction is described in Scheme 4.15. Diethanolamine was added at the end of the reaction to precipitate the dialkyl borane byproduct.



Scheme 4.15. The mechanism of the stereospecific reduction of aldehyde **142** by *S*-Alpine-Borane.

Subsequently, acetylation and debenzylation of the deuterated alcohol **143** followed by RuO_4-oxidation of the deprotected hydroxymethyl group of compound **145** afforded protected L-serine derivative **146** in quantitative yield. The ^{1}H NMR spectrum of **146**, as well as other doubly protected intermediates, had unusual signals and splitting patterns, most notably for the protons at α- and β-position of the amino acid. This indicated the presence of rotamers that are interchanging conformation on the timescale of the NMR experiment. To investigate this hypothesis, the ^{1}H NMR experiments for protected L-serine derivative **146** were repeated at varying temperatures from 5 ℃ to 55 ℃ as shown in Figure 4.18. As the temperature increased, conversion between the rotamers became faster and the signal splitting was significantly reduced, confirming the presence of rotamers.

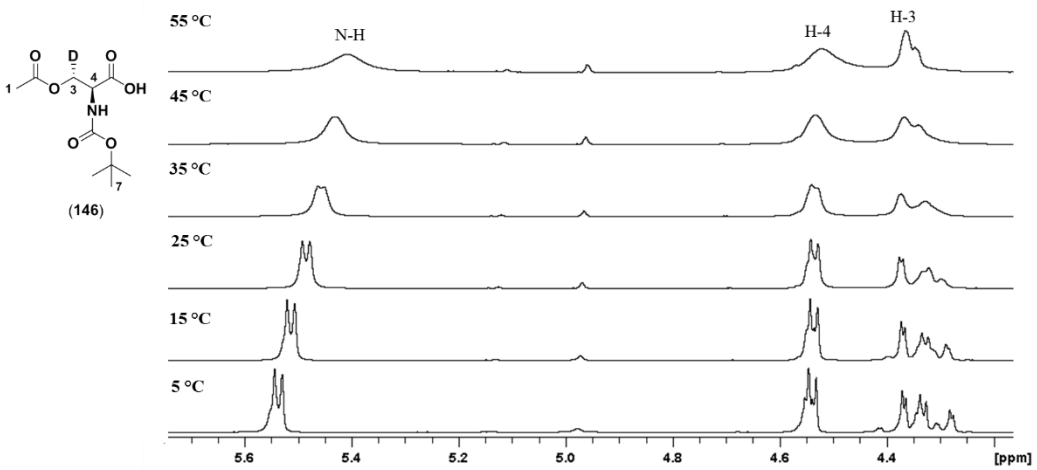


Figure 4.18. A comparison between 4.20-5.70 ppm of the ^{1}H NMR spectra in $CDCl_3$ at 600 MHz of (3*R*)-*O*-acetyl-*N*-(*tert*-butoxycarbonyl)-L-serine-3-^{2}H **146** at varying temperatures.

After protection of the carboxyl functionality as its *tert*-butyl ester, the deuterated L-serine derivative **147** was transformed into the corresponding tosylate **149** in 55% yield *via* deacetylation and tosylation. The tosylate **149** obtained was then treated with potassium thioacetate to give deuterium-labelled L-cysteine derivative **150** in an S_N2-type displacement. Deprotection of compound **150** was carried out in refluxing 6 M HCl solution and the resulting L-cysteine hydrochloride was purified by ion-exchange column chromatography (Dowex® 50WX8 hydrogen form, 200-400 mesh). The eluent containing L-cysteine was aerated in air for two days to give the (2*R*,2'*R*,3*S*,3'*S*)-[3,3'-$^{2}H_2$]cystine **151** as a white solid.

The synthetic route proved to be reproducible and gave more than 20% total yield in our hands. The ^{1}H and ^{13}C NMR spectra (Figure 4.19) of the product were consistent with that reported by Oba *et al.*[201]

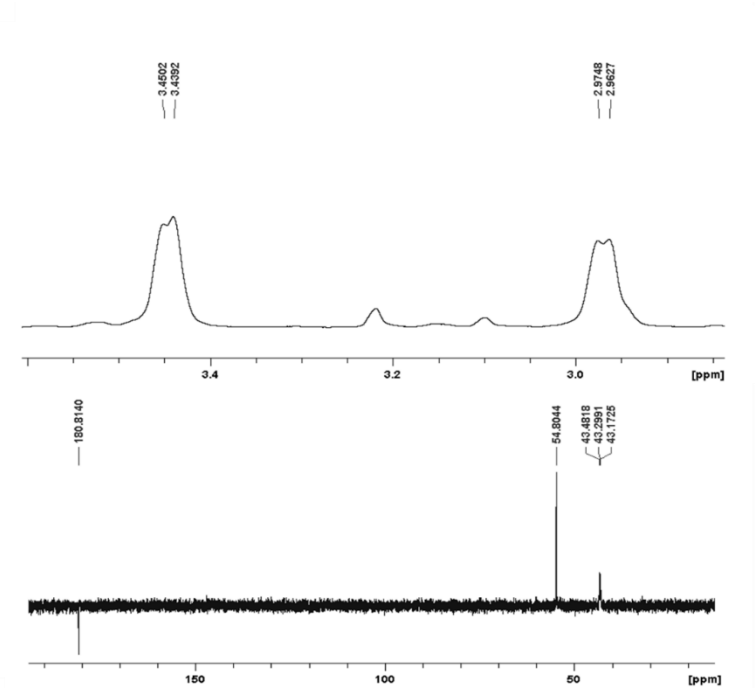


Figure 4.19. The ^{1}H (at 300 MHz) and ^{13}C NMR (at 500 MHz) spectra of the synthesised (2*R*,2'*R*,3*S*,3'*S*)-[3,3'-$^{2}H_2$]cystine **151** in 2.5% NaOD in D_2O.

4.5.3 Incorporation of stereospecifically deuterium-labelled L-cystine into watasemycin

With the successful synthesis of the stereospecifically deuterium-labelled L-cystine, the incorporation approach to probe the stereochemical course of Sven0515-catalysed methylation was investigated. *S. coelicolor* M1152/SV-2_E03*::Ssp*I was initially grown in YD medium for two days to ensure good growth, then pelleted, transferred to supplemented minimal medium (SMM), since this medium gives more equal production levels of thiazostatin and watasemycin (Figure 4.20). SMM is a simple production medium for secondary metabolites, with glucose as the primary carbon source.

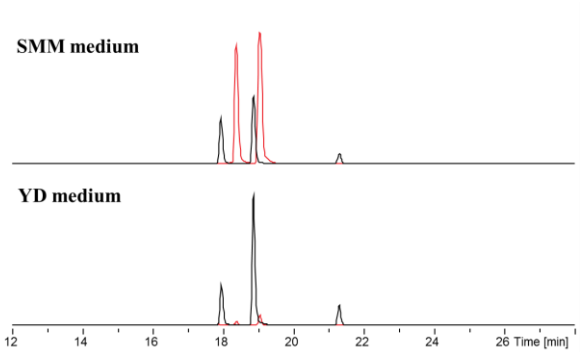


Figure 4.20. A comparison of EICs for thiazostatin (red traces) and watasemycin (black traces) from the LC-MS analyses of the ethanol extracts of *S. coelicolor* M1152/SV-2_E03*::Ssp*I culture supernatant grown in SMM and YD medium.

After incubation of the culture for 24 hours, labelled L-cystine was added. The feeding procedure was repeated at 24 hour intervals until a total incubation time of 5 days at 30 ℃. The synthesised labelled L-cystine fed into the culture could be easily converted by bacteria to L-cysteine with the *pro-S* hydrogen atom replaced by a deuterium label at the β-position.

Following extraction with ethanol, LC-MS analysis was carried out. Due to the incorporation of two molecules of L-cysteine into thiazostatin and watasemycin, the deuterium label may not only appear at the methylating position, C-5', but also C-5''. Therefore, feeding of deuterium-labelled cysteine should lead to production of thiazostatin and its isotopomers harboring one deuterium label at C-5' or C-5'', or two deuterium labels at both of the positions (left column, Figure 4.21).

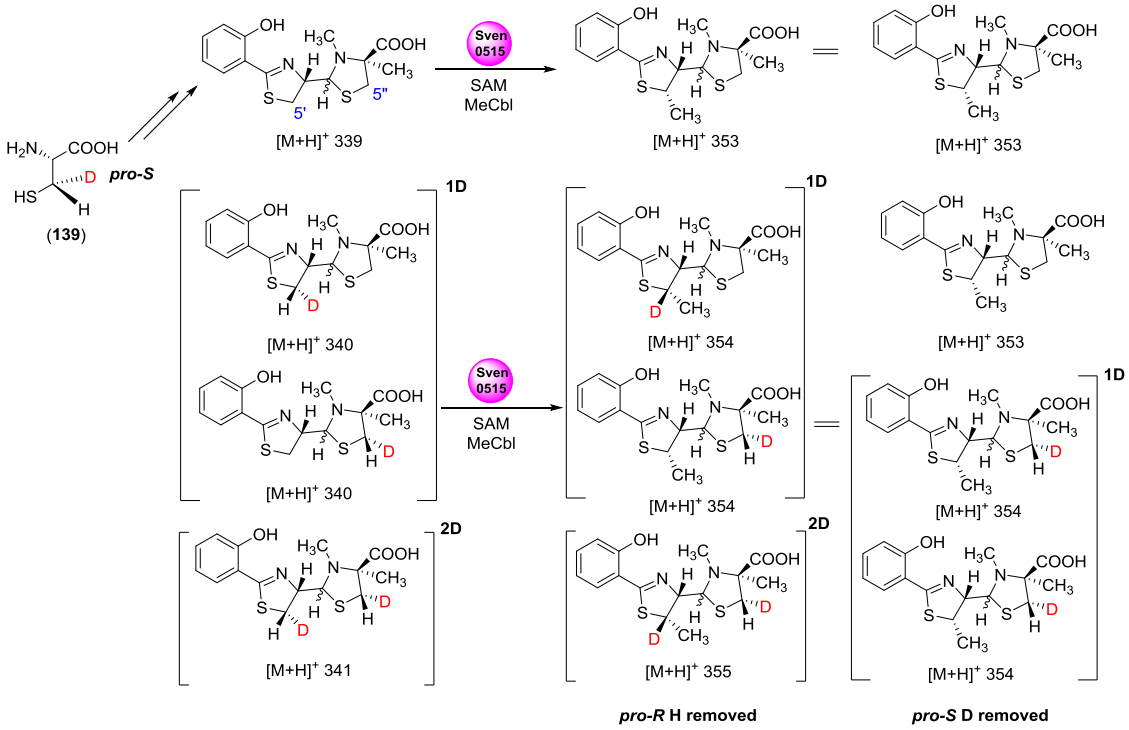


Figure 4.21. Possible deuterium-labelled isotopomers of thiazostatin and watasemycin produced after feeding with (2*R*, 3*S*)-[3-$^{2}H_{1}$] cysteine **139**.

If Sven0515 stereospecifically abstracts the *pro-R* hydrogen atom from C-5', then feeding of *pro-S* deuterium labelled L-cysteine should result in labelled watasemycin and its isotopomers with retention of the *pro-S* deuterium label. As with thiazostatin, the watasemycin isotopomers could include a deuterium label(s) at C-5' and/or C-5", as shown in the middle column of Figure 4.21. The ratio of the watasemycin isotopomer labelled with two deuterium to that labelled with one deuterium atoms should correspond to the ratio of the thiazostatin isotopomers.

Alternatively, if Sven0515 stereospecifically abstracts the *pro-S* hydrogen atom from C-5' of thiazostatin, then feeding of *pro-S* deuterium labelled cysteine should result in loss of the deuterium at C-5' from watasemycin. In this case, there would not be a watasemycin isotopomer labelled with two deuterium atoms but only isotopomers labelled with one deuterium at C-5" (right column, Figure 4.21).

Finally, a non-stereospecific process would also result in production of isotopomer of watasemycin with two deuterium labels. However, compared to the stereospecific process, the ratio of the watasemycin isotopomer with two deuterium labels to the isotopomer with one deuterium label should be lower than for thiazostatin, due to a partial loss of the *pro-S* deuterium at C-5'.

A comparison of the HRMS of thiazostatin and watasemycin resulting from LC-MS analysis of the ethanol extracts of *S. coelicolor* M1152/SV-2_E03*::Ssp*I fed with (2*R*,2'*R*,3*S*,3'*S*)-[3,3'-$^{2}H_2$]cystine **151** is shown in Figure 4.22. It shows that both thiazostatin and watasemycin were produced as isotopomers harboring two deuterium labels, which correspond to the peak at *m/z* = 341.0943 and *m/z* = 355.1099 for thiazostatin and watasemycin, respectively.

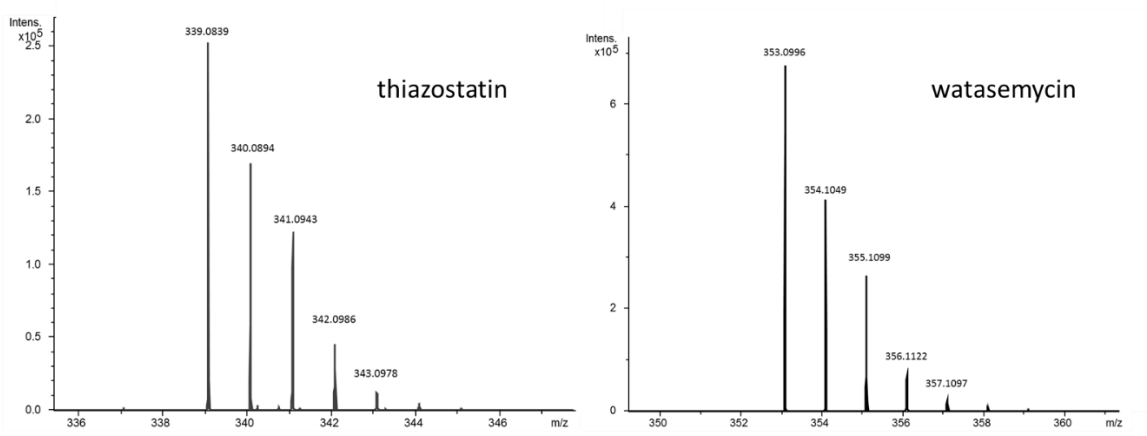


Figure 4.22. A comparison of the HRMS for thiazostatin and watasemycin from LC-MS analyses of the ethanol extracts of *S. coelicolor* M1152/SV-2_E03*::Ssp*I fed with (2*R*,2'*R*,3*S*,3'*S*)-[3,3'-$^{2}H_2$]cystine.

In order to determine the relative amount of each isotopomer of thiazostatin and watasemycin more accurately, a detailed analysis of the mass spectra is required. Taking thiazostatin as an example (Figure 4.23), the peak at *m/z* = 340.0894 is a combination of the natural isotopomer(s) of unlabelled thiazostatin and the singly deuterium-labelled thiazostatin from our feeding experiment with deuterium-labelled cysteine. Natural isotopomers of a compound exist due to the natural abundance of isotopes of the chemical elements composing the molecule with the ^{13}C-isotopomer mainly considered. The calculation is necessary due to the inability of the mass spectrometer to distinguish compounds containing a carbon-13 atom from those containing a deuterium atom.

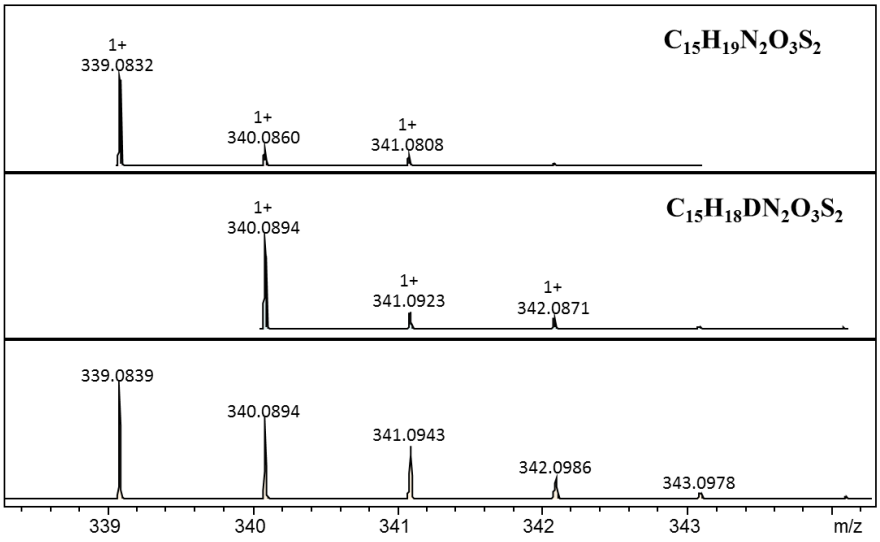


Figure 4.23. An example of deciphering the HRMS of the metabolic products from feeding experiments with deuterium-labelled precursors, including the isotopic distribution of natural thiazostatin (top) and singly deuterium-labelled thiazostatin (middle) and the actual mass spectrum of thiazostatin detected (bottom).

The isotope distribution in each molecule we analysed was obtained from the LC-MS data analysis software (Bruker Compass DataAnalysis). The relative amount of singly deuterium-labelled thiazostatin (*m/z* = 340.0894) is determined as 49.9% of unlabelled thiazostatin since 18.8% of the total peak at 340.0894 is due to the natural isotopomer(s) of unlabelled thiazostatin (Table 4.2).

Table 4.2. Calculation of the relative amounts of the deuterium-labelled thiazostatins from the HRMS.

m/z $[M+H]^+$	Relative amount (%) of each peak	+1 Due to natural isotopic abundance	+2 Due to natural isotopic abundance	Relative amount (%) of thiazostatins
339	100.0	-	-	**100** (0D)
340	68.7	**100** × 18.8%	-	**49.9** (1D)
341	47.6	**49.9** × 18.8%	**100** × 11.2%	**27.0** (2D)

For calculation of the relative amount of the doubly deuterium-labelled thiazostatin (*m/z* = 341.0957 for $[M+H]^+$ $C_{15}H_{17}D_2N_2O_3S_2$), it becomes more complicated due to contribution of the +2 isotopomer(s) of unlabelled thiazostatin and the +1 isotopomer(s) of singly deuterium-labelled thiazostatin. By subtracting these two contributions, the amount of doubly deuterium-labelled thiazostatin was determined as 27.0% of unlabelled

thiazostatin. Thus, the ratio of the doubly deuterium-labelled thiazostatin to singly deuterium-labelled thiazostatin can be calculated as 0.54:1.

We then determined the relative amounts of singly deuterium- and doubly deuterium-labelled watasemycins as 38.9% and 22.2%, respectively, compared to unlabelled watasemycin (Table 4.3). The ratio of the doubly deuterium-labelled watasemycin to singly deuterium-labelled watasemycin is 0.57:1.

Table 4.3. Calculation of the relative amounts of the deuterium-labelled watasemycins from the HRMS.

m/z $[M+H]^+$	Relative amount (%) of each peak	+1 due to natural isotopic abundance	+2 due to natural isotopic abundance	Relative amount (%) of watasemycins
353	100.0	-	-	**100** (0D)
354	58.9	**100** × 20.0%	-	**38.9** (1D)
355	39.3	**38.9** × 20.0%	**100** × 9.3%	**22.2** (2D)

The fact that the ratio of doubly deuterium-labelled watasemycin to singly deuterium-labelled watasemycin is very close to that of thiazostatin strongly suggests retention of the *pro-S* deuterium label during Sven0515-catalysed methylation, as previously outlined in Figure 4.21.

Although repeated feeding experiments produced varied levels of labelled thiazostatin and watasemycin, the ratio in watasemycin was always comparable and often higher than that in thiazostatin, further confirming the result. Thus it can be concluded that Sven0515 stereospecifically abstracts the *pro-R* hydrogen atom from C-5' and proceeds with inversion of configuration at that position to give 5'*S* configuration (Figure 4.24). Sven0515-catalysed methylation appears to share the same mechanism as other class B RS methylases, in which inversion of stereochemistry also occurs. The result also reciprocally supports our reassignment of the configuration at C-5' as *S*.

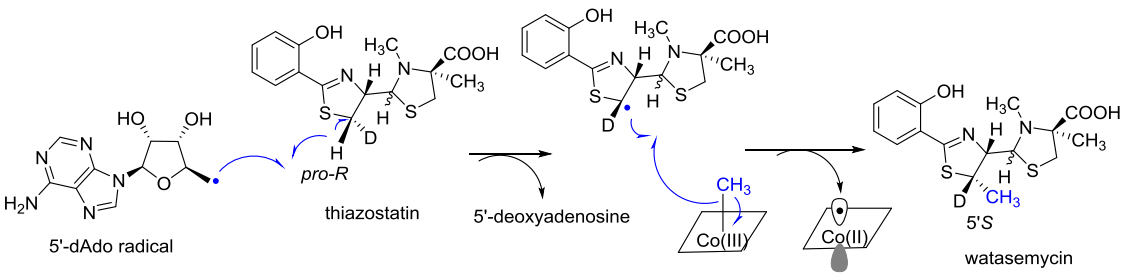


Figure 4.24. Proposed stereochemical course during Sven0515-catalysed methylation at C-5' of thiazostatin based on the feeding experiment.

In addition, it should be noted that the 3D peak (*m/z* 342 and 356 for $[M+H]^+$ of thiazostatin and watasemycin, respectively) in the HRMS of thiazostatin and watasemycin appears more intense than expected. Examination of the L-cysteine metabolic pathway revealed that the β-carbon is transferred to tetrahydrofolate to give N^5,N^{10}-methylenetetrahydrofolate in primary metabolism, which is later incorporated into homocysteine. Homocysteine is then used to regenerate methionine, which is a necessary component for SAM. Thus, we suspect that part of the 3D peak may be due to thiazostatin and watasemycin with two deuterium labels at C-5' and C-5" and another label at one of the methyl groups. Reducing the concentration of labelled cystine used in feeding or co-feeding of L-methionine (2-4 mM) at the same time proved to be effective for minimising the intensity of the 3D peak while retention of deuterium label at C-5' was still observed with similar results as above.

4.6 Proposed pathway for watasemycin biosynthesis

Based on above results, a biosynthetic pathway for 2-hydroxyphenylthiazolines from the *sven0503-sven0517* gene cluster is proposed (Figure 4.25).

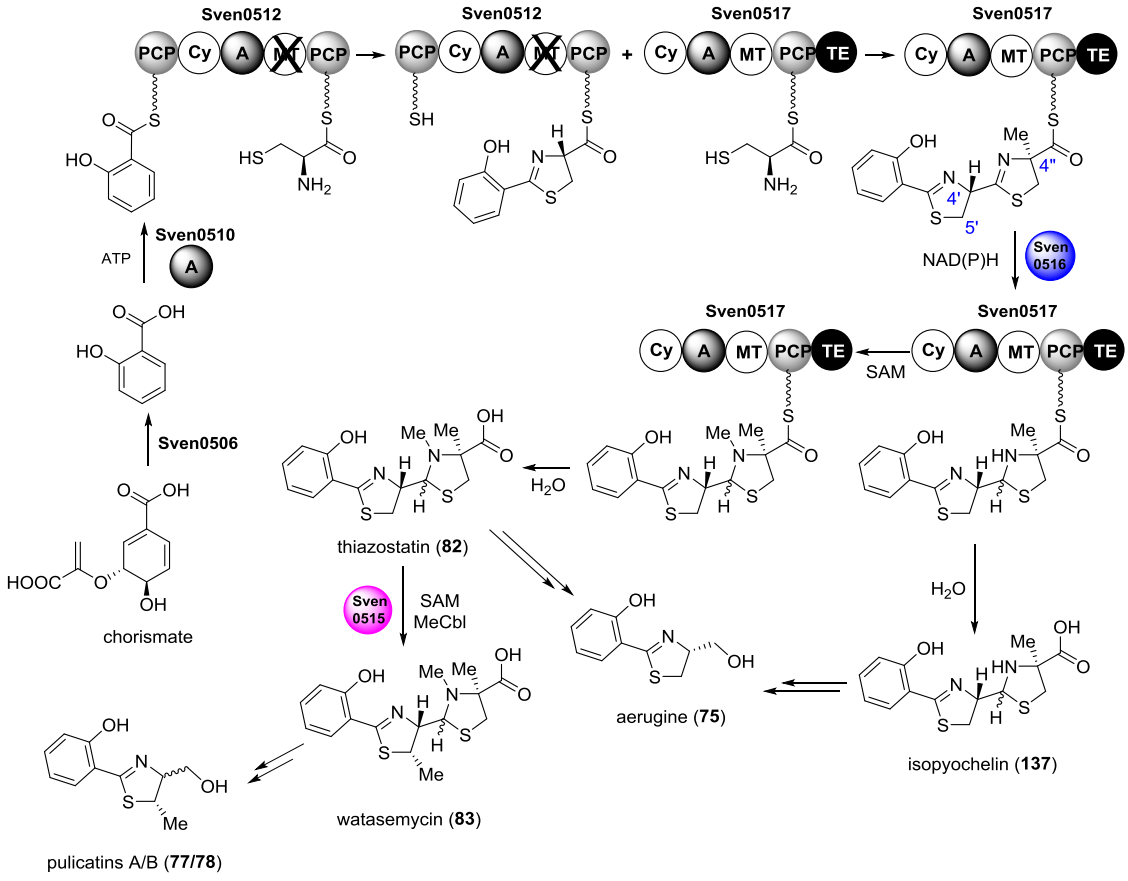


Figure 4.25. Proposed pathway for the biosynthesis of thiazostatin **82**, watasemycin **83**, isopyochelin **137**, aerugine **75** and pulicatins A/B **77**/**78** in *S. venezuelae* ATCC 10712.

Sven0506 catalyses formation of salicylate from chorismate. Salicylate is adenylated by the A domain, Sven0510, in the presence of ATP and loaded onto the *N*-terminal PCP domain of the NRPS Sven0512. The internal A domain of Sven0512 catalyses adenylation of the first molecule of L-cysteine, which is then covalently linked to its C-terminal PCP domain. The Cy domain in Sven0512 catalyses condensation of the thioesters of L-cysteine and salicylate and subsequent cyclisation to give the first thiazoline ring. During this process, the stereochemistry of L-cysteine is retained due to lack of a functional MT domain to carry out an epimerisation. A second molecule of L-cysteine is adenylated by the A domain of Sven0517 and linked to its PCP domain. The Cy domain of Sven0517 catalyses a similar condensation and cyclisation reaction to generate a 2-hydroxyphenyl-*bis*-thiazolinyl thioester. Until this point, all the reactions are analogous to the biosynthesis of enantio-pyochelin **76** in *P. protegens* (Figure 1.21).

The MT domain in Sven0517 probably catalyses the methylation at C-4" with inversion of stereochemistry. The thiazoline reductase Sven0516 then catalyses reduction of the

second thiazoline ring. The selective reduction of the second thiazoline by Sven0516 seems to be due to the opposite stereochemistry of the two cysteine-derived stereocenters. Thioester hydrolysis catalysed by the C-terminal TE domain of Sven0517 at this stage results in the production of isopyochelin. Hydrolysis after *N*-methylation by the MT domain of Sven0517 leads to the production of thiazostatin, which is methylated by the class B RS methylase Sven0515 to yield watasemycin. Aerugine **75** is proposed to be produced from hydrolytic cleavage and reduction of thiazostatin/isopyochelin while pulicatins A/B **77**/**78** derive from watasemycin in the same way.

4.7 Summary

Bioinformatics analyses of the *S. venezuelae* ATCC10712 genome sequence revealed the presence of a putative biosynthetic gene cluster, *sven0503-sven0517*, for pyochelin-like metabolites. Heterologous expression of this gene cluster in *S. coelicolor* M1152 resulted in identification and characterisation of the main metabolic products, watasemycin and thiazostatin, as well as other 2-hydroxyphenylthiazolines, including aerugine and pulicatins A/B, which are putatively identified by high resolution LC-MS. A novel compound, isopyochelin, has also been found to be one of the metabolic products from the gene cluster. The absolute configuration of isopyochelin at C-4" has been unambiguously established as *S* based on comparison with synthetic standards.

Both Sven0508 and Sven0516 share sequence similarities to the thiazoline reductase PchK. Gene deletion of *sven0508* did not affect the production of all the 2-hydroxyphenylthiazolines from the *sven0503-sven0517* gene cluster, suggesting that it is not involved in the biosynthetic pathway. Deletion of *sven0516* resulted in abolishment of all the 2-hydroxyphenylthiazoline metabolic products from the gene cluster, indicating it plays an essential role as the putative thiazoline reductase. A deletion mutant of *sven0515*, which putatively encodes a homologue of class B RS methylases, lost the ability to produce watasemycin and pulicatins A/B, implying a critical role of Sven0515 in methylation at C-5'. Further feeding experiments, supplying the extract from the Δ*sven0515* strain to the Δ*sven0516* strain showed complete conversion of thiazostatin to watasemycin, suggesting that thiazostatin is the substrate for Sven0515-catalysed methylation.

The absolute stereochemistry of watasemycin and thiazostatin has been reassigned as 4'*S*, 5'*S*, 4"*S* and 4'*S*, 4"*S*, respectively, based on the experimental results and analysis. The reassignment relies on the assumption that watasemycin and thiazostatin have the same 4"*S* configuration as isopyochelin. Moreover, based on these results, a biosynthetic pathway has been proposed for the 2-hydroxyphenylthiazolines from the *sven0503-sven0517* gene cluster in *S. venezuelae* ATCC10712.

The stereochemical course of the methylation catalysed by Sven0515 on the chemically inert 5' carbon of thiazostatin has been investigated by incorporation experiments using chemically synthesised L-cysteine stereospecifically deuterium labelled at the β-position. LC-MS analyses showed that the *pro-R* hydrogen atom at C-5' of thiazostatin is

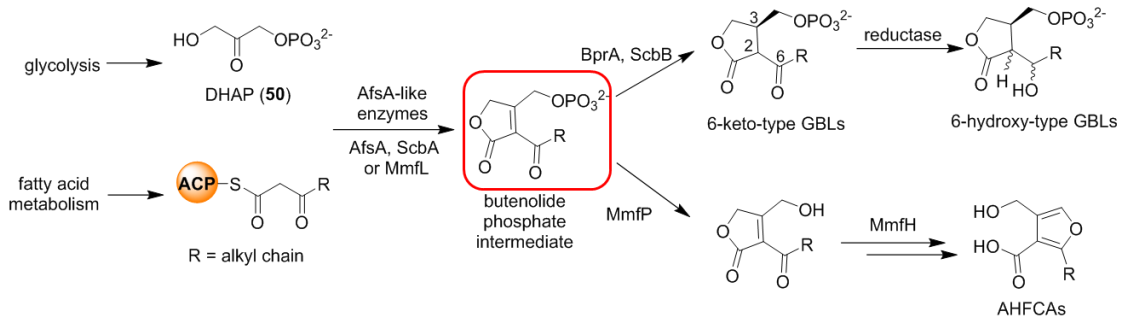
substituted by the transferred methyl group while the *pro-S* deuterium is retained. The stereochemical course of the Sven0515-catalysed methylation reaction therefore proceeds with inversion of configuration, consistent with the mechanism of C-methylation by other class B RS methylases. This also reciprocally supports the reassignment of the stereochemistry of watasemycin **83** and thiazostatin **82**.

5. Conclusions and future work

5.1 MmfL catalyses the formation of a butenolide intermediate in AHFCA biosynthesis

A TBDMS-protected derivative of the MmfL reaction product has been synthesised using the methodology reported by Sello and co-workers,[168] involving condensation of an appropriate acyl Meldrum's acid and TBDMS mono-protected dihydroxyacetone to give the corresponding β-ketoester, which cyclises during purification by flash chromatography on silica. Deprotection using a hydrogen fluoride-pyridine complex led to the desired butenolide, which was confirmed by its ESI-MS/MS analysis. It should be noted that the TBDMS-protected butenolide is unstable and degraded quickly to a species corresponding to the deprotected butenolide.

The dephosphorylated product of the MmfL-catalysed reaction of DHAP and the NAC β-ketothioester has been identified to be the butenolide by comparing to the synthetic standard. Thus, it can be concluded that MmfL catalyses the reaction between a β-ketothioester and DHAP to form a phosphorylated butenolide intermediate, which is an analogous reaction to that catalysed by AfsA during the γ-butyrolactone signalling molecule A-factor biosynthesis in *S. griseus*.[67] Therefore, the formation of GBLs and AHFCAs are proposed to proceed *via* a shared butenolide phosphate intermediate (Scheme 5.1).[74-75] Each type of enzyme (i.e., AfsA, MmfL or ScbA) presumably has a different preference for the length of the alkyl chain in the β-ketothioester substrate.



Scheme 5.1. General scheme for the biosynthesis of *Streptomyces* signalling molecules, GBLs and AHFCAs, which are proposed to share a butenolide phosphate intermediate.

As the biosynthetic roles of both MmfL and MmfP have been established *in vitro*, it would be of particular interest to study the function of another key enzyme, MmfH, in AHFCA biosynthesis. Based on previous work,[77] one promising way would be to elucidate the X-ray crystal structure of MmfH.

5.2 Interactions between DNA, ligand and ArpA-like repressors

5.2.1 Three operators for MmfR binding

Three conserved MAREs within the methylenomycin biosynthetic gene cluster have been confirmed as the binding sites for MmfR, including the operator upstream of *mmyR* and the operators located in the *mmyY-mmyB, mmfR-mmfL* intergenic regions. The order of affinity of MmfR for the three operators has been shown *in vitro* as follows: the operator upstream of *mmyB* > the operator between *mmfR* and *mmfL* > the operator upstream of *mmyR*.

This order partially supports the proposed regulatory cascade in methylenomycin biosynthesis. When AHFCA production reaches a critical concentration, they bind to the MmfR repressor and cause derepression of MmfR binding to the operator within the *mmfL-mmfR* intergenic region, leading to a significant increase of the production of AHFCAs. The increased concentration of AHFCAs is then able to release MmfR from the *mmyB* promoter, which has the highest affinity, and thus the transcriptional activator MmyB can be produced. MmyB binds to its operator regions ('B-boxes'), initiating the transcription of the structural genes to produce methylenomycins.

Previous gene deletion experiments showed that MmyR-like repressors likely act as an off-switch to stop the production of specialised metabolites, and appear to be unresponsive to signalling molecules. Due to lack of direct evidence on the function of MmyR, further studies are required to fully reveal the mechanism of the transcriptional regulation involving MmfR/MmyR/AHFCAs in methylenomycin biosynthesis.

5.2.2 Molecular basis for AHFCA recognition by MmfR

Site-directed mutagenesis based on the X-ray crystal structure of MmfR with AHFCA2 bound was used to investigate ligand binding and recognition by MmfR. Tyr85Ala, Tyr85Phe, Tyr144Ala, Tyr144Phe and Gln130Glu mutants of MmfR were prepared and their interaction with DNA and AHFCA1 were assessed *in vitro* using EMSAs. Key residues involved in interactions with AHFCAs in MmfR were identified, in particular tyrosines at positions 85 and 144, glutamine at position 130 and leucine at position 110.

The hydroxyl group of Tyr85 and the amide group in the side chain of Gln130 were confirmed to play important roles in the interaction of MmfR with the ligand through hydrogen bond formation with the carboxyl group of AHFCAs. Alteration of these residues attenuates the ability of the ligand to release MmfR from its operators. For Tyr144, it is likely that its backbone amide participates in ligand recognition and binding while its side chain may also be involved in the formation of the ligand binding pocket. Leu110 is probably important for proper folding and stability of MmfR. These results have illuminated the molecular basis for AHFCA recognition by MmfR.

5.2.3 MmfR is not responsive to γ-butyrolactones

A representative of the γ-butyrolactone class of signalling molecules, SCB1, was synthesised in this study to determine whether there is any cross-talk between the regulation systems mediated by GBLs and AHFCAs. The effect of SCB1 on the binding of MmfR to its operator was investigated *in vitro* using EMSAs. The data showed that SCB1 is unable to release MmfR from its operator. This demonstrated that ligand binding for MmfR is highly specific, supporting the hypothesis that GBLs and AHFCAs mediate pathway-specific regulation of antibiotic biosynthesis in *S. coelicolor* A3(2).

5.2.4 SgnR in the regulation of gaburedin biosynthesis

As a close homologue of MmfR, SgnR was proposed to be able to bind AHFCA ligands. *In silico* analysis also identified three ARE sequences in the gaburedin biosynthetic gene cluster, which were proposed to be the binding sites for SgnR. These predictions have been experimentally confirmed by *in vitro* studies using EMSAs.

It has been shown that SgnR is able to bind the MARE sequence from *S. coelicolor* in a pre-test, and the three ARE sequences from the *S. venezuelae gbn* gene cluster, including the operator upstream of *gbnA* and the operators located in the *sgnR-sgnL, gbnR-sgnH* intergenic regions. The SgnR-DNA complexes can be dissociated from its operators by addition of AHFCA6, indicating that gaburedin biosynthesis in *S. venezuelae* is regulated by a similar system to that used for the regulation of methylenomycin biosynthesis in *S. coelicolor*.

Additionally, the key amino acids in MmfR that have been demonstrated in the above studies to be important for ligand binding or protein stability, are conserved in SgnR (Figure 5.1), further supporting the similar roles played by SgnR and MmfR in regulation of specialised metabolite biosynthesis.

```
SgnR      --------------MATPRSQPKQERARRTKVHILQSAAELFAERGYATVTLQDVAERAE
MmfR      MTSAQQPTPFAVRSNVPRGPHPQQERSIKTRAQILEAASEIFASRGYRGASVKDVAERVG
                        .    :*:***: :*:.:**::*:*:**.***  .:::*****.

SgnR      MTKGAVYFHYTNKEALAVAVVQEHYARWPEILKGAEGDHAEPFDMLTAVLDTVTRAFARD
MmfR      MTKGAVYFHFPSKESLAIAVVEEHYARWPAAMEEIRIQGFTPLETVEEMLHRAAQAFRDD
          *********: .**:**:***:*******  ::  . :   *:: :  :*. .::**  *

SgnR      IVVQAGARLQIERALIDAELPEPYVGWEDYLTRLIAEARDAGQLRDGVEPRAAARVLVSA
MmfR      PVMQAGARLQSERAFIDAELPLPYVDWTHLLEVPLQDAREAGQLRAGVDPAAAARSLVAA
           *:******* ***:****** *** * . *   : :**:***** **:* **** **:*

SgnR      FFGMQHISDVLSGRSDLTERYEELRTVLLEGLRR
MmfR      FFGMQHVSDNLHQRADIMERWQELRELMFFALRA
          ******:** *  *:*: **::*** ::: .**
```

Figure 5.1. Sequence alignment of SgnR and MmfR. The key amino acid residues identified from MmfR are highlighted in yellow.

5.3 Watasemycin biosynthesis in *S. venezuelae* ATCC10712

5.3.1 Isopyochelin is a shunt metabolite in the watasemycin biosynthetic pathway

Bioinformatics analyses of the *S. venezuelae* ATCC 10712 genome sequence revealed the presence of a putative biosynthetic gene cluster, *sven0503-sven0517*, for pyochelin-like metabolites. Most of the genes encode proteins with high sequence similarity to those encoded by genes within the *S. scabies* pyochelin biosynthetic gene cluster. Heterologous expression of the gene cluster in *S. coelicolor* M1152 revealed thiazostatin and watasemycin as the main metabolic products of the cluster, as well as other putative 2-hydroxyphenylthiazolines identified by LC-MS, including aerugine and pulicatins A/B.

A metabolite with molecular formula identical to pyochelin was also identified. Comparison to synthetic standards showed that it is actually not pyochelin, but isopyochelin, which is a structural isomer of pyochelin with the C-4" position methylated instead of the nitrogen atom of the thiazolidine. The configuration at C-4" of isopyochelin was established as *S* by comparing the metabolite to synthetic stereoisomers by chiral LC-MS.

5.3.2 Sven0516 is the only thiazoline reductase required for watasemycin biosynthesis

Sven0516 is similar in sequence to PchK from *P. protegens*, which has been proposed to function as a thiazoline reductase in enantio-pyochelin biosynthesis. Deletion of *sven0516* (carried out by Dr. Yuki Inahashi) abolished the production of all the 2-hydroxyphenylthiazoline metabolic products, indicating its key role as a thiazoline reductase in the watasemycin biosynthetic pathway. Production of aerugine and pulicatins, which ostensibly do not require a thiazoline reductase for their biosynthesis, was also abolished in the mutant strain, suggesting that they may be derived from hydrolysis and reduction from thiazostatin/isopyochelin and watasemycin, respectively. The presence of a PchK-like thiazoline reductase is consistent with the assignment of *S* configuration at C-4"of isopyochelin, as enantiopyochelin has the same *S* configuration at C-4".

Another PchK homologue, Sven0508, has been found not to be involved in watasemycin biosynthesis since deletion of the gene encoding it resulted in a similar metabolic profile to the wild type strain (results of Dr. Yuki Inahashi).

5.3.3 Sven0515 catalyses C-methylation of thiazostatin to produce watasemycin

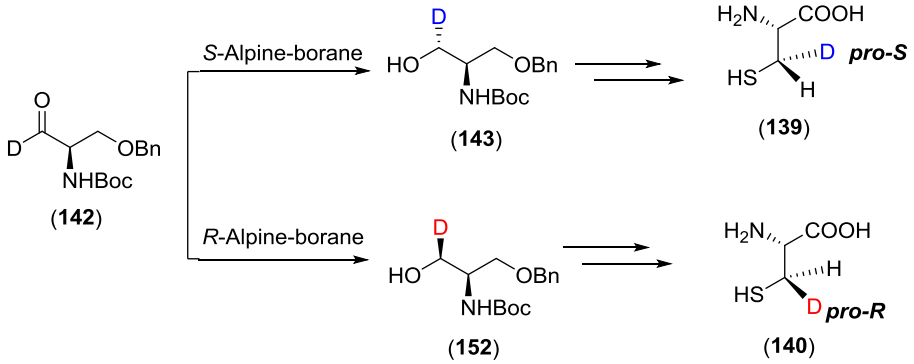
Deletion of *sven0515* (carried out by Dr. Yuki Inahashi) abolished the production of pulicatins and watasemycin, both of which contain a methyl group at C-5' on the thiazoline ring, with no effect on the production of thiazostatin, aerugine and isopyochelin. This revealed an essential role for Sven0515, a class B RS methylase, in the C-methylation of the thiazoline ring in watasemycin biosynthesis. Feeding of natural thiazostatin extracted from the culture of the Δ*sven0515* mutant to the Δ*sven0516* strain resulted in its complete transformation to watasemycin, indicating that thiazostatin is the substrate of Sven0515. The Sven0515-catalysed methylation of thiazostatin is the final step in watasemycin biosynthesis.

Class B RS methylases are known to catalyse methylation of un-activated carbon centres in the biosynthesis of a wide variety of natural products, including aminoglycosides, aminocoumarins, β-lactams, enediynes, polyketides, phosphonates, and RiPPs. To our knowledge, the Sven0515-catalysed methylation of thiazostatin is the first example of

such a reaction that has been experimentally validated to be found in the biosynthesis of a nonribosomal peptide.

The stereochemical course of the methylation on the chemically inert 5' carbon of thiazostatin has been investigated by incorporation experiments using chemically synthesised L-cystine with the *pro-S* hydrogen atom replaced by a deuterium atom at the β-position. LC-MS analyses showed retention of the deuterium label at C-5' in watasemycin, indicating that the *pro-R* hydrogen atom of thiazostatin is substituted by the appended methyl group while the *pro-S* deuterium is retained. The stereochemical inversion during the Sven0515-catalysed methylation is consistent with other members of class B RS methylases, such as GenD1 in gentamicin biosynthesis and MoeK5 in moenomycin biosynthesis.[149,165]

Feeding with synthetic L-cysteine with the *pro-R* hydrogen atom replaced by a deuterium atom at the β-position would further support the proposed stereochemical course. In this case, the *pro-R* deuterium atom should be abstracted during methylation, resulting in singly deuterium-labelled watasemycin. The synthesis of this deuterium-labelled cysteine could be achieved by using the same synthetic approach with the corresponding stereospecific reducing agent, *R*-Alpine-Borane (Scheme 5.2).[201]



Scheme 5.2. Synthetic route to L-cysteine with the *pro-S* or *pro-R* hydrogen atom replaced by a deuterium atom at the β-position.

Enzymatic assays *in vitro* using the purified and reconstituted Sven0515 enzyme and synthetic thiazostatin substrate harboring a stereospecifically-incorporated deuterium label at the site of methylation would provide direct evidence for the stereochemical course of the methylation reaction. Furthermore, despite the important role played by class B RS methylases in antibiotic biosynthesis, no X-ray crystal structures of enzymes belonging to this family have been reported to date. Thus, it will be useful to determine

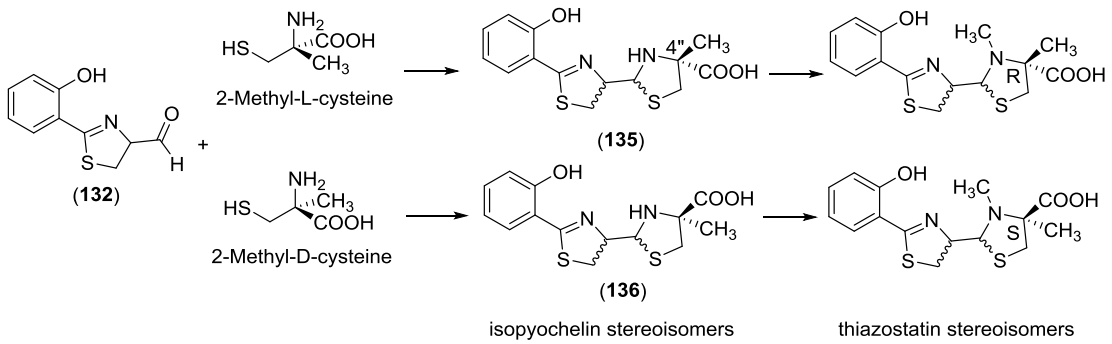
the X-ray crystal of Sven0515, a member of this family of methylases, which would provide a clear picture of the structure-function relationships of these enzymes. Due to the oxygen-sensitive nature of class B RS methylases, crystallisation trials for Sven0515 would need to be set up in an anaerobic chamber.

This would develop a better understanding of the catalytic mechanism employed by class B RS methylases to methylate inert carbon centres. Such reactions, which are key steps in the biosynthesis of numerous natural products, are currently impossible using anthropogenic methodologies. Thus, the information obtained will underpin development of new methodologies in biotechnology and production of new or improved antimicrobial compounds.

5.3.4 Stereochemistry of thiazostatin and watasemycin

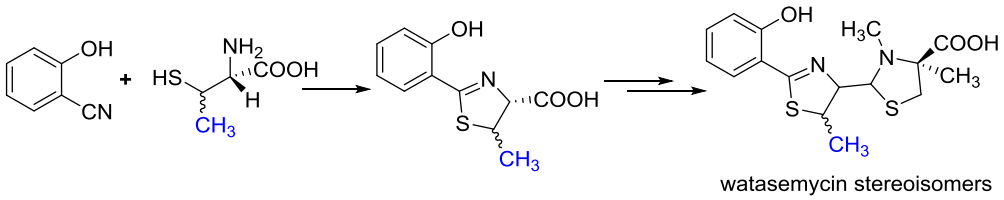
The relative stereochemistry of thiazostatin and watasemycin has been reported before this study. However, their absolute stereochemistry hasn't been elucidated since their discovery. Our absolute stereochemical assignment of thiazostatin and watasemycin is inferred on the basis of the 4"*S* absolute configuration determined for isopyochelin, as they are all derived from the same biosynthetic pathway.

Thus, it will be important to experimentally confirm the absolute configuration at C-4" of thiazostatin, which is the substrate for Sven0515. This could be achieved in the same way as the configuration of isopyochelin was determined, involving synthesis of thiazostatin stereoisomers, followed by subsequent chiral LC-MS comparison with the natural metabolites (Scheme 5.3).



Scheme 5.3. Proposed route to characterise the absolute configuration at C-4" of thiazostatin.

The configuration at C-5' of watasemycin could be characterised using the same methodology. However, condensation with a β-methyl-cysteine, instead of a normal cysteine, in the first step is required to introduce a methyl group at C-5' of the final product (Scheme 5.4).



Scheme 5.4. Proposed route to characterise the absolute configuration at C-5' of watasemycin.

During the synthesis of pyochelin, it has been noticed that epimerisation occurs at the C-4' position under the reaction conditions, which has been reported in the literature.[181-182] This makes preparation of an authentic standard with a specific configuration at C-4' impossible. Moreover, it appears that feeding of isotope-labelled precursors, such as 2-^{2}H-L-cysteine, would also be ineffective, due to the fact that the deuterium atom at the α-position of the cysteine may be washed out by primary metabolism before incorporation into the biosynthesis, e.g., by a pyridoxal phosphate (PLP)-mediated exchange reaction.[203] However, the C-4' configuration can be deduced from the finding that the C-4' hydrogen substituent and the C-5' methyl group in watasemycin are *anti* to each other on the basis of NOE studies.

6. Experimental

6.1 Synthetic chemistry

6.1.1 General information

Dry solvents, including THF, DMF, DCM, Et_2O and toluene, were obtained from the solvent towers in the O'Reilly Group (Department of Chemistry, University of Warwick) and stored over activated 4 Å molecular sieves under argon. All other solvents were purchased from Sigma-Aldrich, Fisher Scientific or VWR and used without further purification. All reagents were purchased from Sigma-Aldrich, Fisher Scientific, VWR, Alfa Aesar, Acros Organic, Scientific Laboratory Supplies, Atlantic Research Chemicals or Carbosynth and used without further purification. All air- and moisture-sensitive manipulations were carried out with standard Schlenk techniques under argon.

Flash column chromatography was conducted on Sigma-Aldrich silica gel (40-63 μm, 60 Å). TLC was performed on aluminium backed sheets pre-coated with Merck silica gel 60 F_{254} and visualised by UV radiation, potassium permanganate or ninhydrin stains. Preparative TLC was performed on Analtech 20 cm × 20 cm silica gel plate with a thickness of 2000 μm on glass support for purification of small amounts of product which are UV-visible. Solvents were evaporated using a Buchi Rotavapor R-200 equipped with a Buchi Vacuum pump V-700.

The ^{1}H- and ^{13}C-NMR spectra were obtained using a Bruker AV-250 MHz, Avance III HD 300, 400 or 500 MHz spectrometer. The variable temperature NMR experiments were obtained using a Bruker AV-600 MHz spectrometer. Deuterated solvents were purchased from Sigma-Aldrich and used as supplied. All chemical shifts are given as δ values in ppm with reference to residual HDO (δ_H 4.79), $CHCl_3$ (δ_H 7.26) or TMS (δ_H 0.00) if residue $CHCl_3$ peak was difficult to be distinguished (e.g. overlapping with other peaks), the central peak of $CDCl_3$ (δ_C 77.00) or DMSO-d_6 (δ_C 39.52). Data for ^{1}H NMR spectra are reported as follows: chemical shift (δ, ppm), multiplicity (s = singlet, d = doublet, t = triplet, dd = doublet of doublets, td = triplet of doublets, ddd = doublet of doublet of doublets, br = broad, m = multiplet), integration, assignment and coupling constant (Hz).

Low resolution ESI mass spectra were recorded using an Agilent 6130B single quad spectrometer. HRMS spectra were measured by the University of Warwick mass

spectrometry service on Bruker MaXis or MaXis impact mass spectrometer. Melting points were measured in open capillaries using a Stuart SMP10 melting point apparatus and were uncorrected.

General procedure for condensation to yield pyochelin and isopyochelin

To a solution of 2-(2-hydroxyphenyl)-4,5-dihydrothiazole-4-carbaldehyde (25 mg, 0.12 mmol) in a mixture of 5 mL ethanol and 1.3 mL water was added CH_3COOK (82 mg, 0.84 mmol) and the hydrochloride salt of the corresponding methylated cysteine (30 mg, 0.17 mmol). After being stirred at RT for 24 hours, 15 mL water was added and the mixture washed once with hexane. The aqueous phase was then acidified to pH 5 and extracted with EtOAc. The combined EtOAc layers were washed with brine, dried over $MgSO_4$ and concentrated under vacuum to yield crude pyochelin or isopyochelin as a yellow solid. Due to the formation of four diastereoisomers in this step and the fact that the free carboxylic acids cannot be isolated in pure form due to their rapid interconversion, NMR assignments of the product were difficult. Therefore, these compounds were analysed by LC-MS directly.

General procedure for acyl Meldrum's acid synthesis starting from carboxylic acid

To a solution of carboxylic acid (1 eq) in DCM was added EDC.HCl (1.5 eq) and the mixture cooled to 0 ℃. Then DMAP (3 eq) dissolved in DCM was added, followed by Meldrum's acid (1 eq) dissolved in DCM. After stirring at RT for 18 hours, the reaction mixture was diluted with DCM, washed with 1 M HCl, water and brine. The organic phase was dried over $MgSO_4$ and concentrated under vacuum. The residue was purified by column chromatography to yield the product.

General procedure for butenolide synthesis

A solution of monosilyl DHA (1 eq) and acyl Meldrum's acid (1.2 eq) in 2 mL toluene was refluxed for 3 hours at which point was added more acyl Meldrum's acid (0.5 eq). After stirring for an additional 5 hours, the reaction mixture was allowed to stand at -20 ℃ for 19 hours and purified by column chromatography (hexanes/acetone = 60:1 v/v) without removal of toluene to yield the desired product.

General procedure for butenolide reduction

To a solution of $NaBH_3CN$ (2 eq) in acetic acid was dropwise added a solution of butenolide (1 eq) in acetic acid and the mixture stirred at RT for 8 hours. After removal of acetic acid under vacuum, the oily residue was diluted with EtOAc and washed with 5% $NaHCO_3$ solution and brine. The organic phase was then dried over $MgSO_4$ and concentrated under vacuum. The residue was purified by flash column chromatography to yield the product.

The reduction would result in the formation of three stereogenic centres in the molecule so NMR spectra contain peaks from a mixture of different stereoisomers. This results in excessive peaks and inaccurate integration. The diastereoisomers were isolated in the next step after removal of the silyl group.

General procedure for silyl deprotection of γ-butyrolactones

A 0.1 M solution of TBDMS-protected γ-butyrolactone in a 6:3:1 mixture of THF/HCOOH/H_2O was stirred at RT overnight. The mixture was adjusted to pH 4 with saturated $NaHCO_3$ solution and extracted with EtOAc. The combined EtOAc layers were dried over $MgSO_4$ and concentrated under vacuum. The diastereoisomers were purified and separated by column chromatography (pet ether/EtOAc = 1:1 v/v).

General procedure for β-ketoester synthesis

A solution of acyl Meldrum's acid in methanol was refluxed overnight. After removal of the solvent under vacuum, the residue was purified by column chromatography to afford the desired β-ketoester.

General procedure for AHFCA cyclisation

To a 0.5 M solution of β-ketoester (1 eq) and 2,2-dimethyl-1,3-dioxan-5-one (1 eq) in methanol was added scandium III triflate (0.1 eq) and the reaction mixture stirred at RT overnight. After removal of the solvent under vacuum, the residue was purified by column chromatography to give the desired furan.

General procedure for hydrolysis of AHFCA methyl ester

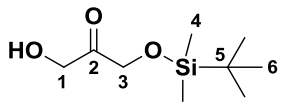
To a solution of the furan ester (1 eq) in a 1:1 mixture of THF/water was added LiOH (2.5 eq) and the reaction mixture stirred at RT overnight. After removal of the THF under vacuum, the remaining aqueous phase was washed with hexane, acidified with 1 M HCl until a white precipitate formed, and extracted with Et_2O. The combined organic layers were washed with brine, dried over $MgSO_4$ and concentrated to afford the product.

General procedure for β-ketothioester synthesis

To a 0.5 M solution of acyl Meldrum's acid (1.5 eq) in 1,2-dichloroethane under argon was added *N*-acetylcysteamine (1 eq) and the reaction mixture refluxed for 3 hours. After removal of the solvent under vacuum, the residue was purified by column chromatography with a plug of silica impregnated with $CuSO_4$ (10 cm) on top to afford the desired product.

6.1.2 Synthesis of monosilyl DHA

1-((*tert*-Butyldimethylsilyl)oxy)-3-hydroxypropan-2-one 102



To a mixture of dihydroxyacetone (5.00 g, 55.6 mmol) and imidazole (1.51 g, 22.2 mmol) in 50 mL DMF at 0 ℃ was added a solution of *tert*-butyldimethylsilyl chloride (2.67 g, 17.8 mmol) in 10 mL DMF and the reaction stirred at RT for 17 hours. Then 50 mL water was added and the mixture extracted with Et_2O. The organic phase was washed with brine, dried over $MgSO_4$ and concentrated under vacuum. The residue was purified by column chromatography (pet ether/EtOAc = 5:1 v/v) to give 1-((*tert*-butyldimethylsilyl)oxy)-3-hydroxypropan-2-one (1.82 g, 50%) as a colourless oil.

^{1}H NMR (300 MHz, $CDCl_3$) δ: 0.12 (s, 6H, H-4), 0.90 (s, 9H, H-6), 2.99 (br s, 1H, O-H), 4.31 (s, 2H, H-1), 4.50 (s, 2H, H-3), further signals at 3.43 - 4.24 due to formation of the intermolecular ketal compounds.

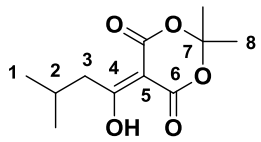
^{13}C NMR (75 MHz, $CDCl_3$) δ: 210.88 (C-2), 25.74 (C-6), 18.15 (C-5), -5.56 (C-4), further signals at 63.07 - 111.38.

HR-MS: *m/z* calculated for $C_9H_{20}NaO_3Si$ $[M+Na]^+$: 227.1074; found: 227.1076.

6.1.3 Synthetic MmfL substrate and enzymatic product

6.1.3.1 Synthesis of NAC β-ketothioester

2,2-Dimethyl-5-(3-methylbutanoyl)-1,3-dioxane-4,6-dione 99



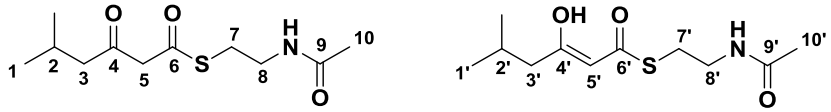
To a solution of Meldrum's acid (1.79 g, 12.4 mmol) in 25 mL DCM was added triethylamine (2.52 g, 25.0 mmol) and DMAP (0.30 g, 2.5 mmol) and the mixture stirred for 10 mins at RT before being cooled to 0 ℃. Then 3-methylbutanoyl chloride (1.50 g, 12.4 mmol) was added dropwise and the reaction warmed to RT and stirred overnight. The resulting mixture was washed with 1 M HCl and water, dried over $MgSO_4$ and concentrated under vacuum. The residue was purified by column chromatography (pet ether/EtOAc = 2:1 v/v) to yield 2,2-dimethyl-5-(3-methylbutanoyl)-1,3-dioxane-4,6-dione (2.80 g, 99%) as a yellow oil.

^{1}H NMR (300 MHz, $CDCl_3$) δ: 0.89 (d, 6H, H-1, J = 7.0 Hz), 1.61 (s, 6H, H-8), 2.03 - 2.12 (m, 1H, H-2), 2.85 (d, 2H, H-3, J = 7.0 Hz), 15.19 (br s, 1H, O-H)

^{13}C NMR (75 MHz, $CDCl_3$) δ: 170.22 (C-4), 159.89 (C-6), 104.33 (C-7), 91.62 (C-5), 43.45 (C-3), 27.05 (C-2), 26.40 (C-8), 22.17 (C-1)

LR-MS: *m/z* calculated for $C_{11}H_{15}O_5$ $[M-H]^-$ found: 227.1.

***S*-(2-Acetamidoethyl) 5-methyl-3-oxohexanethioate 61**



Following the general procedure for β-ketothioester synthesis, work up gave *S*-(2-acetamidoethyl) 5-methyl-3-oxohexanethioate (0.21 g, 60%) as a yellowish oil. The keto-enol form ratio was nearly **2:1**, estimated from integration of the H-5 and H-5' peaks

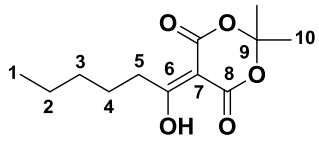
in ^{1}H NMR.

^{1}H NMR (400 MHz, $CDCl_3$) δ: 0.89 (d, 6H, H-1, J = 6.5 Hz), 1.94 (s, 3H, H-10), 1.99 (br s, 2H, H-3'), 2.11 (nonet, 1H, H-2, J = 6.5 Hz), 2.37 (d, 2H, H-3, J = 7.0 Hz), 3.05 (t, 2H, H-7, J = 6.0 Hz), 3.42 (t, 2H, H-8, J = 6.0 Hz), 3.65 (s, 2H, H-5), 5.40 (s, 1H, H-5'), 6.23 (br s, 1H, N-H), 12.55 (br s, 1H, O-H)

^{13}C NMR (100 MHz, $CDCl_3$) δ: 194.08 (C-4'), 192.21 (C-4), 176.54 (C-6), 170.51 (C-9), 170.36 (C-9'), 100.04 (C-5'), 57.45 (C-5), 52.10 (C-3), 43.95 (C-3'), 39.74 (C-8'), 39.04 (C-8), 29.04 (C-7), 27.67 (C-7'), 26.33 (C-2), 24.23 (C-10), 22.29 (C-1)

HR-MS: *m/z* calculated for $C_{11}H_{19}NNaO_3S$ $[M+Na]^+$: 268.0978; found: 268.0985.

5-(1-Hydroxyhexylidene)-2,2-dimethyl-1,3-dioxane-4,6-dione 100



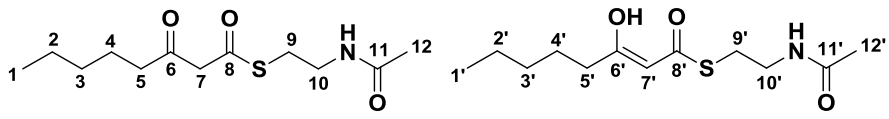
Following the general procedure for acyl Meldrum's acid synthesis, work up gave 5-(1-hydroxyhexylidene)-2,2-dimethyl-1,3-dioxane-4,6-dione (1.67 g, 67%) as a yellow oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ: 0.78 (t, 3H, H-1, J = 7.0 Hz), 1.17 - 1.31 (m, 4H, H-2 and 3), 1.54 - 1.58 (m, 2H, H-4), 1.61(s, 6H, H-10), 2.93 (t, 2H, H-5, J = 7.5 Hz), 15.18 (br s, 1H, O-H)

^{13}C NMR (100 MHz, $CDCl_3$) δ: 170.23 (C-6), 159.77 (C-8), 104.37 (C-9), 90.96 (C-7), 35.28 (C-5), 31.14 (C-3), 26.38 (C-10), 25.46 (C-4), 21.93 (C-2), 13.48 (C-1)

LR-MS: *m/z* calculated for $C_{12}H_{17}O_5$ $[M-H]^-$ found: 241.1.

***S*-(2-Acetamidoethyl) 3-oxooctanethioate 101**



Following the general procedure for β-ketothioester synthesis, work up gave *S*-(2-acetamidoethyl) 3-oxooctanethioate (0.35 g, 80%) as a white solid. The keto-enol form

ratio was nearly **3:1**, estimated from integration of the H-5 and H-5' peaks in ^{1}H NMR.

^{1}H NMR (400 MHz, $CDCl_3$) δ: 0.86 - 0.90 (m, 3H, H-1), 1.24 - 1.31 (m, 4H, H-2 and 3), 1.58 (quintet, 2H, H-4, J = 7.0 Hz), 1.96 (s, 3H, H-12), 2.16 (t, 2H, H-5', J = 7.5 Hz), 2.51 (t, 2H, H-5, J = 7.5 Hz), 3.05 - 3.09 (m, 2H, H-9), 3.42 - 3.48 (m, 2H, H-10), 3.68 (s, 2H, H-7), 5.45 (s, 1H, H-7'), 6.02 (br s, 1H, N-H), 12.59 (br s, 1H, O-H)

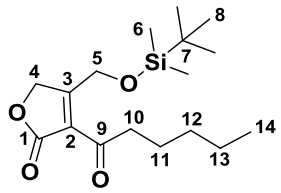
^{13}C NMR (100 MHz, $CDCl_3$) δ: 194.24 (C-6'), 192.34 (C-6), 177.63 (C-8), 170.41 (C-11), 99.09 (C-7'), 57.14 (C-7), 43.37 (C-5), 39.86 (C-5'), 39.14 (C-10), 34.82 (C-10'), 31.19 (C-9'), 31.07 (C-9), 29.17 (C-3), 27.78 (C-3'), 25.86 (C-4'), 23.11 (C-12), 23.05 (C-4), 22.32 (C-2), 22.28 (C-2'), 13.82 (C-1)

HR-MS: *m/z* calculated for $C_{12}H_{21}NNaO_3S$ $[M+Na]^+$: 282.1134; found: 282.1147.

Melting point: 89-90 ℃.

6.1.3.2 Synthetic enzymatic product

4-(((*tert*-Butyldimethylsilyl)oxy)methyl)-3-hexanoylfuran-2(5*H*)-one 106



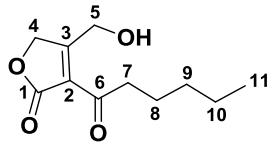
Following the general procedure for butenolide synthesis, work up gave 4-(((*tert*-butyldimethylsilyl)oxy)methyl)-3-hexanoylfuran-2(5*H*)-one (0.17 g, 30%) as a yellowish oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ: 0.08 (s, 6H, H-6), 0.81 - 0.92 (m, 12H, H-8 and 14), 1.29 - 1.32 (m, 4H, H-12 and 13), 1.59 (quintet, 2H, H-11, J = 7.0 Hz), 2.95 (t, 2H, H-10, J = 7.5 Hz), 4.97 (s, 2H, H-5), 5.05 (s, 2H, H-4)

^{13}C NMR (100 MHz, $CDCl_3$) δ: 197.08 (C-9), 181.16 (C-1), 170.61 (C-3), 122.43 (C-2), 70.17 (C-4), 61.81 (C-5), 41.52 (C-10), 31.19 (C-11), 30.80 (C-7), 25.65 (C-8), 22.80 (C-12), 22.44 (C-13), 13.85 (C-14), -5.70 (C-6)

HR-MS: *m/z* calculated for $C_{17}H_{30}NaO_4Si$ $[M+Na]^+$: 349.1806; found: 349.1801.

3-Hexanoyl-4-(hydroxymethyl)furan-2(5*H*)-one 107

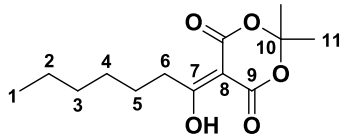


To a solution of 4-(((*tert*-butyldimethylsilyl)oxy)methyl)-3-hexanoylfuran-2(5*H*)-one (2 mg) in 200 μL THF at 0 ℃ was added 2 drops of hydrogen fluoride-pyridine and the mixture stirred for 40 mins at 0 ℃. The reaction was quenched with 500 μL saturated $NaHCO_3$ solution, concentrated under vacuum to remove the THF and extracted with EtOAc. The organic layers were combined, concentrated and redissolved in 700 μL of 1:1 v/v methanol/water. The sample was then filtered (0.4 μm) for LC-MS analyses.

HR-MS: *m/z* calculated for $C_{11}H_{16}NaO_4$ $[M+Na]^+$: 235.0941; found: 235.0948.

6.1.4 Synthesis of SCB6

5-(1-Hydroxyheptylidene)-2,2-dimethyl-1,3-dioxane-4,6-dione 115



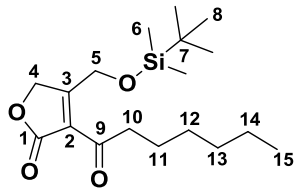
Following the general procedure for acyl Meldrum's acid synthesis, work up gave 5-(1-hydroxyheptylidene)-2,2-dimethyl-1,3-dioxane-4,6-dione (1.13 g, 55%) as a yellow oil.

^{1}H NMR (300 MHz, $CDCl_3$) δ: 0.76 (t, 3H, H-1, J = 6.5 Hz), 1.13 - 1.24 (m, 6H, H-2, 3 and 4), 1.29 (quintet, 2H, H-5, J = 7.0 Hz), 1.61 (s, 6H, H-11), 2.94 (t, 2H, H-6, J = 7.5 Hz), 15.19 (br s, 1H, O-H)

^{13}C NMR (75 MHz, $CDCl_3$) δ: 170.25 (C-7), 159.79 (C-9), 104.38 (C-10), 90.95 (C-8), 35.36 (C-6), 31.07 (C-3), 28.71 (C-4), 26.39 (C-11), 25.77 (C-5), 22.12 (C-2), 13.65 (C-1)

HR-MS: *m/z* calculated for $C_{13}H_{21}O_5$ $[M+H]^+$: 257.1384; found: 257.1374.

4-(((*tert*-Butyldimethylsilyl)oxy)methyl)-3-heptanoylfuran-2(5*H*)-one 116



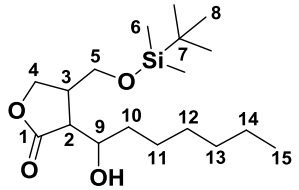
Following the general procedure for butenolide synthesis, work up gave 4-(((*tert*-butyldimethylsilyl)oxy)methyl)-3-heptanoylfuran-2(5*H*)-one (125 mg, 30%) as a yellow oil.

^{1}H NMR (300 MHz, $CDCl_3$) δ: 0.10 (s, 6H, H-6), 0.88 (t, 3H, H-15, J = 5.0 Hz), 0.90 (s, 9H, H-8), 1.25 - 1.29 (m, 6H, H-12, 13 and 14), 1.57 - 1.66 (m, 2H, H-11), 2.97 (t, 2H, H-10, J = 7.5 Hz), 4.99 (s, 2H, H-5), 5.06 (s, 2H, H-4)

^{13}C NMR (100 MHz, $CDCl_3$) δ: 70.20 (C-4), 61.84 (C-5), 41.60 (C-10), 31.59 (C-13), 28.72 (C-12), 25.66 (C-8), 23.07 (C-11), 22.46 (C-14), 13.99 (C-15), C-1, 2, 3, 6, 7 and 9 signals missing

HR-MS: *m/z* calculated for $C_{18}H_{33}O_4Si$ $[M+H]^+$: 341.2143; found: 341.2125.

4-(((*tert*-Butyldimethylsilyl)oxy)methyl)-3-(1-hydroxyheptyl)dihydrofuran-2(3*H*)-one 117



Following the general procedure for butenolide synthesis, work up gave 4-(((*tert*-butyldimethylsilyl)oxy)methyl)-3-(1-hydroxyheptyl)dihydrofuran-2(3*H*)-one (77 mg, 61%) as a colourless oil.

HR-MS: *m/z* calculated for $C_{18}H_{36}NaO_4Si$ $[M+Na]^+$: 367.2275; found: 367.2281.

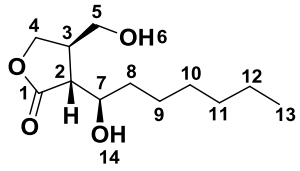
SCB6 stereoisomers 118

Following the general silyl deprotection procedure, work up and column chromatography gave a pair of SCB6 and its enantiomer (24 mg, 47%, higher polarity) and a pair of their

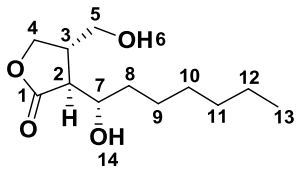
diastereoisomers (20 mg, 39%, lower polarity) as colourless oils.

(3*R*,4*R*)-3-((*R*)-1-Hydroxyheptyl)-4-(hydroxymethyl)dihydrofuran-2(3*H*)-one, SCB6

(3*S*,4*S*)-3-((*S*)-1-Hydroxyheptyl)-4-(hydroxymethyl)dihydrofuran-2(3*H*)-one, SCB6 enantiomer



SCB6



SCB6 enantiomer

^{1}H NMR (400 MHz, $CDCl_3$) δ: 0.88 (t, 3H, H-13, J = 7.0 Hz), 1.25 - 1.34 (m, 8H, H-9, 10, 11 and 12), 1.47 - 1.63 (m, 2H, H-8), 2.49 (br s, 2H, H-6 and 14), 2.65 (dd, 1H, H-2, J = 9.5 and 4.5 Hz), 2.72 - 2.82 (m, 1H, H-3), 3.68 (dd, 1H, H-5a, J = 10.5 and 6.5 Hz), 3.75 (dd, 1H, H-5b, J = 10.5 and 5.0 Hz), 3.98 (t, 1H, H-4a, J = 9.0 Hz), 4.02 (ddd, 1H, H-7, J = 9.0, 4.5 and 3.0 Hz), 4.42 (t, 1H, H-4b, J = 8.5 Hz)

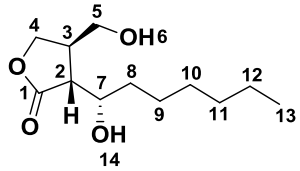
^{13}C NMR (125 MHz, $CDCl_3$) δ: 177.21 (C-1), 70.86 (C-7), 68.33 (C-4), 63.00 (C-5), 49.05 (C-2), 40.10 (C-3), 33.95 (C-8), 31.70 (C-11), 29.08 (C-9), 25.76 (C-10), 22.56 (C-12), 14.04 (C-13)

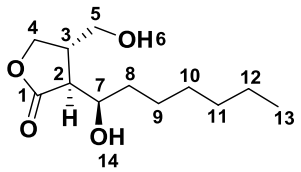
HR-MS: *m/z* calculated for $C_{12}H_{22}NaO_4$ $[M+Na]^+$: 253.1410; found: 253.1411.

SCB6 diastereoisomers

(3*R*,4*R*)-3-((*S*)-1-Hydroxyheptyl)-4-(hydroxymethyl)dihydrofuran-2(3*H*)-one

(3*S*,4*S*)-3-((*R*)-1-Hydroxyheptyl)-4-(hydroxymethyl)dihydrofuran-2(3*H*)-one





^{1}H NMR (400 MHz, $CDCl_3$) δ: 0.88 (t, 3H, H-13, J = 7.0 Hz), 1.25 - 1.35 (m, 8H, H-9, 10, 11 and 12), 1.45 - 1.64 (m, 2H, H-8), 2.00 (br s, 2H, H-6 and 14), 2.56 (dd, 1H, H-2, J = 7.5 and 3.5 Hz), 2.81 - 2.90 (m, 1H, H-3), 3.71 (dd, 1H, H-5a, J = 10.5 and 6.0 Hz),

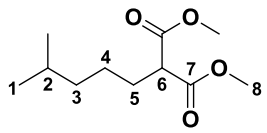
3.75 (dd, 1H, H-5b, *J* = 10.5 and 5.5 Hz), 4.08 - 4.14 (m, 2H, H-4a and 7), 4.42 (t, 1H, H-4b, *J* = 8.5 Hz)

^{13}C NMR (125 MHz, $CDCl_3$) δ: 178.41 (C-1), 70.80 (C-7), 69.37 (C-4), 63.32 (C-5), 48.05 (C-2), 38.07 (C-3), 34.81 (C-8), 31.71 (C-11), 29.06 (C-9), 25.76 (C-10), 22.56 (C-12), 14.03 (C-13)

HR-MS: *m/z* calculated for $C_{12}H_{22}NaO_4$ $[M+Na]^+$: 253.1410; found: 253.1411.

6.1.5 Synthesis of 6-methylheptanoic acid

Dimethyl 2-(4-methylpentyl)malonate 119



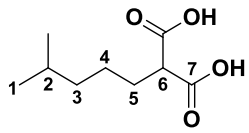
To a suspension of NaH (0.72 g, 30.0 mmol) in 60 mL THF at 0 ℃ was slowly added dimethyl malonate (2.90 g, 22.0 mmol) and the mixture stirred at the same temperature for 30 mins before 1-bromo-4-methylpentane (3.30 g, 20.0 mmol) and TBAI (1.48 g, 4.0 mmol) were added. After being stirred at 0 ℃ for an additional 30 mins, the reaction was allowed to warm to RT and heated to reflux overnight. The mixture was concentrated, acidified with 3 M HCl and extracted with DCM. The combined organic layers were washed with water and brine, dried over $MgSO_4$ and concentrated under vacuum. The residue was purified by column chromatography (pet ether/EtOAc = 10:1 v/v) to give dimethyl 2-(4-methylpentyl)malonate (3.54 g, 82%) as a colourless oil.

^{1}H NMR (300 MHz, $CDCl_3$) δ: 0.80 (d, 6H, H-1, *J* = 6.5 Hz), 1.09 - 1.17 (m, 2H, H-3), 1.20 - 1.30 (m, 2H, H-4), 1.48 (nonet, 1H, H-2, *J* = 6.5 Hz), 1.82 (quartet, 2H, H-5, *J* = 8.0 Hz), 3.31 (t, 1H, H-6, *J* = 7.5 Hz), 3.68 (s, 6H, H-8)

^{13}C NMR (75 MHz, $CDCl_3$) δ: 169.81 (C-7), 52.25 (C-8), 51.55 (C-6), 38.25 (C-3), 28.88 (C-5), 27.52 (C-2), 24.98 (C-4), 22.33 (C-1)

HR-MS: *m/z* calculated for $C_{11}H_{21}O_4$ $[M+H]^+$: 217.1434; found: 217.1441.

2-(4-Methylpentyl)malonic acid 120

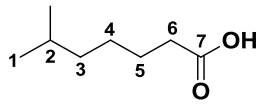


To a solution of KOH (6.40 g, 114.3 mmol) in a mixture of 20 mL water and 40 mL methanol was added dimethyl 2-(4-methylpentyl)malonate (1.35 g, 6.3 mmol) and the resulting mixture stirred at RT overnight. After removal of the methanol under vacuum, the remaining aqueous solution was washed with Et_2O, acidified with 3 M HCl and extracted with DCM. The combined organic layers were dried over $MgSO_4$ and concentrated to yield 2-(4-methylpentyl)malonic acid (1.17 g, 100%) as white crystals.

^{1}H NMR (300 MHz, $CDCl_3$) δ: 0.87 (d, 6H, H-1, J = 6.5 Hz), 1.17 - 1.25 (m, 2H, H-3), 1.34 - 1.44 (m, 2H, H-4), 1.55 (nonet, 1H, H-2, J = 6.5 Hz), 1.92 (quartet, 2H, H-5, J = 8.0 Hz), 3.44 (t, 1H, H-6, J = 7.5 Hz), 10.73 (br s, 2H, COO-H)

LR-MS: *m/z* calculated for $C_9H_{15}O_4$ $[M-H]^-$ found: 187.1.

6-Methylheptanoic acid 121



A solution of 2-(4-methylpentyl)malonic acid (2.93 g, 15.6 mmol) in 30 mL DMF was refluxed overnight. Then 30 mL of 5% $NaHCO_3$ solution was added and the mixture washed with hexane, acidified with 3 M HCl and extracted with DCM. The combined DCM layers were washed with brine, dried over $MgSO_4$ and concentrated under vacuum to give 6-methylheptanoic acid (2.20 g, 98%) as a yellowish oil.

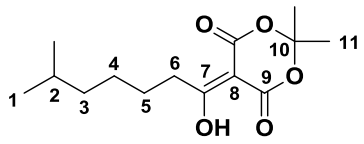
^{1}H NMR (300 MHz, $CDCl_3$) δ: 0.85 (d, 6H, H-1, J = 7.0 Hz), 1.14 - 1.21 (m, 2H, H-3), 1.27 - 1.38 (m, 2H, H-4), 1.52 (nonet, 1H, H-2, J = 6.5 Hz), 1.60 (quintet, 2H, H-5, J = 7.5 Hz), 2.34 (t, 2H, H-6, J = 7.5 Hz), 11.73 (br s, 1H, COO-H)

^{13}C NMR (75 MHz, $CDCl_3$) δ: 180.68 (C-7), 38.49 (C-3), 34.14 (C-6), 27.76 (C-2), 26.81 (C-4), 24.86 (C-5), 22.49 (C-1)

HR-MS: *m/z* calculated for $C_8H_{17}O_2$ $[M+H]^+$: 145.1223; found: 145.1222.

6.1.6 Synthesis of SCB1

5-(1-Hydroxy-6-methylheptylidene)-2,2-dimethyl-1,3-dioxane-4,6-dione 122



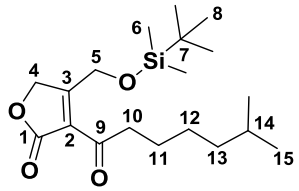
Following the general procedure for acyl Meldrum's acid synthesis, work up gave 5-(1-hydroxy-6-methylheptylidene)-2,2-dimethyl-1,3-dioxane-4,6-dione (2.05 g, 48%) as a yellow oil.

^{1}H NMR (300 MHz, $CDCl_3$) δ: 0.79 (d, 6H, H-1, J = 6.5 Hz), 1.12 - 1.17 (m, 2H, H-3), 1.28 - 1.38 (m, 2H, H-4), 1.47 (nonet, 1H, H-2, J = 6.5 Hz), 1.61 (quintet, 2H, H-5, J = 7.5 Hz), 1.66 (s, 6H, H-11), 2.99 (t, 2H, H-6, J = 7.5 Hz), 15.23 (br s, 1H, O-H)

^{13}C NMR (75 MHz, $CDCl_3$) δ: 170.25 (C-7), 159.79 (C-9), 104.38 (C-10), 90.96 (C-8), 38.16 (C-3), 35.42 (C-6), 27.49 (C-2), 26.86 (C-4), 26.42 (C-11), 26.06 (C-5), 22.23 (C-1)

LR-MS: *m/z* calculated for $C_{14}H_{21}O_5$ $[M-H]^-$ found: 269.2.

4-(((*tert*-Butyldimethylsilyl)oxy)methyl)-3-(6-methylheptanoyl)furan-2(5*H*)-one 123



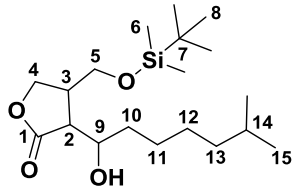
Following the general procedure for butenolide synthesis, work up gave 4-(((*tert*-butyldimethylsilyl)oxy)methyl)-3-(6-methylheptanoyl)furan-2(5*H*)-one (172 mg, 33%) as a yellow oil.

^{1}H NMR (300 MHz, $CDCl_3$) δ: 0.09 (s, 6H, H-6), 0.84 (d, 6H, H-15, J = 6.5 Hz), 0.89 (s, 9H, H-8), 1.10 - 1.36 (m, 4H, H-12 and 13), 1.47 - 1.63 (m, 3H, H-11 and 14), 2.96 (t, 2H, H-10, J = 7.5 Hz), 4.98 (s, 2H, H-5), 5.06 (s, 2H, H-4)

^{13}C NMR (75 MHz, $CDCl_3$) δ: 197.08 (C-9), 181.23 (C-1), 170.64 (C-3), 122.41 (C-2), 70.19 (C-4), 61.83 (C-5), 41.64 (C-10), 38.69 (C-13), 29.66 (C-7), 27.81 (C-14), 26.85 (C-12), 25.65 (C-8), 23.35 (C-11), 22.55 (C-15), -5.68 (C-6)

HR-MS: *m/z* calculated for $C_{19}H_{35}O_4Si$ $[M+H]^+$: 355.2299; found: 355.2306.

4-(((*tert*-Butyldimethylsilyl)oxy)methyl)-3-(1-hydroxy-6-methylheptyl) dihydrofuran-2(3*H*)-one 124



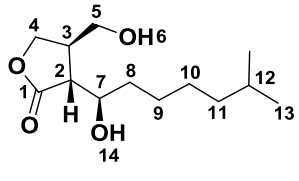
Following the general procedure for butenolide reduction, work up gave 4-(((*tert*-butyldimethylsilyl)oxy)methyl)-3-(1-hydroxy-6-methylheptyl)dihydrofuran-2(3*H*)-one (85 mg, 49%) as a colourless oil.

HR-MS: *m/z* calculated for $C_{19}H_{38}NaO_4Si$ $[M+Na]^+$: 381.2432; found: 381.2439.

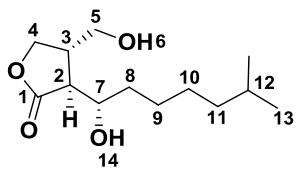
SCB1 stereoisomers 125

Following the general silyl deprotection procedure, work up and column chromatography gave a pair of SCB1 and its enantiomer (25 mg, 43%, higher polarity) and a pair of their diastereoisomers (17 mg, 29%, lower polarity) as colourless oils.

(3*R*,4*R*)-3-((*R*)-1-Hydroxy-6-methylheptyl)-4-(hydroxymethyl)dihydrofuran-2(3*H*)-one, SCB1
(3*S*,4*S*)-3-((*S*)-1-Hydroxy-6-methylheptyl)-4-(hydroxymethyl)dihydrofuran-2(3*H*)-one, SCB1 enantiomer



SCB1



SCB1 enantiomer

^{1}H NMR (400 MHz, $CDCl_3$) δ: 0.86 (d, 6H, H-13, J = 6.5 Hz), 1.14 - 1.20 (m, 2H, H-11), 1.26 - 1.39 (m, 4H, H-9 and 10), 1.47 - 1.55 (m, 2H, H-8), 1.57 - 1.64 (m, 1H, H-12),

2.65 (dd, 1H, H-2, *J* = 9.5 and 4.5 Hz), 2.74 - 2.81 (m, 1H, H-3), 3.68 (dd, 1H, H-5a, *J* = 10.5 and 6.5 Hz), 3.75 (dd, 1H, H-5b, *J* = 10.5 and 5.0 Hz), 3.98 (t, 1H, H-4a, *J* = 9.0 Hz), 4.01 (ddd, 1H, H-7, *J* = 9.0, 4.5 and 3.0 Hz), 4.42 (t, 1H, H-4b, *J* = 8.5 Hz), O-H signal missing

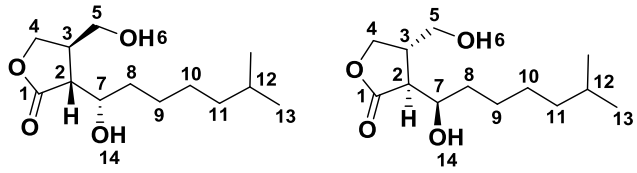
^{13}C NMR (100 MHz, $CDCl_3$) δ: 70.82 (C-7), 68.32 (C-4), 62.98 (C-5), 49.14 (C-2), 40.08 (C-3), 38.82 (C-11), 33.94 (C-8), 27.88 (C-12), 27.18 (C-10), 26.08 (C-9), 22.60 (C-13), C-1 signal missing

HR-MS: *m/z* calculated for $C_{13}H_{24}NaO_4$ $[M+Na]^+$: 267.1567; found: 267.1565.

SCB1 diastereoisomers

(3*R*,4*R*)-3-((*S*)-1-Hydroxy-6-methylheptyl)-4-(hydroxymethyl)dihydrofuran-2(3*H*)-one

(3*S*,4*S*)-3-((*R*)-1-Hydroxy-6-methylheptyl)-4-(hydroxymethyl)dihydrofuran-2(3*H*)-one



^{1}H NMR (300 MHz, $CDCl_3$) δ: 0.86 (d, 6H, H-13, *J* = 6.5 Hz), 1.14 - 1.21 (m, 2H, H-11), 1.28 - 1.38 (m, 4H, H-9 and 10), 1.47 - 1.62 (m, 3H, H-8 and 12), 1.98 (br s, 2H, H-6 and 14), 2.57 (dd, 1H, H-2, *J* = 7.5 and 4.0 Hz), 2.80 - 2.92 (m, 1H, H-3), 3.71 (dd, 1H, H-5a, *J* = 10.0 and 4.5 Hz), 3.75 (dd, 1H, H-5b, *J* = 10.0 and 5.0 Hz), 4.08 - 4.15 (m, 2H, H-4a and 7), 4.42 (t, 1H, H-4b, *J* = 8.5 Hz)

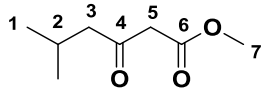
^{13}C NMR (75 MHz, $CDCl_3$) δ: 178.50 (C-1), 70.80 (C-7), 69.42 (C-4), 63.32 (C-5), 48.10 (C-2), 38.85 (C-11), 38.09 (C-3), 34.85 (C-8), 27.89 (C-12), 27.17 (C-10), 26.07 (C-9), 22.59 (C-13)

HR-MS: *m/z* calculated for $C_{13}H_{24}NaO_4$ $[M+Na]^+$: 267.1567; found: 267.1565.

6.1.7 Synthesis of AHFCAs

6.1.7.1 AHFCA1

Methyl 5-methyl-3-oxohexanoate 110



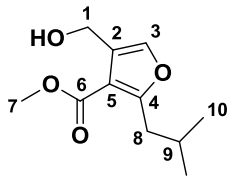
Following the general procedure for β-ketoester synthesis, work up gave methyl 5-methyl-3-oxohexanoate (1.59 g, 82%) as a yellow oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ: 0.78 (d, 6H, H-1, J = 6.5 Hz), 1.97 - 2.04 (m, 1H, H-2), 2.28 (d, 2H, H-3, J = 7.0 Hz), 3.30 (s, 2H, H-5), 3.58 (S, 3H, H-7)

^{13}C NMR (100 MHz, $CDCl_3$) δ: 202.08 (C-4), 167.33 (C-6), 51.84 (C-7), 51.44 (C-3), 49.01 (C-5), 23.94 (C-2), 22.02 (C-1)

HR-MS: *m/z* calculated for $C_8H_{14}NaO_3$ $[M+Na]^+$: 181.0835, found: 181.0831.

Methyl 4-(hydroxymethyl)-2-isobutylfuran-3-carboxylate 113



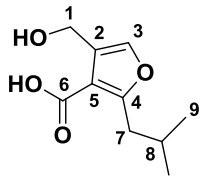
Following the general procedure for AHFCA cyclisation, work up gave methyl 4-(hydroxymethyl)-2-isobutylfuran-3-carboxylate (0.17 g, 63%) as a colourless oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ: 0.90 (d, 6H, H-10, J = 6.5 Hz), 2.02 (nonet, 1H, H-9, J = 7.0 Hz), 2.79 (d, 2H, H-8, J = 7.0 Hz), 3.68 (br s, 1H, O-H), 3.85 (s, 3H, H-7), 4.53 (s, 2H, H-1), 7.23 (s, 1H, H-3)

^{13}C NMR (75 MHz, $CDCl_3$) δ: 165.73 (4), 164.17 (C-6), 138.42 (C-3), 125.71 (C-2), 112.69 (C-5), 55.74 (C-1), 51.64 (C-7), 36.93 (C-8), 28.31 (C-9), 22.37 (C-10)

HR-MS: *m/z* calculated for $C_{11}H_{16}NaO_4$ $[M+Na]^+$: 235.0941, found: 235.0941.

4-(Hydroxymethyl)-2-isobutylfuran-3-carboxylic acid, AHFCA1 26



Following the general procedure for AHFCA methyl ester hydrolysis, work up gave 4-(hydroxymethyl)-2-isobutylfuran-3-carboxylic acid (0.11 g, 70%) as a yellowish solid.

^{1}H NMR (400 MHz, $CDCl_3$) δ: 0.93 (d, 6H, H-9, J = 6.5 Hz), 2.09 (nonet, 1H, H-8, J = 7.0 Hz), 2.88 (d, 2H, H-7, J = 7.0 Hz), 4.61 (s, 2H, H-1), 7.28 (s, 1H, H-3), 7.80 (br s, 2H, O-H and COO-H)

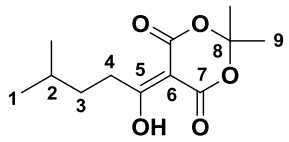
^{13}C NMR (100 MHz, $CDCl_3$) δ: 170.15 (C-4), 166.17 (C-6), 138.79 (C-3), 125.27 (C-2), 112.15 (C-5), 55.58 (C-1), 36.79 (C-7), 28.25 (C-8), 22.35 (C-9)

HR-MS: *m/z* calculated for $C_{10}H_{14}NaO_4$ $[M+Na]^+$: 221.0784, found: 221.0786.

Melting point: 91-93 ℃.

6.1.7.2 AHFCA3

5-(1-Hydroxy-4-methylpentylidene)-2,2-dimethyl-1,3-dioxane-4,6-dione 111



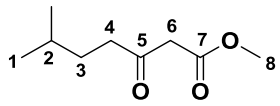
Following the general procedure for acyl Meldrum's acid synthesis, work up gave 5-(1-hydroxy-4-methylpentylidene)-2,2-dimethyl-1,3-dioxane-4,6-dione (1.05 g, 84%) as a yellow oil.

^{1}H NMR (300 MHz, $CDCl_3$) δ: 0.91 (d, 6H, H-1, J = 6.0 Hz), 1.48 - 1.58 (m, 2H, H-3), 1.64 (nonet, 1H, H-2, J = 6.5 Hz), 1.70 (s, 6H, H-9), 3.03 (t, 2H, H-4, J = 7.5 Hz), 15.26 (br s, 1H, O-H)

^{13}C NMR (75 MHz, $CDCl_3$) δ: 198.54 (C-5), 170.49 (C-7), 104.66 (C-8), 91.07 (C-6), 34.83 (C-3), 33.83 (C-4), 28.04 (C-2), 26.69 (C-9), 22.11 (C-1)

LR-MS: *m/z* calculated for $C_{12}H_{17}O_5$ $[M-H]^-$ found: 241.2.

Methyl 6-methyl-3-oxoheptanoate 112



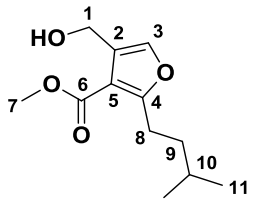
Following the general procedure for β-ketoester synthesis, work up gave methyl 6-methyl-3-oxoheptanoate (0.71 g, 95%) as a yellow oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ: 0.77 (d, 6H, H-1, J = 6.5 Hz), 1.35 - 1.39 (m, 2H, H-3), 1.44 (nonet, 1H, H-2, J = 6.5 Hz), 2.43 (t, 2H, H-4, J = 7.5 Hz), 3.36 (s, 2H, H-6), 3.61 (s, 3H, H-8)

^{13}C NMR (100 MHz, $CDCl_3$) δ: 202.69 (C-5), 167.44 (C-7), 51.93 (C-8), 48.64 (C-6), 40.77 (C-4), 31.88 (C-3), 27.21 (C-2), 21.99 (C-1)

HR-MS: *m/z* calculated for $C_9H_{16}NaO_3$ $[M+Na]^+$: 195.0992, found: 195.0998.

Methyl 4-(hydroxymethyl)-2-isopentylfuran-3-carboxylate 114



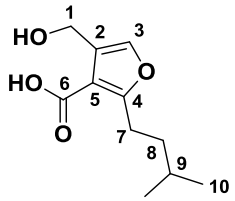
Following the general procedure for AHFCA cyclisation, work up gave methyl 4-(hydroxymethyl)-2-isopentylfuran-3-carboxylate (0.19 g, 57%) as a yellow oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ: 0.89 (d, 6H, H-11, J = 5.5 Hz), 1.46 - 1.54 (m, 3H, H-9 and 10), 2.90 (t, 2H, H-8, J = 7.5 Hz), 3.70 (br s, 1H, O-H), 3.84 (s, 3H, H-7), 4.51 (s, 2H, H-1), 7.20 (s, 1H, H-3)

^{13}C NMR (75 MHz, $CDCl_3$) δ: 165.68 (C-4), 165.10 (C-6), 138.25 (C-3), 125.78 (C-2), 111.87 (C-5), 55.71 (C-1), 51.67 (C-7), 36.92 (C-8), 27.67 (C-10), 26.22 (C-9), 22.23 (C-11)

HR-MS: *m/z* calculated for $C_{12}H_{18}NaO_4$ $[M+Na]^+$: 249.1097, found: 249.1095.

4-(Hydroxymethyl)-2-isopentylfuran-3-carboxylic acid, AHFCA3 42



Following the general procedure for AHFCA methyl ester hydrolysis, work up gave 4-(hydroxymethyl)-2-isopentylfuran-3-carboxylic acid (0.11 g, 65%) as a yellowish solid.

^{1}H NMR (400 MHz, $CDCl_3$) δ: 0.85 (d, 6H, H-10, J = 5.5 Hz), 1.47 - 1.50 (m, 3H, H-8 and 9), 2.92 (t, 2H, H-7, J = 7.0 Hz), 4.53 (s, 2H, H-1), 7.20 (s, 1H, H-3), 7.55 (br s, 2H, O-H and COO-H)

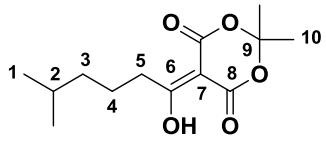
^{13}C NMR (100 MHz, $CDCl_3$) δ: 169.85 (C-4), 167.00 (C-6), 138.60 (C-3), 125.24 (C-2), 111.36 (C-5), 55.51 (C-1), 36.67 (C-7), 27.68 (C-9), 26.19 (C-8), 22.22 (C-10)

HR-MS: *m/z* calculated for $C_{11}H_{16}NaO_4$ $[M+Na]^+$: 235.0941, found: 235.0948.

Melting point: 75-76 ℃.

6.1.7.3 AHFCA6

5-(1-Hydroxy-5-methylhexylidene)-2,2-dimethyl-1,3-dioxane-4,6-dione 126



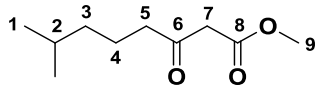
Following the general procedure for acyl Meldrum's acid synthesis, work up gave 5-(1-hydroxy-5-methylhexylidene)-2,2-dimethyl-1,3-dioxane-4,6-dione (1.35 g, 53%) as a yellow oil.

^{1}H NMR (250 MHz, $CDCl_3$) δ: 0.87 (d, 6H, H-1, J = 6.5 Hz), 1.23 - 1.32 (m, 2H, H-3), 1.57 (nonet, 1H, H-2, J = 6.5 Hz), 1.62 - 1.69 (m, 2H, H-4), 1.71 (s, 6H, H-10), 3.03 (t, 2H, H-5, J = 7.5 Hz), 15.28 (br s, 1H, O-H)

^{13}C NMR (75 MHz, $CDCl_3$) δ: 170.51 (C-6), 160.11 (C-8), 104.68 (C-9), 91.18 (C-7), 38.40 (C-3), 35.81 (C-5), 27.64 (C-2), 26.72 (C-10), 23.96 (C-4), 22.36 (C-1)

HR-MS: *m/z* calculated for $C_{13}H_{20}NaO_5$ $[M+Na]^+$: 279.1203, found: 279.1217.

Methyl 7-methyl-3-oxooctanoate 127



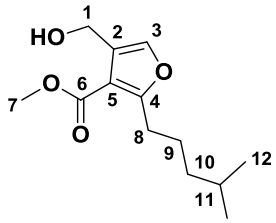
Following the general procedure for β-ketoester synthesis, work up gave methyl 7-methyl-3-oxooctanoate (0.93 g, 95%) as a yellow oil.

^{1}H NMR (300 MHz, $CDCl_3$) δ: 0.73 (d, 6H, H-1, J = 6.5 Hz), 0.99 - 1.13 (m, 2H, H-3), 1.33 - 1.50 (m, 3H, H-2 and 4), 2.39 (t, 2H, H-5, J = 7.5 Hz), 3.32 (s, 2H, H-7), 3.58 (s, 3H, H-9)

^{13}C NMR (75 MHz, $CDCl_3$) δ: 202.41 (C-6), 167.34 (C-8), 51.81 (C-9), 48.59 (C-7), 42.82 (C-5), 37.85 (C-3), 27.47 (C-2), 22.07 (C-1), 20.95 (C-4)

HR-MS: *m/z* calculated for $C_{10}H_{18}NaO_3$ $[M+Na]^+$: 209.1148, found: 209.1149.

Methyl 4-(hydroxymethyl)-2-(4-methylpentyl)furan-3-carboxylate 129



Following the general procedure for AHFCA cyclisation, work up gave methyl 4-(hydroxymethyl)-2-(4-methylpentyl)furan-3-carboxylate (0.24 g, 85%) as a colourless oil.

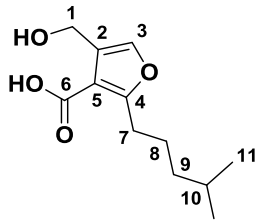
^{1}H NMR (300 MHz, $CDCl_3$) δ: 0.85 (d, 6H, H-12, J = 6.5 Hz), 1.15 - 1.26 (m, 2H, H-10), 1.48 - 1.68 (m, 3H, H-9 and 11), 2.89 (t, 2H, H-8, J = 7.5 Hz), 3.57 (br s, 1H, O-H), 3.85 (s, 3H, H-7), 4.54 (s, 2H, H-1), 7.22 (s, 1H, H-3)

^{13}C NMR (75 MHz, $CDCl_3$) δ: 165.66 (C-4), 164.85 (C-6), 138.25 (C-3), 125.73 (C-2), 111.93 (C-5), 55.69 (C-1), 51.64 (C-7), 38.35 (C-10), 28.35 (C-8), 27.61 (C-11), 25.76

(C-9), 22.42 (C-12)

HR-MS: *m/z* calculated for $C_{13}H_{20}NaO_4$ $[M+Na]^+$: 263.1254; found: 263.1256.

4-(Hydroxymethyl)-2-(4-methylpentyl)furan-3-carboxylic acid, AHFCA6 45



Following the general procedure for AHFCA methyl ester hydrolysis, work up gave 4-(hydroxymethyl)-2-(4-methylpentyl)furan-3-carboxylic acid (0.17 g, 95%) as a colourless solid.

^{1}H NMR (400 MHz, $CDCl_3$) δ: 0.88 (d, 6H, H-11, J = 6.5 Hz), 1.23 (quartet, 2H, H-9, J = 7.5 Hz), 1.57 (nonet, 1H, H-10, J = 6.5 Hz), 1.69 (quintet, 2H, H-8, J = 8.0 Hz), 2.98 (t, 2H, H-7, J = 7.5 Hz), 4.60 (s, 2H, H-1), 7.28 (s, 1H, H-3), O-H and COO-H proton signal missing

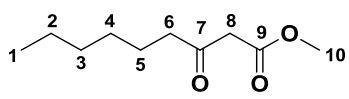
^{13}C NMR (100 MHz, $CDCl_3$) δ: 169.97 (C-4), 167.14 (C-6), 138.70 (C-3), 125.45 (C-2), 111.30 (C-5), 55.63 (C-1), 38.38 (C-9), 28.43 (C-7), 27.72 (C-10), 25.75 (C-8), 22.48 (C-11)

HR-MS: *m/z* calculated for $C_{12}H_{18}NaO_4$ $[M+Na]^+$: 249.1097; found: 249.1108.

Melting point: 80 - 82 ℃.

6.1.7.4 AHFCA7

Methyl 3-oxononanoate 128



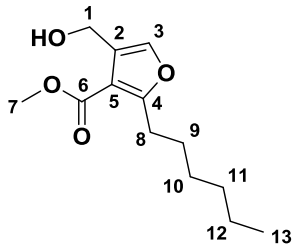
Following the general procedure for β-ketoester synthesis, work up gave methyl 3-oxononanoate (0.35 g, 80%) as a colourless oil.

^{1}H NMR (300 MHz, $CDCl_3$) δ: 0.84 (t, 3H, H-1, J = 6.5 Hz), 1.23 - 1.30 (m, 6H, H-2, 3 and 4), 1.55 (quintet, 2H, H-5, J = 7.0 Hz), 2.50 (t, 2H, H-6, J = 7.5 Hz), 3.42 (s, 2H, H-8), 3.70 (s, 3H, H-10)

^{13}C NMR (75 MHz, $CDCl_3$) δ: 202.79 (C-7), 167.62 (C-9), 52.20 (C-10), 48.90 (C-8), 42.97 (C-6), 31.43 (C-3), 28.56 (C-4), 23.31 (C-5), 22.36 (C-2), 13.90 (C-1)

HR-MS: *m/z* calculated for $C_{10}H_{18}NaO_3$ $[M+Na]^+$: 209.1148, found: 209.1151.

Methyl 2-hexyl-4-(hydroxymethyl)furan-3-carboxylate 130



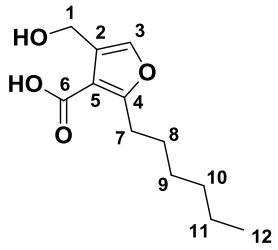
Following the general procedure for AHFCA cyclisation, work up gave methyl 2-hexyl-4-(hydroxymethyl)furan-3-carboxylate (0.25 g, 74%) as a yellow oil.

^{1}H NMR (300 MHz, $CDCl_3$) δ: 0.87 (t, 3H, H-13, J = 6.5 Hz), 1.24 - 1.33 (m, 6H, H-10, 11 and 12), 1.63 (quintet, 2H, H-9, J = 7.5 Hz), 2.92 (t, 2H, H-8, J = 7.5 Hz), 3.51 (br s, 1H, O-H), 3.87 (s, 3H, H-7), 4.54 (s, 2H, H-1), 7.23 (s, 1H, H-3)

^{13}C NMR (75 MHz, $CDCl_3$) δ: 165.68 (C-4), 164.90 (C-6), 138.25 (C-3), 125.73 (C-2), 111.94 (C-5), 55.69 (C-1), 51.64 (C-7), 31.37 (C-8), 28.79 (C-11), 28.15 (C-9), 27.85 (C-10), 22.44 (C-12), 13.97 (C-13)

HR-MS: *m/z* calculated for $C_{13}H_{20}NaO_4$ $[M+Na]^+$: 263.1254; found: 263.1260.

2-Hexyl-4-(hydroxymethyl)furan-3-carboxylic acid, AHFCA7 46



Following the general procedure for AHFCA methyl ester hydrolysis, work up gave 2-hexyl-4-(hydroxymethyl)furan-3-carboxylic acid (0.22 g, 95%) as a yellowish solid.

^{1}H NMR (400 MHz, $CDCl_3$) δ: 0.88 (t, 3H, H-12, J = 6.5 Hz), 1.25 - 1.36 (m, 6H, H-9, 10 and 11), 1.68 (quintet, 2H, H-8, J = 7.0 Hz), 3.00 (t, 2H, H-7, J = 7.5 Hz), 4.60 (s, 2H, H-1), 7.28 (s, 1H, H-3), O-H and COO-H proton signal missing

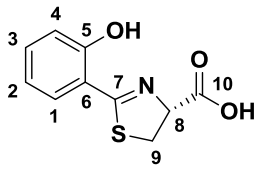
^{13}C NMR (100 MHz, $CDCl_3$) δ: 170.17 (C-4), 167.17 (C-6), 138.68 (C-3), 125.43 (C-2), 111.33 (C-5), 55.61 (C-1), 31.44 (C-7), 28.87 (C-10), 28.23 (C-8), 27.82 (C-9), 22.49 (C-11), 14.04 (C-12)

HR-MS: *m/z* calculated for $C_{12}H_{18}NaO_4$ $[M+Na]^+$: 249.1097; found: 249.1108.

Melting point: 80-81 ℃.

6.1.8 Synthesis of pyochelin and isopyochelin

(*R*)-2-(2-Hydroxyphenyl)-4,5-dihydrothiazole-4-carboxylic acid 131



A solution of 2-hydroxy-2-benzonitrile (1.90 g, 16.0 mmol) and L-cysteine (3.86 g, 31.9 mmol) in a 1:1 v/v mixture of methanol and phosphate buffer pH 6.4 (75 mL) was stirred under argon at 50 ℃ for 4 days. The insoluble materials were removed by filtration and the filtrate concentrated under vacuum. To the aqueous phase was added 20 mL water and 10 mL 1 M HCl, and the mixture extracted with DCM. The combined

organic layers were washed with water and brine, dried over $MgSO_4$ and concentrated under vacuum. The residue was recrystallised using a mixture of ethanol and hexane to yield (*R*)-2-(2-hydroxyphenyl)-4,5-dihydrothiazole-4-carboxylic acid (2.85 g, 80%) as a yellow solid.

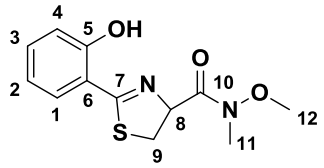
^{1}H NMR (300 MHz, $CDCl_3$): δ 3.61 - 3.73 (m, 2H, H-9), 5.42 (dd, 1H, H-8, J = 9.5 and 8.0 Hz), 6.86 - 6.92 (m, 1H, H-4), 7.03 (dd, 1H, H-2, J = 8.5 and 0.5 Hz), 7.35 - 7.39 (m, 1H, H-3), 7.43 (dd, 1H, H-1, J = 8.0 and 1.5 Hz), 10.79 (br s, 2H, O-H and COO-H)

^{13}C NMR (75 MHz, DMSO-d_6): δ 172.71 (C-10), 171.29 (C-7), 158.38 (C-5), 133.72 (C-3), 130.53 (C-1), 119.34 (C-2), 116.87 (C-4), 115.60 (C-6), 76.32 (C-8), 33.42 (C-9)

HR-MS: *m/z* calculated for $C_{10}H_{10}NO_3S$ $[M+H]^+$: 224.0381, found: 224.0378.

Melting point: 114-116 ℃.

2-(2-Hydroxyphenyl)-*N*-methoxy-*N*-methyl-4,5-dihydrothiazole-4-carboxamide 134



To a solution of (*R*)-2-(2-hydroxyphenyl)-4,5-dihydrothiazole-4-carboxylic acid (2.85 g, 12.8 mmol) in 30 mL THF was added CDMT (2.70 g, 15.4 mmol) and *N*-methylmorpholine (4.2 mL, 38.2 mmol) at 0 ℃ and the mixture stirred for 1 hour, followed by addition of *N,O*-dimethylhydroxylamine hydrochloride (1.25 g, 12.8 mmol). The reaction mixture was then stirred at RT overnight and quenched with 50 mL water. After removal of THF under vacuum, the remaining aqueous solution was acidified to pH 4 and extracted with DCM. The combined organic layers were washed with water and brine, dried over $MgSO_4$ and concentrated under vacuum. The residue was purified by column chromatography (pet ether/EtOAc = 2:1 v/v) to yield 2-(2-hydroxyphenyl)-*N*-methoxy-*N*-methyl-4,5-dihydrothiazole-4-carboxamide (1.43 g, 42%) as a yellow oil.

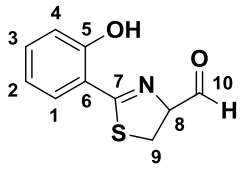
^{1}H NMR (300 MHz, $CDCl_3$): δ 3.27 (s, 3H, H-11), 3.46 (dd, 1H, H-9a, J = 11.0 and 9.5 Hz), 3.75 (t, 1H, H-9b, J = 10.0 Hz), 3.81 (s, 3H, H-12), 5.68 (t, 1H, H-8, J= 9.0 Hz),

6.86 (td, 1H, H-4, *J* = 7.5 and 1.0 Hz), 6.97 (dd, 1H, H-2, *J* = 8.5 and 1.0 Hz), 7.31 - 7.37 (m, 1H, H-3), 7.41 (dd, 1H, H-1, *J* =8.0 and 1.5 Hz), 12.32 (br s, 1H, O-H)

^{13}C NMR (75 MHz, $CDCl_3$): δ 173.98 (C-10), 169.62 (C-7), 158.87 (C-5), 133.25 (C-3), 130.71 (C-1), 118.86 (C-2), 116.96 (C-4), 116.03 (C-6), 74.51 (C-8), 61.72 (C-12), 32.76 (C-9), 32.45 (C-11)

HR-MS: *m/z* calculated for $C_{12}H_{14}N_2NaO_3S$ $[M+Na]^+$: 289.0617, found: 289.0618.

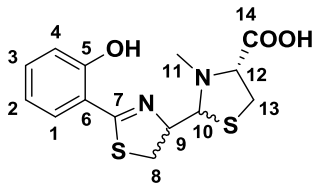
2-(2-Hydroxyphenyl)-4,5-dihydrothiazole-4-carbaldehyde 132



To a solution of 2-(2-hydroxyphenyl)-*N*-methoxy-*N*-methyl-4,5-dihydrothiazole-4-carboxamide (0.45 g, 1.7 mmol) in 20 mL dry Et_2O at -20 ℃ was added $LiAlH_4$ (0.19 g, 5.0 mmol) and the reaction mixture stirred at the same temperature for 20 mins or until the starting material had disappeared. Then 2 mL methanol, 20 mL saturated NH_4Cl solution and 10 mL 5% (v/v) H_2SO_4 solution were successively added. The organic phase was conserved and the aqueous phase extracted with DCM. The combined organic layers were washed with brine, dried over $MgSO_4$ and concentrated under vacuum to yield the crude 2-(2-hydroxyphenyl)-4,5-dihydrothiazole-4-carbaldehyde (0.21 g, 60%) as a yellow oil. The aldehyde was unstable and therefore immediately used in the next step without further purification.

HR-MS: *m/z* calculated for $C_{10}H_{10}NO_2S$ $[M+H]^+$: 208.0427, found: 208.0432.

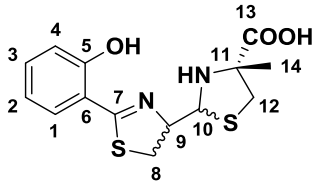
Pyochelin 74



Following the general condensation procedure with *N*-methyl-L-cysteine hydrochloride, work up gave a mixture of pyochelin stereoisomers (29 mg, 74%) as a yellow solid.

HR-MS: *m/z* calculated for $C_{14}H_{17}N_2O_3S_2$ $[M+H]^+$: 325.0675, found: 325.0681.

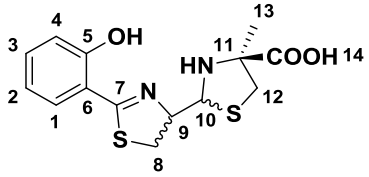
4"*R*-Isopyochelin 135



Following the general condensation procedure with 2-methyl-L-cysteine hydrochloride, work up gave a mixture of 4"*R*-isopyochelin stereoisomers (25 mg, 64%) as a yellow solid.

HR-MS: *m/z* calculated for $C_{14}H_{17}N_2O_3S_2$ $[M+H]^+$: 325.0675, found: 325.0677.

4"*S*-Isopyochelin 136

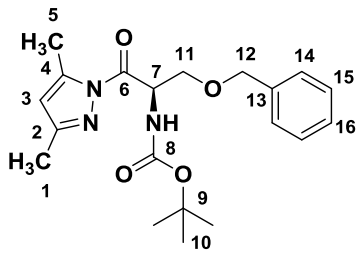


Following the general condensation procedure with 2-methyl-D-cysteine hydrochloride, work up gave a mixture of 4"*S*-isopyochelin stereoisomers (31 mg, 79%) as a yellow solid.

HR-MS: *m/z* calculated for $C_{14}H_{17}N_2O_3S_2$ $[M+H]^+$: 325.0675, found: 325.0674.

6.1.9 Synthesis of stereospecifically deuterium-labelled L-cystine

***tert*-Butyl (*R*)-(3-(benzyloxy)-1-(3,5-dimethyl-1H-pyrazol-1-yl)-1-oxopropan-2-yl)carbamate 141**



To a solution of *O*-benzyl-*N*-(*tert*-butoxycarbonyl)-D-serine (7.60 g, 25.8 mmol), 3,5-dimethylpyrazole (2.72 g, 28.3 mmol) and DMAP (0.50 g, 4.1 mmol) in 40 mL DCM was slowly added a solution of DCC (5.84 g, 28.3 mmol) in 10 mL DCM at 0 ℃ and the mixture stirred at the same temperature for 0.5 hour. The reaction was then allowed to warm up to RT and stirred overnight. After addition of 20 mL n-hexane, the precipitated DCU was filtered to be removed. The filtrate was concentrated under vacuum and the residue was purified by column chromatography (pet ether/EtOAc = 9:1 v/v) to yield *tert*-butyl (*R*)-(3-(benzyloxy)-1-(3,5-dimethyl-1H-pyrazol-1-yl)-1-oxopropan-2-yl)carbamate (6.81 g, 71%) as a white solid.

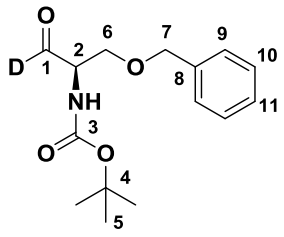
^{1}H NMR (300 MHz, $CDCl_3$) δ: 1.45 (s, 9H, H-10), 2.16 (s, 3H, H-1), 2.53 (s, 3H, H-5), 3.83 (d, 1H, H-11a, J = 10.0 Hz), 4.05 (d, 1H, H-11b, J = 10.5 Hz), 4.44 (d, 1H, H-12a, J = 12.5 Hz), 4.52 (d, 1H, H-12b, J = 12.5 Hz), 5.64 (br s, 2H, H-7 and N-H), 5.93 (s, 1H, H-3), 7.17 - 7.29 (m, 5H, H-14, 15 and 16)

^{13}C NMR (75 MHz, $CDCl_3$) δ: 170.38 (C-6), 155.41 (C-8), 152.35 (C-2), 144.19 (C-4), 137.61 (C-13), 128.18 (C-15), 127.50 (C-16), 127.34 (C-14), 111.12 (C-3), 79.71 (C-9), 72.79 (C-12), 70.31 (C-11), 54.14 (C-7), 28.23 (C-10), 14.12 (C-5), 13.65 (C-1)

HR-MS: *m/z* calculated for $C_{20}H_{28}N_3O_4$ $[M+H]^+$: 374.2074; found: 374.2077.

Melting point: 66-67 ℃.

***tert*-Butyl (*R*)-(1-(benzyloxy)-3-oxopropan-2-yl-3-^{2}H)carbamate 142**



To a solution of $LiAlD_4$ (0.53 g, 12.3 mmol) in 12 mL THF was added a solution of *tert*-butyl (*R*)-(3-(benzyloxy)-1-(3,5-dimethyl-1H-pyrazol-1-yl)-1-oxopropan-2-yl) carbamate (2.30 g, 6.2 mmol) in 60 mL at -40 ℃ under argon and the mixture stirred for 1 hour. The reaction was quenched by adding 10% aqueous citric acid slowly and allowed to warm up to RT. After removal of the insoluble materials by filtration through celite, the filtrate was diluted with Et_2O, washed with aqueous $KHSO_4$ solution, dried

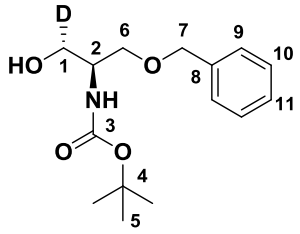
over $MgSO_4$ and concentrated to yield *tert*-butyl (*R*)-(1-(benzyloxy)-3-oxopropan-2-yl-3-^{2}H)carbamate (1.73 g, 100%) as a colourless oil. The crude product was used in the next step without further purification.

^{1}H NMR (400 MHz, $CDCl_3$) δ: 1.44 (s, 9H, H-5), 3.69 (dd, 1H, H-6a, J = 9.5 and 3.5 Hz), 3.95 (dd, 1H, H-6b, J = 9.5 and 3.0 Hz), 4.29 (m, 1H, H-2), 4.45 (d, 1H, H-7a, J = 12.0 Hz), 4.50 (d, 1H, H-7b, J = 12.0 Hz), 5.53 (br d, 1H, N-H, J = 6.5 Hz), 7.24 - 7.34 (m, 5H, H-9, 10 and 11)

^{13}C NMR (100 MHz, $CDCl_3$) δ: 198.63 (t, C-1, J = 27.0 Hz), 155.44 (C-3), 137.13 (C-8), 128.27 (C-10), 127.73 (C-11), 127.47 (C-9), 79.90 (C-4), 73.25 (C-7), 67.60 (C-6), 59.74 (C-2), 28.08 (C-5)

HR-MS: *m/z* calculated for $C_{15}H_{20}DNNaO_4$ $[M+Na]^+$: 303.1426; found: 303.1430.

tert-Butyl ((2*S*,3*R*)-1-(benzyloxy)-3-hydroxypropan-2-yl-3-^{2}H)carbamate 143



To a solution of *tert*-butyl (*R*)-(1-(benzyloxy)-3-oxopropan-2-yl-3-^{2}H)carbamate (2.00 g, 7.1 mmol) in 70 mL THF was added a solution of 0.5 M *S*-Alpine-Borane in THF (14.3 mL, 7.2 mmol) at RT under argon. After being stirred overnight, the reaction was refluxed for 1 hour and cooled to RT. To the mixture was added acetaldehyde (0.8 mL) and, after being stirred for 15 mins, the solvent was removed under vacuum. The residue was dissolved in dry Et_2O and to the solution was added diethanolamine (2.8 mL) at 0 ℃ under argon. After being stirred for 30 mins, the precipitated solids were removed by filtration and the filtrate was washed with water, dried over $MgSO_4$ and concentrated. The crude product was purified by column chromatography (pet ether/EtOAc = 2:1 v/v) to yield *tert*-butyl ((2*S*,3*R*)-1-(benzyloxy)-3-hydroxypropan-2-yl-3-^{2}H)carbamate (1.92 g, 95%) as a colourless oil.

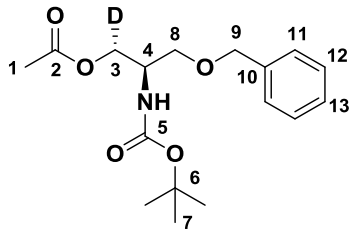
^{1}H NMR (400 MHz, $CDCl_3$) δ: 1.43 (s, 9H, H-5), 3.20 (br s, 1H, O-H), 3.56 (dd, 1H, H-6a, J = 9.5 and 4.5 Hz), 3.61 (dd, 1H, H-6b, J = 9.0 and 4.0 Hz), 3.73 (s, 1H, H-1),

3.78 (br s, 1H, H-2), 4.50 (s, 2H, H-7), 5.25 (s, 1H, N-H), 7.25 - 7.35 (m, 5H, H-9, 10 and 11)

^{13}C NMR (100 MHz, $CDCl_3$) δ: 155.91 (C-3), 137.64 (C-8), 128.28 (C-10), 127.62 (C-11), 127.47 (C-9), 79.39 (C-4), 73.17 (C-7), 69.93 (C-6), 62.64 (t, C-1, J = 22.0 Hz), 51.40 (C-2), 28.21 (C-5)

HR-MS: *m/z* calculated for $C_{15}H_{22}DNNaO_4$ $[M+Na]^+$: 305.1575; found: 305.1582.

(1*R*,2*R*)-3-(Benzyloxy)-2-((*tert*-butoxycarbonyl)amino)propyl-1-^{2}H acetate 144



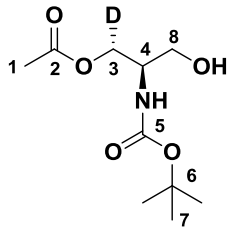
To a solution of *tert*-butyl ((2*S*,3*R*)-1-(benzyloxy)-3-hydroxypropan-2-yl-3-^{2}H)carbamate (2.55 g, 9.0 mmol) in 20 mL DCM was added triethylamine (3.8 mL, 27.1 mmol), acetic anhydride (1.40 g, 13.6 mmol) and DMAP (0.33 g, 2.7 mmol) and the mixture stirred at RT for 5 hours. After removal of the solvent, 30 mL water was added and the mixture extracted with Et_2O. The combined organic phase was washed with aqueous $KHSO_4$ solution, dried over $MgSO_4$ and concentrated under vacuum to yield (1*R*,2*R*)-3-(benzyloxy)-2-((*tert*-butoxycarbonyl)amino)propyl-1-^{2}H acetate (2.90 g, 99%) as a colourless oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ: 1.44 (s, 9H, H-7), 2.00 (s, 3H, H-1), 3.49 (dd, 1H, H-8a, J = 9.5 and 5.0 Hz), 3.55 (dd, 1H, H-8b, J = 9.5 and 3.0 Hz), 4.03 (br s, 1H, H-4), 4.17 (d, 1H, H-3, J = 6.0 Hz), 4.48 (d, 1H, H-9a, J = 12.0 Hz), 4.52 (d, 1H, H-9b, J = 12.0 Hz), 5.06 (br d, 1H, N-H, J = 8.0 Hz), 7.27 - 7.35 (m, 5H, H-11, 12 and 13)

^{13}C NMR (100 MHz, $CDCl_3$) δ: 170.64 (C-2), 155.25 (C-5), 137.55 (C-10), 128.16 (C-12), 127.53 (C-13), 127.43 (C-11), 79.36 (C-6), 72.95 (C-9), 68.50 (C-8), 63.02 (t, C-3, J = 22.5 Hz), 48.77 (C-4), 28.07 (C-7), 20.50 (C-1)

HR-MS: *m/z* calculated for $C_{17}H_{24}DNNaO_5$ $[M+Na]^+$: 347.1688; found: 347.1700.

(1*R*,2*R*)-2-((*tert*-Butoxycarbonyl)amino)-3-hydroxypropyl-1-^{2}H acetate 145



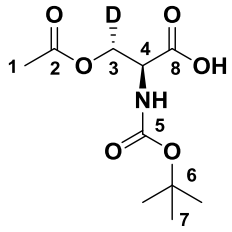
A mixture of (1*R*,2*R*)-3-(benzyloxy)-2-((*tert*-butoxycarbonyl)amino)propyl-1-^{2}H acetate (6.30 g, 19.4 mmol) and 10% palladium on carbon (1.80 g) in 120 mL methanol was hydrogenated at RT for 2 hours. The catalyst was removed by filtration through celite and the filtrate was concentrated to give (1*R*,2*R*)-2-((*tert*-butoxycarbonyl)amino)-3-hydroxypropyl-1-^{2}H acetate (4.44 g, 98%) as a colourless oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ: 1.44 (s, 9H, H-7), 2.08 (s, 3H, H-1), 3.63 (dd, 1H, H-8a, J =11.5 and 5.0 Hz), 3.68 (dd, 1H, H-8b, J = 11.5 and 4.5 Hz), 3.88 (br s, 1H, H-4), 4.15 (d, 1H, H-3, J = 5.5 Hz), 4.57 (br s, 1H, O-H), 5.49 (br d, 1H, N-H, J = 7.0 Hz)

^{13}C NMR (100 MHz, $CDCl_3$) δ: 171.07 (C-2), 155.72 (C-5), 79.33 (C-6), 62.68 (t, C-3, J = 22.5 Hz), 61.00 (C-8), 50.56 (C-4), 27.91 (C-7), 20.36 (C-1)

HR-MS: *m/z* calculated for $C_{10}H_{18}DNNaO_5$ $[M+Na]^+$: 257.1218; found: 257.1215.

(3*R*)-*O*-Acetyl-*N*-(*tert*-butoxycarbonyl)-L-serine-3-^{2}H 146



To a suspension of $NaIO_4$ (48.3 g, 225.7 mmol) and $RuCl_3 \cdot nH_2O$ (1.40 g) in 115 mL water in an ice bath was slowly added a solution of (1*R*,2*R*)-2-((*tert*-butoxycarbonyl)amino)-3-hydroxypropyl-1-^{2}H acetate (4.44 g, 19.0 mmol) in 115 mL acetone. Then the reaction mixture was allowed to stir at RT for 1.5 hours. To the reaction was added isopropanol (60 mL) and the mixture stirred in an ice bath for 1 hour. The precipitated RuO_2 was removed by filtration through celite and the filtrate concentrated under vacuum to remove the acetone and extracted with chloroform. The

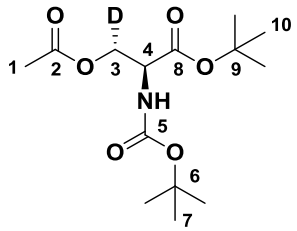
combined chloroform layers were washed with brine, dried over $MgSO_4$ and concentrated to yield (3*R*)-*O*-acetyl-*N*-(*tert*-butoxycarbonyl)-L-serine-3-^{2}H (4.70 g, 100%) as a colourless oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ: 1.46 (s, 9H, H-7), 2.08 (s, 3H, H-1), 4.43 (d, 1H, H-3, J =3.5 Hz), 4.62 (dd, 1H, H-4, J = 8.0 and 4.0 Hz), 5.56 (br d, 1H, N-H, J = 8.0 Hz), 11.00 (s, 1H, O-H)

^{13}C NMR (100 MHz, $CDCl_3$) δ: 172.96 (C-8), 170.80 (C-2), 155.35 (C-5), 80.38 (C-6), 63.83 (t, C-3, J = 22.0 Hz), 52.43 (C-4), 28.01 (C-7), 20.39 (C-1)

HR-MS: *m/z* calculated for $C_{10}H_{16}DNNaO_6$ $[M+Na]^+$: 271.1011; found: 271.1009.

tert-Butyl (_R_)-_O_-acetyl-_N_-(_tert_-butoxycarbonyl)-L-serinate-3-^{2}H 147



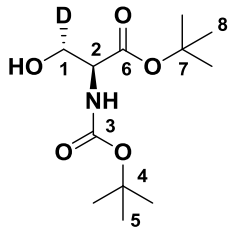
To a refluxing solution of (3*R*)-*O*-acetyl-*N*-(*tert*-butoxycarbonyl)-L-serine-3-*d* (4.90 g, 19.8 mmol) in 75 mL toluene was added a mixture of *tert*-butyl alcohol (5.97 g, 80.6 mmol) and *N*,*N*-dimethylformamide dineopentyl acetal (12.44 g, 53.8 mmol) and the reaction mixture refluxed for 6.5 hours. Then the cooled reaction was diluted with EtOAc, washed with saturated $NaHCO_3$ solution and brine, and dried over $MgSO_4$. The organic layer was concentrated under vacuum and the residue purified by column chromatography (pet ether/EtOAc = 9:1 v/v) to yield *tert*-butyl (*R*)-*O*-acetyl-*N*-(*tert*-butoxycarbonyl)-L-serinate-3-^{2}H (4.50 g, 75%) as a colourless oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ: 1.35 (s, 9H, H-7), 1.36 (s, 9H, H-10), 1.94 (s, 3H, H-1), 4.32 - 4.34 (m, 2H, H-3 and 4), 5.28 (br d, 1H, N-H, J = 7.0 Hz)

^{13}C NMR (100 MHz, $CDCl_3$) δ: 170.17 (C-8), 168.49 (C-2), 155.00 (C-5), 82.34 (C-9), 79.70 (C-6), 64.13 (t, C-3, J = 23.5 Hz), 53.17 (C-4), 28.06 (C-10), 27.66 (C-7), 20.36 (C-1)

HR-MS: *m/z* calculated for $C_{14}H_{25}DNO_6$ $[M+H]^+$: 305.1817; found: 305.1824.

tert-Butyl (*R*)-(_tert_-butoxycarbonyl)-L-serinate-3-^{2}H 148



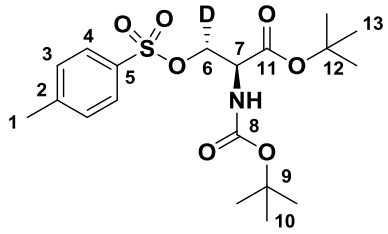
To a solution of *tert*-butyl (*R*)-*O*-acetyl-*N*-(*tert*-butoxycarbonyl)-L-serinate-3-^{2}H (4.20 g, 13.8 mmol) in 65 mL THF was added a solution of LiOH (0.50 g, 20.7 mmol) in 50 mL water and the reaction stirred at RT for 3 hours. The mixture was extracted with EtOAc and the combined organic layers washed successively with water and brine and dried over $MgSO_4$. After removal of the solvent under vacuum, the residue was purified by column chromatography (pet ether/EtOAc = 4:1 v/v) to yield *tert*-butyl (*R*)-(*tert*-butoxycarbonyl)-L-serinate-3-^{2}H (2.24 g, 62%) as a colourless oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ: 1.36 (s, 9H, H-5), 1.39 (s, 9H, H-8), 3.23 (br s, 1H, O-H), 3.79 (br s, 1H, H-1), 4.14 (br s, 1H, H-2), 5.52 (br d, 1H, N-H, J = 7.0 Hz)

^{13}C NMR (100 MHz, $CDCl_3$) δ: 169.80 (C-6), 155.73 (C-3), 82.08 (C-7), 79.71 (C-4), 63.02 (t, C-1, J = 23.0 Hz), 56.09 (C-2), 28.11 (C-8), 27.78 (C-5)

HR-MS: *m/z* calculated for $C_{12}H_{22}DNNaO_5$ $[M+Na]^+$: 285.1531; found: 285.1538.

tert-Butyl (*R*)-*N*-(_tert_-butoxycarbonyl)-*O*-tosyl-L-serinate-3-^{2}H 149



To a solution of *tert*-butyl (*R*)-(*tert*-butoxycarbonyl)-L-serinate-3-^{2}H (2.24 g, 8.5 mmol) in 10 mL pyridine was added 4-toluenesulfonyl chloride (3.00 g, 15.7 mmol) at 0 ℃ and the reaction stirred at 4 ℃ overnight. The mixture was quenched with 10 mL water and extracted with EtOAc. The combined organic layers were washed successively with 1 M HCl, water and brine, and dried over $MgSO_4$. After removal of the solvent under vacuum, the residue was purified by column chromatography (pet ether/EtOAc = 6:1 v/v) to yield

tert-butyl (*R*)-*N*-(*tert*-butoxycarbonyl)-*O*-tosyl-L-serinate-3-^{2}H (3.10 g, 89%) as a white solid.

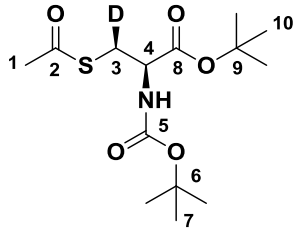
^{1}H NMR (400 MHz, $CDCl_3$) δ: 1.41 (s, 9H, H-10), 1.43 (s, 9H, H-13), 2.44 (s, 3H, H-1), 4.35 - 4.38 (m, 2H, H-6 and 7), 5.27 (br d, 1H, N-H, J = 7.5 Hz), 7.35 (d, 2H, H-3, J = 8.0 Hz), 7.77 (d, 2H, H-4, J = 8.5 Hz)

^{13}C NMR (100 MHz, $CDCl_3$) δ: 167.29 (C-11), 154.91 (C-8), 144.97 (C-2), 132.42 (C-5), 129.83 (C-3), 127.88 (C-4), 83.22 (C-12), 80.07 (C-9), 69.60 (t, C-6, J = 23.5 Hz), 53.21 (C-7), 28.14 (C-13), 27.73 (C-10), 21.54 (C-1)

HR-MS: *m/z* calculated for $C_{19}H_{29}DNO_7S$ $[M+H]^+$: 417.1800; found: 417.1807.

Melting point: 91 - 92 ℃.

tert-Butyl (_S_)-_S_-acetyl-_N_-(_tert_-butoxycarbonyl)-L-cysteinate-3-^{2}H 150



To a solution of *tert*-butyl (*R*)-*N*-(*tert*-butoxycarbonyl)-*O*-tosyl-L-serinate-3-^{2}H (0.80 g, 1.9 mmol) in 5 mL DMF was added potassium thioacetate (0.43 g, 3.8 mmol) under argon at 0 ℃ and the reaction stirred at RT overnight. The solvent was removed under vacuum and the residue extracted with EtOAc. The combined organic phase was washed with water and brine, dried over $MgSO_4$ and concentrated under vacuum. The crude product was purified by column chromatography (pet ether/EtOAc = 6:1 v/v) to yield *tert*-butyl (*S*)-*S*-acetyl-*N*-(*tert*-butoxycarbonyl)-L-cysteinate-3-^{2}H (0.61 g, 99%) as a white solid.

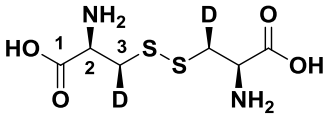
^{1}H NMR (400 MHz, $CDCl_3$) δ: 1.29 (s, 9H, H-7), 1.32 (s, 9H, H-10), 2.19 (s, 3H, H-1), 3.25 (br d, 1H, H-3, J = 4.0 Hz), 4.24 (br dd, 1H, H-4, J = 8.0 and 4.5 Hz), 5.21 (br d, 1H, N-H, J = 7.5 Hz)

^{13}C NMR (100 MHz, $CDCl_3$) δ: 193.95 (C-2), 169.06 (C-8), 154.78 (C-5), 82.12 (C-9), 79.36 (C-6), 53.26 (C-4), 30.75 (t, C-3, J = 22.0 Hz), 30.10 (C-1), 27.94 (C-10), 27.52 (C-7)

HR-MS: *m/z* calculated for $C_{14}H_{24}DNNaO_5S$ $[M+Na]^+$: 343.1408; found: 343.1400.

Melting point: 51-53 ℃.

(2*R*,2′*R*,3*S*,3′*S*)-[3,3'-^{2}H$_2$]Cystine 151



A solution of *tert*-butyl (*S*)-*S*-acetyl-*N*-(*tert*-butoxycarbonyl)-L-cysteinate-3-^{2}H (0.61 g, 1.9 mmol) in 20 mL of 6 M HCl was refluxed overnight. The cooled reaction mixture was washed with chloroform and concentrated. The residue was treated with prewashed and activated ion-exchange resin (Dowex® 50WX8 hydrogen form, 200-400 mesh) and the adsorbed amino acid eluted with 1 M ammonia solution in methanol. The fractions, which gave a positive ninhydrin reaction, were combined, aerated for 48 hours, and concentrated to give (2*R*,2′*R*,3*S*,3′*S*)-[3,3'-^{2}H$_2$]cystine (0.18 g, 78%) as a white solid.

^{1}H NMR (300 MHz, 2.5% NaOD-D_2O) δ: 2.97 (d, 1H, H-3, J = 3.5 Hz), 3.45 (d, 1H, H-2, J = 3.5 Hz)

^{13}C NMR (125 MHz, 2.5% NaOD-D_2O) δ: 180.81 (C-1), 54.80 (C-2), 43.30 (t, C-3, J = 22.0 Hz)

HR-MS: *m/z* calculated for $C_6H_{11}D_2N_2O_4S_2$ $[M+H]^+$: 243.0437; found: 243.0440.

Melting point: > 200 ℃ dec. (open capillary) The sample showed slight discolorization at 190 ℃ and turned black from 200 ℃.

6.2 Analytical chemistry

6.2.1 Extraction protocol

For extracts from liquid bacterial cultures, ethanol was added (in equal volume to the volume of culture media used) and the mixture shaken for 1 hour (180 rpm, 30 ℃) before the insoluble materials were pelleted at 13200 rpm for 10 mins. The supernatant was then filtered (0.4 μm) for LC-MS analyses.

6.2.2 Liquid chromatography-mass spectrometry (LC-MS)

Unless otherwise stated, LC-MS analyses were performed on a Dionex Ultimate 3000 HPLC instrument equipped with a ZORBAX Eclipse Plus C18 column (2.1 × 100 mm, 1.8 μm) coupled to a Bruker MaXis Impact mass spectrometer [ESI in positive ion mode; full scan 50-2500 *m/z*; end plate offset, -500 V; capillary, -4500 V; nebulizer gas (N_2), 1.4 bar; dry gas (N_2), 8 L/min; dry temperature, 200 ℃]. The solvents used for elution of the column were as follows: solvent A, 0.1% formic acid in water; solvent B, 0.1% formic acid in acetonitrile. The elution profile was as follows: 0-5 min, 5% B; 5-17 min, linear gradient from 5% B to 100% B; 17-22 min, 100% B; 22-25 min, linear gradient from 100% B to 5% B; 25-34 min, 5% B. The flow rate was 0.2 mL/min. The mass spectrometer was calibrated with 10 mM sodium formate at the beginning of each run.

For comparison of MS/MS fragmentation pattern of synthetic butenolide and the dephosphorylated product of the MmfL-catalysed reaction, LC-MS/MS analyses were performed on an Agilent 1260 Infinity HPLC equipped with an Agilent Eclipse XDB-C18 column (4.6 × 150 mm, 5 μm) coupled to a Bruker AmaZon X ion trap mass spectrometer [ESI in positive ion mode; full scan 100-3000 *m/z*; end plate offset, -500 V; capillary, -4500 V; nebulizer gas (N_2), 40 p.s.i.; dry gas (N_2), 10 L/min; dry temperature, 200 ℃]. The solvents used for elution of the column were as follows: solvent A, 0.1% formic acid in water; solvent B, 0.1% formic acid in acetonitrile. The elution profile was as follows: 0-5 min, 5% B; 5-30 min, linear gradient from 5% B to 100% B; 30-35 min, 100% B; 35-40 min, linear gradient from 100% B to 5% B. The flow rate was 1.0 mL/min.

For protein samples (25-45 kDa), they are diluted in distilled water to a final concentration of 3-4 μM. LC-MS analyses were performed on a Dionex Ultimate 3000 HPLC instrument equipped with an ACE C4-300 column (2.1 × 100 mm, 5 μm) coupled

to a Bruker MaXis II mass spectrometer [ESI in positive ion mode; full scan 200-3000 *m/z*; end plate offset, -500 V; capillary, -4500 V; nebulizer gas (N_2), 1.8 bar; dry gas (N_2), 9 L/min; dry temperature, 200 ℃]. The solvents used for elution of the column were as follows: solvent A, 0.1% formic acid in water; solvent B, 0.1% formic acid in acetonitrile. The elution profile was as follows: 0-2.5 min, 5% B; 2.5-13.4 min, linear gradient from 5% B to 100% B; 13.4-18.4 min, 100% B; 18.4-21.4 min, linear gradient from 100% B to 5% B; 21.4-30 min, 5% B. The flow rate was 0.2 mL/min.

6.2.3 Chiral LC-MS analysis

An Agilent 1260 Infinity HPLC coupled to a Bruker AmaZon X ion trap mass spectrometer [ESI in positive ion mode; full scan 100-3000 *m/z*; end plate offset, -500 V; capillary, -4500 V; nebulizer gas (N_2), 10 p.s.i.; dry gas (N_2), 4 L/min; dry temperature, 180 ℃] was used for isopyochelin samples on a homochiral stationary phase. Synthetic isopyochelins and the ethanol extract of culture broth from *S. coelicolor* M1152/SV-2_E03::*SspI* were analysed on a ChiralPAK IE analytical column (4.6 × 150 mm, 5 μm) in reverse phase mode. The solvents used for elution of the column were as follows: solvent A, 0.1% formic acid in water; solvent B, 0.1% formic acid in acetonitrile. The elution profile was as follows: 0-5 min, 20% B; 5-30 min, linear gradient from 20% B to 100% B; 30-35 min, 100% B; 35-40 min, linear gradient from 100% B to 20% B. The flow rate was 0.8 mL/min.

6.2.4 Circular dichroism (CD) spectroscopy

The instrument used was a Jasco J-815 CD Spectrometer for analyses of purified MmfR and its mutant proteins as well as the ordered 2-methyl-L-cysteine and 2-methyl-D-cysteine. Concentrated proteins were diluted in 5 mM phosphate buffer pH 7.0 to a final concentration of 0.1 or 0.01 mg/mL and measured in the range from 180 to 260 nm. Chemical samples were dissolved in distilled water to a final concentration of 1 mg/mL and measured in the range from 180 to 300 nm. The samples were loaded on a 1 mm path length cuvette and instrument settings were as follows: sensitivity standard, data pitch 0.2 nm, bandwidth 1.0 nm, data integration time (D.I.T) 1 s, scanning speed 100 nm/min, accumulation 4. Each spectrum was calibrated by subtracting a reference spectrum from a sample containing only buffer or water.

6.3 Biology

6.3.1 Bacterial strains

The wild type *S. coelicolor* M1152 strain was obtained from Dr. Vincent Poon. Strains of *S. coelicolor* M1152/SV-2_E03*::Ssp*I and deletion mutants Δ*sven0508*, Δ*sven0515* and Δ*sven0516* were constructed by Dr. Yuki Inahashi in the Challis group.

6.3.2 Culture media

Mannitol soya flour medium (SFM) was prepared by dissolving Bacto agar (20 g), mannitol (20 g) and soya flour (20 g) in tap water (1 L) and autoclaving before use.

Yeast-extract dextrose (YD) liquid medium was prepared by dissolving yeast extract (10 g) and D-glucose (10 g) in distilled water (1 L) and autoclaving before use.

LB medium was prepared using LB Broth, Miller (Fisher) to a final concentration of 25 g/L. For solid LB, Bacto agar was added (15 g/L).

Supplemented minimal medium (SMM) was made up with the following ingredients added to the concentrations as follows: Bacto casaminoacids (2 g) and TES buffer (5.73 g) were dissolved in distilled water (1 L), the pH adjusted to 7.2 using 10 N NaOH and autoclaved. At the time of use, the following ingredients were added: NaH_2PO_4 + K_2HPO_4 (50 mM each, autoclaved, 10 mL per litre of culture), $MgSO_4$ (1 M, autoclaved, 5 mL per litre of culture), D-glucose (50% w/v, autoclaved, 18 mL per litre of culture) and trace elements solution (1 mL per litre of culture). Trace elements solution contained 0.1 g each of $ZnSO_4.7H_2O$, $FeSO_4.7H_2O$, $MnCl_2.4H_2O$, $CaCl_2.6H_2O$ and NaCl in distilled water (1 L). The solution was stored at 4 ℃ and refreshed every 4 weeks.

6.3.3 Preparation of bacterial spores

Mannitol SFM was used for the production of bacterial spores. Media was melted and apramycin added to a final concentration of 50 μg/mL for growth of the *S. coelicolor* M1152/SV-2_E03*::Ssp*I and Δ*sven0508*, Δ*sven0515* and Δ*sven0516* mutant strains. Each plate of SFM was inoculated with 20 μL of bacterial spore stocks. After incubating at 30 ℃ for 1 week, 5 mL sterile water was added to each plate and an inoculation loop was used to scrape the spores on the surface. The spores suspended in the water were

collected and filtered through a non-absorbent cotton wool using a syringe to remove mycelia. The suspension was then centrifuged at 3000 rpm for 10 mins to pellet the spores and the supernatant removed, leaving approximately 1 mL of suspension. Spore suspension was transferred to a cryotube and an equal volume of sterile 50% glycerol was added and the resulting spore stocks in 25% glycerol stored at -80 ℃ for future use.

6.3.4 Production of 2-hydroxyphenylthiazoline metabolites

Unless otherwise stated, strains of *S. coelicolor* M1152/SV-2_E03*::Ssp*I and deletion mutants Δ*sven0508*, Δ*sven0515* and Δ*sven0516* were grown (30 ℃, 180 rpm) in liquid YD medium for production of the metabolites. Seed cultures were prepared in 50 mL of YD medium, inoculated with 20 μL of spores. After incubation for 2 days, 2% of this culture was used to inoculate 250-mL Erlenmeyer flasks, each containing 50 mL of YD medium. The resulting culture was grown for 5 days. For analysis of these cultures, samples were prepared for LC-MS analyses as described in Section 6.2.1 and 6.2.2.

6.3.5 Feeding the extract from Δ*sven0515* mutant to Δ*sven0516* mutant

50 mL of the 5-day-old culture of the Δ*sven0515* mutant was first centrifuged at 5,000 rpm for 10 mins to remove the mycelium. Then the supernatant was acidified with 1 M HCl to pH 5.0 and extracted with EtOAc. The EtOAc layers were combined, dried over $MgSO_4$ and concentrated under vacuum to give the residue containing the metabolites from the culture of the Δ*sven0515* mutant. The extract was dissolved in 3 mL of YD medium and fed into a 50 mL culture of the Δ*sven0516* mutant which was incubated (30 ℃, 180 rpm) for 1 day after inoculation with 2% of the corresponding seed culture. The resulting culture was grown for 4 days and prepared for LC-MS analyses as described in Section 6.2.1 and 6.2.2.

6.3.6 Feeding experiment with deuterium-labelled precursors

50 mL of YD medium was inoculated with 20 μL spores of *S. coelicolor* M1152/SV-2_E03*::SspI* and incubated (30 ℃, 180 rpm) for 2 days. Then the cells were pelleted at 3000 rpm for 5 mins, washed once with SMM medium and resuspended in 50 mL of SMM medium. 10 mL of SMM medium was inoculated with 0.2 mL (2% v/v) of this bacterial suspension and incubated (30 ℃, 180 rpm). After 24 hours, a suspension

of precursor (2*R*,2′*R*,3*S*,3′*S*)-[3,3'-$^{2}H_2$]cystine or 2,3,3-$^{2}H_3$-L-cysteine (CK Isotopes) in sterile water was added to give a final concentration of 1.25 mM. In some instances, a solution of L-methionine (MP Biomedicals) was also added at the same time. This feeding procedure was repeated at 24 hour intervals until a total incubation time of 5 days, after which the metabolites were extracted and analysed as described in Section 6.2.1 and 6.2.2.

6.3.7 Chemical transformations

For subsequent transformations, 50 μL of chemically competent *E. coli* TOP10 (Invitrogen), *E. coli* BL21 star (DE3) (Invitrogen) or *E. coli* BL21 star (DE3)pLysS (Invitrogen) cells suspension was incubated on ice for 30 mins after a gentle mixing with 1 μL plasmid. The cells were heat-shocked in a 42 ℃ water bath for 30 seconds and immediately transferred on ice. 250 μL of room temperature LB (or SOC) medium was added and cells incubated (37 ℃, 180 rpm) for 1 hour before being spread in two different portions (50 and 200 μL) on LB plates containing 100 μg/mL of ampicillin and incubated at 37 ℃ for 16 hours. Single colonies were picked and incubated further in 10 mL LB containing ampicillin (37 ℃, 180 rpm) for 16 hours. 0.6 mL of each culture was taken and added to 0.6 mL of sterile 50% glycerol for storage at -80 ℃.

For transformations using the *E. coli* TOP10 cells (IBA StarGate), most of the procedures were same as above except the following modifications according to the manufacturer's protocol: cells were heat-shocked at 37 ℃ for 5 mins, and 900 mL of LB medium was used for the 45-min pre-incubation before spreading the cells onto selection plates.

6.3.8 Plasmid DNA preparation

10 μL of glycerol stock of *E. coli* TOP10 cells carrying the plasmid of interest was used to inoculate 10 ml LB medium containing 100 μg/mL of ampicillin. The liquid culture was shaken for 16 hours (37 ℃, 180 rpm). The culture was centrifuged at 2000 rpm for 10 mins and the supernatant discarded. GeneJET plasmid miniprep kit (Thermo Scientific) was used for isolation of the plasmid DNA as per the manufacturer's guidelines. The pelleted cells were resuspended in 250 μL of the resuspension solution by vortexing. The cells were lysed by adding 250 μL of the lysis solution and mixed by inversion. 350 μL of the neutralization solution was then added before centrifugation to remove the cell debris and chromosomal DNA. The supernatant was loaded onto a

GeneJET spin column and centrifuged before the adsorbed plasmid DNA was washed with 500 μL of the wash solution twice. The plasmid DNA was eluted by addition of 50 μL of sterile water on the purification column membrane and centrifuged after incubation at RT for 2 mins. The concentration of purified plasmid DNA solution was determined using a NanoDrop Lite spectrophotometer (Thermo Scientific). The plasmid DNA was stored at -20 ℃.

6.3.9 Restriction digests

Unless manufacturer's guidelines recommended otherwise, 0.5 μL of restriction endonuclease enzyme was incubated with 2 μL of the recommended buffer and 15 μL of the prepared plasmid DNA at 37 ℃ for 2 hours, after which 4 μL of 6X DNA loading dye was added to the reaction mixture and an agarose gel run (1%, w/v) of the resulting fragments. Reactions using FastDigest enzymes were incubated for 30 mins.

6.3.10 Polymerase Chain Reaction (PCR)

An Eppendorf Mastercycler gradient and an Eppendorf Mastercycler personal PCR machine were used for variable annealing temperature and constant annealing temperature PCR reactions, respectively. Unless otherwise stated, the following quantities of each component were used per reaction: sterile water (35.5 μL), DMSO (5 μL), $MgCl_2$ buffer (5 μL), dNTPs (2.5 mM each, 1 μL), Fermentas Taq DNA polymerase (5 U/μL, 0.5 μL), DNA template (1 μL). 1 μL of each forward and reverse primers (10 μM) were used for each reaction. The DNA template was about 100 ng of cosmid used in each reaction, the volume of which was adjusted according to its concentration. The volume of sterile water was adjusted accordingly to make a total volume of 50 μL.

Cosmid C73-787 was used as the template for amplification of small DNA sequences containing the operators for MmfR and provided by Dr. Christophe Corre. Cosmid SV4 F01 was used as the template for cloning of *sgnR* gene and amplification for the intergenic region containing the operator for SgnR and prepared by Dr. John D. Sidda.

Primers were obtained from Sigma-Genosys and diluted to 100 μM as per the manufacturer's guidelines before use. The typical program below was used for the single annealing temperature PCR reactions: Reaction mixture heated to 95 ℃ (5 mins),

35 cycles consisting of denaturation at 95 ℃ (1 min), annealing (1 min) and elongation at 72 ℃. This was followed by a further elongation step at 72 ℃ for 10 mins. The reaction(s) were then held at 4 ℃ until PCR products were used for agarose gel electrophoresis. The annealing temperature and length of time for elongation were determined based on the primers used and length of the target DNA sequence. The specific conditions and primers used in this study was summarised in Table 6.1.

Table 6.1. Primers and PCR conditions used in this study.

Target	Primers	Annealing temp.	Elongation time	Purpose
sgnR (600 bp)	For 5′-CACCGCGACACCCCGAAGCCAA-3′ Rev 5′-TCAGCGGCGCAGGCC-3′	58 ℃	50 s	Producing PCR product for cloning of *sgnR* gene into pET151/D-TOPO vector
sgnR-sgnL (121 bp)	For 5′-GCCTGTCCGCCTCTCTCGTTC-3′ Rev 5′-CCTTTTCCCCCTCAAGGTGAATC-3′	60 ℃	20 s	Entire intergenic region between *sgnR* and *sgnL* containing the proposed binding site for SgnR
mmyR promoter (150 bp)	For 5′-ATCCTGCCGCGCGGTAGCC-3′ Rev 5′-CAACGCCCGAGTCCTCTCAAG-3′	53 ℃	20 s	Entire intergenic region near *mmyR* promoter containing the proposed binding site for MmfR
mmyR promoter (100 bp)	For 5′-GACATCGTCCACGTGCGCCG-3′ Rev 5′-CAACGCCCGAGTCCTCTCAAG-3′	53 ℃	20 s	Partial sequence near *mmyR* promoter containing the proposed binding site for MmfR
mmfR-mmfL (194 bp)	For 5′-GGCTGCCTTCCTTCGTGTG-3′ Rev 5′-AGGGGCGCTACATCTCCCG-3′	50 ℃	30 s	Entire intergenic region between *mmfR* and *mmfL* containing the proposed binding site for MmfR
mmfR-mmfL (100 bp)	For 5′-CACGGAAACCCATTGCATAATACC-3′ Rev 5′-GGACCACCGGCTGGCTTGC-3′	55 ℃	20 s	Partial sequence between *mmfR* and *mmfL* containing the proposed binding site for MmfR
mmyY-mmyB (230 bp)	For 5′-GGTGAACTCCTTCGGCGAG-3′ Rev 5′-GGCGCCTCACAGTGTCAAAC-3′	55 ℃	20 s	Entire intergenic region between *mmyY* and *mmyB* containing the proposed binding site for MmfR
mmyY-mmyB (98 bp)	For 5′-TCCAAACACCGAGGCCCCG-3′ Rev 5′-GGCGCCTCACAGTGTCAAAC-3′	55 ℃	20 s	Partial sequence between *mmyY* and *mmyB* containing the proposed binding site for MmfR

6.3.11 Agarose gel electrophoresis

Agarose electrophoresis gel (1%, w/v) was prepared by mixing 1 g of LE agarose (SeaKem) and 100 mL of 1X TBE buffer. TBE buffer contained Tris base (10.8 g/L), boric acid (5.5 g/L), and EDTA (1.17 g/L). The mixture was heated in a microwave until complete dissolution of the agarose. After the gel mixture cooled down, 5 μL GelRed was added and the gel poured into the gel plate with appropriate gel comb. The gel was allowed to cool further until it set.

To 5 µL DNA sample, 1 µL of 6X DNA loading dye (New England Biolabs) was added and the samples loaded onto the prepared agarose gel for electrophoresis at 80 V using a BioRad tank and Power PAC 300 with a UVP imaging system UV transilluminator.

6.3.12 Purification of DNA from agarose gels

After the electrophoresis, the bands of interest were cut out and placed in Eppendorf tubes. GeneJET gel extraction kit (Thermo Scientific) was used to extract the DNA from the excised gel slice as per the manufacturer's guidelines. One volume of binding buffer to one volume of gel slice was added and the mixture was incubated at 55 ℃ until melted (10 mins + shaking every 2 mins). For DNA fragment whose size is no more than 500 bp, one volume of isopropanol was also added to the sample. After mixing thoroughly by vortexing, the DNA was then extracted using a GeneJET purification column and washed with the wash buffer before the DNA was eluted by addition of 50 µL of sterile water on the purification column membrane and centrifuged. The concentration of purified DNA solution was determined using a NanoDrop Lite spectrophotometer (Thermo Scientific). The DNA was stored at -20 ℃.

6.3.13 Construction of plasmid pET151-*sgnR*

The *sgnR* gene was amplified by PCR with appropriate primers and conditions as described in Section 6.3.10 (Table 6.1). PCR products were run on a 1% agarose gel, the bands of interest excised, and purified using the GeneJET Gel Extraction Kit as described in Section 6.3.12.

The cloning reaction was carried out according to the Invitrogen TOPO cloning manual. Concentration of the PCR product was determined using a NanoDrop Lite spectrophotometer (Thermo Scientific), and diluted accordingly to fall within the desired range of 1-5 ng/µL for cloning into pET151 vector. To a 6 µL reaction, 1 µL of the diluted PCR product, 1 µL of pET151/D-TOPO vector, 1 µL of the provided salt solution and sterile water were mixed together gently and incubated for 5 mins at room temperature and then placed on ice. *E. coli* TOP10 cells (IBA StarGate) were transformed with 3 µL of the TOPO cloning reaction mixture as described in Section 6.3.7 and poured onto selection LB plates containing 100 µg/mL of ampicillin.

Colonies (typically 5 to 10) were picked from these plates, added to 10 mL LB medium containing 100 μg/mL of ampicillin and incubated at 37 ℃ for 16 hours. Glycerol stocks were prepared and plasmids purified from these overnight cultures as described in Section 6.3.8. The purified plasmid was confirmed as containing the correct *sgnR* gene insert by PCR with the primers same as above and by restriction digests. Plasmids that gave both the expected sized PCR product and expected sized DNA fragments resulting from the restriction digests were sent for sequencing by GATC Biotech, during which the gene of interest was sequenced using the T7 primer pairs: forward 5'-TAATACGACTCACTATAGGG-3' and reverse 5'-CTAGTTATTGCTCAGCGG-3'.

6.3.14 Site-directed mutagenesis of MmfR

The construct pET151/D-TOPO vector with MmfR (6405 bp) was prepared prior to the start of this project and provided by Dr. Christophe Corre. Most of the MmfR mutants were produced by using Q5 site-directed mutagenesis kit (New England Biolabs) according to the manufacturer's guidelines with the primers used for each mutant shown in Table 6.2. The elongation time for the PCR in site-directed mutagenesis was 3.5 mins. The MmfR mutants Y85A and Y144A were prepared at the same time and provided by Kathryn Styles from the Corre group (QuikChange Lightning, Agilent). All constructs were sequence verified using the T7 primer pairs and transformed into chemically competent *E. coli* BL21 star (DE3) or *E. coli* BL21 star (DE3)pLysS cells as described in Section 6.3.7.

Table 6.2. Primers and conditions used for site-directed mutagenesis of MmfR.

MmfR mutant	Primers	Annealing temp.
Y85F	For 5′-GAGGAGCACTTCGCGCGCTGG-3′ Rev 5′-CACCACGGCGATGGCCAG-3′	72 ℃
Y144F	For 5′-CCCCTGCCCTTCGTGGACTGG-3′ Rev 5′-CAGCTCCGCGTCGATGAAG-3′	69 ℃
Q130E	For 5′-TGCCCGGCTGGAGAGTGAGCG-3′ Rev 5′-CCGGCCTGCATCACGGGG-3′	72 ℃
L110G	For 5′-CGAGGAGATGGGCCATCGCGCG-3′ Rev 5′-ACCGTCTCCAGCGGTGTG-3′	67 ℃
L110A	For 5′-CGAGGAGATGGCCCATCGCGCG-3′ Rev 5′-ACCGTCTCCAGCGGTGTG-3′	66 ℃
L110V	For 5′-CGAGGAGATGGTCCATCGCGC-3′ Rev 5′-ACCGTCTCCAGCGGTGTG-3′	71 ℃

6.3.15 Expression and overproduction of recombinant protein

10 μL of glycerol stock of *E. coli* BL21star or *E. coli* BL21 star (DE3)pLysS cells carrying the plasmid of interest was used to inoculate 10 mL LB medium containing 100 μg/mL of ampicillin and incubated for 16 hours (37 ℃, 180 rpm). 5 mL of this preculture was used to inoculate 1 L LB medium (0.5% v/v) containing 100 μg/mL of ampicillin and grown (37 ℃, 180 rpm) for 3-5 hours until optical density (OD) of cells had reached 0.6-0.8 at 600 nm wavelength (Thermo BioMate3 spectrophotometer), when isopropyl β-D-1-thiogalactopyranoside (ITPG, 5 mM stock solution) was added to a final concentration of 0.5 mM. After addition of IPTG, the culture was grown for a further 16 hours (15 ℃, 180 rpm).

6.3.16 Purification of recombinant protein

After incubation, cells were pelleted by centrifugation at 5000 rpm for 15 mins and resuspended in washing buffer (15 mL per litre culture) containing 20 mM Tris-HCl (pH 8), 100 mM NaCl, 20 mM imidazole and 10% v/v glycerol. Phenylmethanesuphonyl fluoride (PMSF) was also added to give a final concentration of 1 mM. Resuspended cells were stored at -80 ℃ until time of use. The resuspended cells were lysed using a Constant Systems E1061 cell disrupter, and the cell debris pelleted by centrifugation at 15,000 rpm for 30 mins. A small quantity of the supernatant was diluted in 1 mL of washing buffer and used as a cell-free extract containing the soluble proteins. Meanwhile, a pipette tip was used to transfer a small quantity of the cell debris to 1 mL of washing buffer, and the resuspended cell debris was used later as a sample containing the insoluble proteins.

The supernatant was filtered (0.2 μm) and purified on an Ächta purifier (Amersham Biosciences, GE Healthcare) at 4 ℃ or using a simple bench-top purification with a 1 mL HisTrap HP nickel affinity column (GE Healthcare). In the bench-top purification, the column was first equilibrated with washing buffer and loaded with the supernatant. Unbound proteins were removed by washing with 5 mL of the above washing buffer and His_6-tagged protein was eluted using 3 mL of each elution buffer with successively increasing amount of imidazole [20 mM Tris-HCl (pH 8), 100 mM NaCl, 10% v/v glycerol and 50 mM, 100 mM, 200 mM or 300 mM imidazole]. The factions were analysed by SDS-PAGE. The factions containing target protein were collected and

concentrated using a Vivaspin ultrafiltration column (GE Healthcare or Sartorius) and buffer-changed to the imidazole-free storage buffer containing 20 mM Tris-HCl (pH 8), 100 mM NaCl and 10% v/v glycerol. The protein was stored on ice, distributed into approximately 50 μL aliquots, flash frozen in liquid nitrogen and stored at -80 ℃.

6.3.17 Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)

Samples were prepared by addition of 6 μL of 2X SDS-PAGE sample to 6 μL sample and heated at 95 ℃ for 5 mins before loading onto the prepared gel. The 2X SDS-PAGE sample buffer was prepared as follows: 0.5 M Tris-HCl, pH 6.8 (2.5 mL), glycerol (2.0 mL), β-mercaptoethanol (0.4 mL), bromophenol blue (0.02 g) and SDS (0.4 g) dissolved in distilled water to make a total volume of 10 mL, and then aliquoted and frozen at -20 ℃ until needed.

15% SDS-PAGE gels were used for analysis of recombinant MmfR protein and its mutants as well as SgnR protein, while 12% gel used for recombinant MmfL protein. The resolving and stacking gels were prepared according to the recipe shown in Table 6.3. Tris buffer pH 8.8, Tris buffer pH 6.8 and tris-glycine SDS-PAGE running buffer were prepared according to the procedure laid out by Sambrook and Russel (2001).[204] Constituents were mixed and applied to the gel mould. After allowing the gel to set for 30 mins, samples were loaded and the gel run at 180 V until the dye front had reached the bottom of the gel. All samples were run concurrently with PageRuler Plus prestained protein ladder (Thermo Scientific). Protein staining of SDS-PAGE gels was performed with Instant Blue (Expedeon).

Table 6.3. Recipe for 15% and 12% SDS-PAGE gels.

Constituents	15% Resolving gel solution	12% Resolving gel solution	5% Stacking gel solution
Distilled water (mL)	1.1	1.6	0.68
29:1 (30%) Acrylamide : bisacrylamide (mL)	2.5	2.0	0.17
Tris buffer (mL)	1.3 (pH 8.8, 1.5 M)	1.3 (pH 8.8, 1.5 M)	0.13 (pH 6.8, 1.0 M)
10% SDS solution (mL)	0.05	0.05	0.01
10% Ammonium persulfate solution (mL)	0.05	0.05	0.01
N,*N*,*N*,*N*-Tetramethylethylenediamine (TEMED) (mL)	0.002	0.002	0.001
Total volume (mL)	5	5	1

6.3.18 Electrophoretic mobility shift assay (EMSA)

Most of DNA sequences containing the operator for MmfR or SgnR used in EMSA experiments were amplified by PCR reactions as described in Section 6.3.10. For hairpin DNA, the ordered single-stranded oligonucleotides were diluted in sterile water to 200 nM and warmed at 95 ℃ for 2 mins and slowly cooled down to promote hairpin formation (Table 6.4). For short annealed DNA fragments, paired single-stranded oligonucleotides were diluted in sterile water to 400 nM each and equal volume of them mixed together before heating at 95 ℃ for 5 mins and slowly cooling down (Table 6.4).

Table 6.4. Oligonucleotides for forming hairpin DNA and short annealed DNA fragments.

Target	Oligonucleotides	Purpose
gbnA28	5′-AGTTGATACCTTCGTGAAGGTTTTTATT-3′ 5′-AATAAAAACCTTCACGAAGGTATCAACT-3′	For preparing annealed 28-bp DNA containing the operator for SgnR upstream of *gbnA*
gbnR-sgnH28	5′-AGAAAATACCTTCCAGACCGAATCTTTT-3′ 5′-AAAAGATTCGGTCTGGAAGGTATTTTCT-3′	For preparing annealed 28-bp DNA containing the operator for SgnR between *gbnR* and *sgnH*
Hairpin mmfRL	5′-GCAATATACCTGCGGGAAGGTATTATGCGAG GCATAATACCTTCCCGCAGGTATATTGC-3′	For preparing hairpin DNA containing the operator for MmfR between *mmfR* and *mmfL*
Hairpin mmyBY	5′-GCAAAAAACCTTCGGGAAGGTTTGACGCGAG GCGTCAAACCTTCCCGAAGGTTTTTTGC-3′	For preparing hairpin DNA containing the operator for MmfR between *mmyY* and *mmyB*
Hairpin mmyR	5′-GCAACATACCCTCGGGAAGGTATGTTGCGAG GCAACATACCTTCCCGAGGGTATGTTGC-3′	For preparing hairpin DNA containing the operator for MmfR upstream of *mmyR*

EMSAs were performed according to the manufacturer's instructions in DIG gel shift kit, 2nd generation (Roche) with some modifications. Binding buffer and loading buffer were prepared according to the recipe in the kit protocol. 5X binding buffer contains 100 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) (pH 7.6), 5 mM ethylenediaminetetraacetic acid (EDTA), 50 mM $(NH_4)SO_4$, 5 mM dithiothreitol (DTT), 1% w/v Tween 20 and 150 mM KCl in distilled water. Loading buffer contains 34% glycerol and 0.2% w/v bromophenol blue in 0.25X TBE buffer. 6% and 10% non-denaturing polyacrylamide gels were freshly prepared as shown in Table 6.5 before use. Most EMSAs were run on 6% non-denaturing polyacrylamide gels, while those with hairpin DNA or short annealed DNA fragments were run on 10% gel due to their smaller size. Constituents were mixed together and added to a gel mould with an appropriate comb. Once set, gel was pre-run in 1X TBE buffer at 150 V at 4 ℃ for 1 hour.

A general reaction procedure is as follows. To a PCR tube on ice was added DNA, protein, 5X binding buffer and distilled water. Sample was incubated at room temperature for

15 mins. Inducer dissolved in DMSO with appropriate concentrations was added to the reaction tube, and the mixture incubated at room temperature for a further 10 mins. The sample was then placed on ice and loading buffer was added to give a final volume of 25 μL before being loaded onto an equilibrated 6% non-denaturing polyacrylamide gel. The gel was run at 80 V at 4 ℃ until the dye had migrated to the bottom of the gel. After electrophoresis, the gel was stained in 100 mL of 1X TBE buffer with 5 μL GelRed at room temperature for 30 mins with agitation before visualisation on a UV transilluminator. A typical run is shown in Table 6.6. Total volume of each reaction did not exceed 25 μL.

Table 6.5. Recipe for 6% and 10% non-denaturing polyacrylamide gels.

Constituents	**Non-denaturing polyacrylamide gel**	
	6%	**10%**
Distilled water (mL)	5.97	4.64
29:1 (30%) Acrylamide: bisacrylamide (mL)	2.0	3.33
5X TBE buffer (mL)	2.0	2.0
10% Ammonium persulfate solution (mL)	0.07	0.07
TEMED (mL)	0.0035	0.0035
Total volume (mL)	10	10

Table 6.6. A typical run of an EMSA assay.

Reaction	1	2	3	4	5	6	7	8
DNA (μL)	4	4	4	4	4	4	4	4
Protein (μL)	-	2	2	2	2	2	2	2
Binding buffer (μL)	4	4	4	4	4	4	4	4
Distilled water (μL)	12	10	6	6	6	6	6	6
	Incubation at room temperature for 15 mins							
Inducer (μL) with increasing concentration	-	-	4	4	4	4	4	4
	Incubation at room temperature for a further 10 mins							
Loading buffer (μL)	5	5	5	5	5	5	5	5
Total volume (μL)	25	25	25	25	25	25	25	25

6.3.19 Enzymatic assay of MmfL

400 μM of *S*-(2-acetamidoethyl) 5-methyl-3-oxohexanethioate or *S*-(2-acetamidoethyl) 3-oxooctanethioate, 400 μM of dihydroxyacetone phophate and 3.0 nmol of recombinant His_6-MmfL protein in storage buffer with a final volume of 200 μL were incubated for 90 mins at room temperature. Then, 1 unit of shrimp alkaline phosphatase (New England

Biolabs) was added and a further 45-min incubation at room temperature was carried out. On completion of the reaction, 200 μL of methanol was added to the enzymatic mixture and allowed to stand for 5 mins. After centrifugation at 14, 200 rpm for 10 mins, the supernatant was filtered (0.4 μm) for LC-MS analyses. The amounts of protein used in assays were tested from the standard reaction. For assay with denatured MmfL, protein solution was boiled for 10 mins and cooled down before use.

References

1. Hopwood, D. A., *Streptomyces in nature and medicine: the antibiotic makers*. Oxford University Press: 2007.

2. Poon, V. PhD thesis. University of Warwick, 2015.

3. McCormick, J. R.; Flärdh, K., *FEMS Microbiol. Rev.* **2012,** *36* (1), 206-31.

4. Barka, E. A.; Vatsa, P.; Sanchez, L.; Gaveau-Vaillant, N.; Jacquard, C.; Klenk, H. P.; Clément, C.; Ouhdouch, Y.; Van Wezeld, G. P., *Microbiol. Mol. Biol. Rev.* **2016,** *80* (1), 1-43.

5. Al-Bassam, M. M.; Bibb, M. J.; Bush, M. J.; Chandra, G.; Buttner, M. J., *PLoS Genet.* **2014,** *10* (8), e1004554.

6. Molle, V.; Buttner, M. J., *Mol. Microbiol.* **2000,** *36* (6), 1265-78.

7. Waksman, S. A.; Woodruff, H. B., *J. Bacteriol.* **1942,** *44* (3), 373-84.

8. Schatz, A.; Waksman, S. A., *Exp. Biol. Med.* **1944,** *57* (2), 244-8.

9. Singh, B.; Mitchison, D. A., *Br. Med. J.* **1954,** *1* (4854), 130-2.

10. Weitzman, D.; de Wend Cayley, F. E.; Wingfield, A. L., *Respir. Med.* **1950,** *44* (4), 98-104.

11. Finken, M.; Kirschner, P.; Meier, A.; Wrede, A.; Bottger, E. C., *Mol. Microbiol.* **1993,** *9* (6), 1239-46.

12. de Lima Procópio, R. E.; da Silva, I. R.; Martins, M. K.; de Azevedo, J. L.; de Araújo, J. M., *Braz. J. Infect. Dis.* **2012,** *16* (5), 466-71.

13. Payne, D. J.; Gwynn, M. N.; Holmes, D. J.; Pompliano, D. L., *Nat. Rev. Drug Discov.* **2007,** *6* (1), 29-40.

14. Alanis, A. J., *Arch. Med. Res.* **2005,** *36* (6), 697-705.

15. Whiley, D. M.; Lahra, M. M.; Unemo, M., *Future Microbiol.* **2015,** *10* (3), 313-6.

16. Watve, M. G.; Tickoo, R.; Jog, M. M.; Bhole, B. D., *Arch. Microbiol.* **2001,** *176* (5), 386-90.

17. Katz, L.; Baltz, R. H., *J. Ind. Microbiol. Biotechnol.* **2016,** *43* (2-3), 155-76.

18. Challis, G. L., *J. Ind. Microbiol. Biotechnol.* **2014,** *41* (2), 219-32.

19. Bentley, S. D.; Chater, K. F.; Cerdeno-Tarraga, A. M.; Challis, G. L.; Thomson, N. R.; James, K. D.; Harris, D. E.; Quail, M. A.; Kieser, H.; Harper, D.; Bateman, A.; Brown, S.; Chandra, G.; Chen, C. W.; Collins, M.; Cronin, A.; Fraser, A.; Goble, A.; Hidalgo, J.; Hornsby, T.; Howarth, S.; Huang, C. H.; Kieser, T.; Larke, L.; Murphy, L.; Oliver, K.; O'Neil, S.; Rabbinowitsch, E.; Rajandream, M. A.; Rutherford, K.; Rutter, S.; Seeger, K.; Saunders, D.; Sharp, S.; Squares, R.; Squares, S.; Taylor, K.; Warren, T.; Wietzorrek, A.; Woodward, J.; Barrell, B. G.; Parkhill, J.; Hopwood, D. A., *Nature* **2002,** *417* (6885), 141-7.

20. Bentley, S. D.; Brown, S.; Murphy, L. D.; Harris, D. E.; Quail, M. A.; Parkhill, J.; Barrell, B. G.; McCormick, J. R.; Santamaria, R. I.; Losick, R.; Yamasaki, M.; Kinashi, H.; Chen, C. W.; Chandra, G.; Jakimowicz, D.; Kieser, H. M.; Kieser, T.; Chater, K. F., *Mol. Microbiol.* **2004,** *51* (6), 1615-28.

21. Haug, I.; Weissenborn, A.; Brolle, D.; Bentley, S.; Kieser, T.; Altenbuchner, J., *Microbiology* **2003,** *149* (2), 505-13.

22. Davis, N. K.; Chater, K. F., *Mol. Microbiol.* **1990,** *4* (10), 1679-91.

23. Wright, L. F.; Hopwood, D. A., *J. Gen. Microbiol.* **1976,** *95* (1), 96-106.

24. Kirby, R.; Hopwood, D. A., *J. Gen. Microbiol.* **1977,** *98* (1), 239-52.

25. Hornemann, U.; Hopwood, D. A., Biosynthesis of Methylenomycin A: A Plasmid-Determined Antibiotic. In *Biosynthesis*, Corcoran, J. W., Ed. Springer Berlin Heidelberg: Berlin, Heidelberg, 1981; pp 123-31.

26. Wright, L. F.; Hopwood, D. A., *J. Gen. Microbiol.* **1976,** *96* (2), 289-97.

27. Haynes, S. W.; Sydor, P. K.; Corre, C.; Song, L.; Challis, G. L., *J. Am. Chem. Soc.* **2011,** *133* (6), 1793-8.

28. Hopwood, D. A.; Wright, H. M., *J. Gen. Microbiol.* **1983,** *129* (12), 3575-9.

29. Gottelt, M.; Kol, S.; Gomez-Escribano, J. P.; Bibb, M.; Takano, E., *Microbiology* **2010,** *156* (8), 2343-53.

30. Gomez-Escribano, J. P.; Song, L.; Fox, D. J.; Yeo, V.; Bibb, M. J.; Challis, G. L., *Chem. Sci.* **2012,** *3* (9), 2716-20.

31. Gurtler, H.; Pedersen, R.; Anthoni, U.; Christophersen, C.; Nielsen, P. H.; Wellington, E. M.; Pedersen, C.; Bock, K., *J. Antibiot.* **1994,** *47* (4), 434-9.

32. Zhao, B.; Lin, X.; Lei, L.; Lamb, D. C.; Kelly, S. L.; Waterman, M. R.; Cane, D. E., *J. Biol. Chem.* **2008,** *283* (13), 8183-9.

33. Challis, G. L.; Ravel, J., *FEMS Microbiol. Lett.* **2000,** *187* (2), 111-4.

34. Lautru, S.; Deeth, R. J.; Bailey, L. M.; Challis, G. L., *Nat. Chem. Biol.* **2005,** *1* (5), 265-9.

35. Barona-Gomez, F.; Wong, U.; Giannakopulos, A. E.; Derrick, P. J.; Challis, G. L., *J. Am. Chem. Soc.* **2004,** *126* (50), 16282-3.

36. Ohnishi, Y.; Ishikawa, J.; Hara, H.; Suzuki, H.; Ikenoya, M.; Ikeda, H.; Yamashita, A.; Hattori, M.; Horinouchi, S., *J. Bacteriol.* **2008,** *190* (11), 4050-60.

37. Ikeda, H.; Ishikawa, J.; Hanamoto, A.; Shinose, M.; Kikuchi, H.; Shiba, T.; Sakaki, Y.; Hattori, M.; Omura, S., *Nat. Biotechnol.* **2003,** *21* (5), 526-31.

38. Pullan, S. T.; Chandra, G.; Bibb, M. J.; Merrick, M., *BMC genomics* **2011,** *12*, 175.

39. Zerikly, M.; Challis, G. L., *ChemBioChem* **2009,** *10* (4), 625-33.

40. Medema, M. H.; Breitling, R.; Bovenberg, R.; Takano, E., *Nat. Rev. Microbiol.* **2011,** *9* (2), 131-7.

41. Rutledge, P. J.; Challis, G. L., *Nat. Rev. Microbiol.* **2015,** *13* (8), 509-23.

42. Gomez-Escribano, J. P.; Bibb, M. J., *Microb. Biotechnol.* **2011,** *4* (2), 207-15.

43. Camilli, A.; Bassler, B. L., *Science* **2006,** *311* (5764), 1113-6.

44. Takano, E., *Curr. Opin. Microbiol.* **2006,** *9* (3), 287-94.

45. Corre, C.; Song, L.; O'Rourke, S.; Chater, K. F.; Challis, G. L., *Proc. Natl. Acad. Sci. U. S. A.* **2008,** *105* (45), 17510-5.

46. Khokhlov, A. S.; Tovarova, I. I.; Borisova, L. N.; Pliner, S. A.; Shevchenko, L. N.; Kornitskaia, E. I.; Ivkina, N. S.; Rapoport, I. A., *Doklady Akademii nauk SSSR* **1967,** *177* (1), 232-5.

47. Hara, O.; Beppu, T., *J. Antibiot.* **1982,** *35* (3), 349-58.

48. Ohnishi, Y.; Furusho, Y.; Higashi, T.; Chun, H. K.; Furihata, K.; Sakuda, S.; Horinouchi, S., *J. Antibiot.* **2004,** *57* (3), 218-23.

49. Horinouchi, S.; Beppu, T., *Mol. Microbiol.* **1994,** *12* (6), 859-64.

50. Hsiao, N.-H.; Nakayama, S.; Merlo, M. E.; de Vries, M.; Bunet, R.; Kitani, S.; Nihira, T.; Takano, E., *Chem. Biol.* **2009,** *16* (9), 951-60.

51. Kitani, S.; Kinoshita, H.; Nihira, T.; Yamada, Y., *J. Bacteriol.* **1999,** *181* (16), 5081-4.

52. Gräfe, U.; Schade, W.; Eritt, I.; Fleck, W.; Radics, L., *J. Antibiot.* **1982,** *35* (12), 1722-3.

53. Sakuda, S.; Yamada, Y., *Tetrahedron Lett.* **1991,** *32* (15), 1817-20.

54. Gräfe, U.; Reinhardt, G.; Schade, W.; Eritt, I.; Fleck, W. F.; Radics, L., *Biotechnol. Lett.* **1983,** *5* (9), 591-6.

55. Yamada, Y.; Sugamura, K.; Kondo, K.; Yanagimoto, M.; Okada, H., *J. Antibiot.* **1987,** *40* (4), 496-504.

56. Kondo, K.; Higuchi, Y.; Sakuda, S.; Nihira, T.; Yamada, Y., *J. Antibiot.* **1989,** *42* (12), 1873-6.

57. Takano, E.; Nihira, T.; Hara, Y.; Jones, J. J.; Gershater, C. J.; Yamada, Y.; Bibb, M., *J. Biol. Chem.* **2000,** *275* (15), 11010-6.

58. O'Rourke, S.; Wietzorrek, A.; Fowler, K.; Corre, C.; Challis, G. L.; Chater, K. F., *Mol. Microbiol.* **2009,** *71* (3), 763-78.

59. Sidda, J. D.; Song, L.; Poon, V.; Al-Bassam, M.; Lazos, O.; Buttner, M. J.; Challis, G. L.; Corre, C., *Chem. Sci.* **2014,** *5* (1), 86-9.

60. Sidda, J. D. PhD thesis. University of Warwick, 2015.

61. Kitani, S.; Miyamoto, K. T.; Takamatsu, S.; Herawati, E.; Iguchi, H.; Nishitomi, K.; Uchida, M.; Nagamitsu, T.; Omura, S.; Ikeda, H.; Nihira, T., *Proc. Natl. Acad. Sci. U. S. A.* **2011,** *108* (39), 16410-5.

62. Arakawa, K.; Tsuda, N.; Taniguchi, A.; Kinashi, H., *ChemBioChem* **2012,** *13* (10), 1447-57.

63. Horinouchi, S.; Kumada, Y.; Beppu, T., *J. Bacteriol.* **1984,** *158* (2), 481-7.

64. Horinouchi, S.; Suzuki, H.; Nishiyama, M.; Beppu, T., *J. Bacteriol.* **1989,** *171* (2), 1206-10.

65. Horinouchi, S.; Nishiyama, M.; Suzuki, H.; Kumada, Y.; Beppu, T., *J. Antibiot.* **1985,** *38* (5), 636-41.

66. Sakuda, S.; Higashi, A.; Tanaka, S.; Nihira, T.; Yamada, Y., *J. Am. Chem. Soc.* **1992,** *114* (2), 663-8.

67. Kato, J. Y.; Funa, N.; Watanabe, H.; Ohnishi, Y.; Horinouchi, S., *Proc. Natl. Acad. Sci. U. S. A.* **2007,** *104* (7), 2378-83.

68. Han, L.; Lobo, S.; Reynolds, K. A., *J. Bacteriol.* **1998,** *180* (17), 4481-6.

69. Li, Y.; Florova, G.; Reynolds, K. A., *J. Bacteriol.* **2005,** *187* (11), 3795-9.

70. Hsiao, N.-H.; Söding, J.; Linke, D.; Lange, C.; Hertweck, C.; Wohlleben, W.; Takano, E., *Microbiology* **2007,** *153* (5), 1394-404.

71. Takano, E.; Chakraburtty, R.; Nihira, T.; Yamada, Y.; Bibb, M. J., *Mol. Microbiol.* **2001,** *41* (5), 1015-28.

72. Kawachi, R.; Akashi, T.; Kamitani, Y.; Sy, A.; Wangchaisoonthorn, U.; Nihira, T.; Yamada, Y., *Mol. Microbiol.* **2000,** *36* (2), 302-13.

73. Waki, M.; Nihira, T.; Yamada, Y., *J. Bacteriol.* **1997,** *179* (16), 5131-7.

74. Shikura, N.; Yamamura, J.; Nihira, T., *J. Bacteriol.* **2002,** *184* (18), 5151-7.

75. Corre, C.; Haynes, S. W.; Malet, N.; Song, L.; Challis, G. L., *Chem. Commun.* **2010,** *46* (23), 4079-81.

76. Davis, J. B.; Bailey, J. D.; Sello, J. K., *Org. Lett.* **2009,** *11* (14), 2984-7.

77. Malet, N. PhD thesis. University of Warwick, 2012.

78. Cuthbertson, L.; Nodwell, J. R., *Microbiol. Mol. Biol. Rev.* **2013,** *77* (3), 440-75.

79. Yang, H. L.; Zubay, G.; Levy, S. B., *Proc. Natl. Acad. Sci. U. S. A.* **1976,** *73* (5), 1509-12.

80. Orth, P.; Schnappinger, D.; Hillen, W.; Saenger, W.; Hinrichs, W., *Nat. Struct. Mol. Biol.* **2000,** *7* (3), 215-9.

81. Retzlaff, L.; Distler, J., *Mol. Microbiol.* **1995,** *18* (1), 151-62.

82. Vujaklija, D.; Horinouchi, S.; Beppu, T., *J. Bacteriol.* **1993,** *175* (9), 2652-61.

83. Hsiao, N.-H.; Gottelt, M.; Takano, E., *Methods Enzymol.* **2009,** *458*, 143-57.

84. Bibb, M. J., *Curr. Opin. Microbiol.* **2005,** *8* (2), 208-15.

85. Xu, G.; Wang, J.; Wang, L.; Tian, X.; Yang, H.; Fan, K.; Yang, K.; Tan, H., *J. Biol. Chem.* **2010,** *285* (35), 27440-8.

86. Haneishi, T.; Terahara, A.; Arai, M.; Hata, T.; Tamura, C., *J. Antibiot.* **1974,** *27* (6), 393-9.

87. Folcher, M.; Gaillard, H.; Nguyen, L. T.; Nguyen, K. T.; Lacroix, P.; Bamas-Jacques, N.; Rinkel, M.; Thompson, C. J., *J. Biol. Chem.* **2001,** *276* (47), 44297-306.

88. Leskiw, B. K.; Mah, R.; Lawlor, E. J.; Chater, K. F., *J. Bacteriol.* **1993,** *175* (7), 1995-2005.

89. Sah, S.; Singh, R., *Agriculture* **2015,** *61* (3), 97-114.

90. Saha, M.; Sarkar, S.; Sarkar, B.; Sharma, B. K.; Bhattacharjee, S.; Tribedi, P., *Environ. Sci. Pollut. Res. Int.* **2016,** *23* (5), 3984-99.

91. Sheldon, J. R.; Heinrichs, D. E., *FEMS Microbiol. Rev.* **2015,** *39* (4), 592-630.

92. Ratledge, C.; Dover, L. G., *Annu. Rev. Microbiol.* **2000,** *54*, 881-941.

93. Crowley, D. E.; Wang, Y. C.; Reid, C. P. P.; Szaniszlo, P. J., *Plant Soil* **1991,** *130* (1), 179-98.

94. Cornish, A. S.; Page, W. J., *BioMetals* **1995,** *8* (4), 332-8.

95. Corbin, J. L.; Bulen, W. A., *Biochemistry* **1969,** *8* (3), 757-62.

96. Page, W. J.; Tigerstrom, M. V., *Microbiology* **1988,** *134* (2), 453-60.

97. Wittmann, S.; Heinisch, L.; Scherlitz-Hofmann, I.; Stoiber, T.; Ankel-Fuchs, D.; Möllmann, U., *BioMetals* **2004,** *17* (1), 53-64.

98. Ward, T. R.; Lutz, A.; Parel, S. P.; Ensling, J.; Gutlich, P.; Buglyo, P.; Orvig, C., *Inorg. Chem.* **1999,** *38* (22), 5007-17.

99. Raymond, K. N.; Dertz, E. A.; Kim, S. S., *Proc. Natl. Acad. Sci. U. S. A.* **2003,** *100* (7), 3584-8.

100. Münzinger, M.; Taraz, K.; Budzikiewicz, H.; Drechsel, H.; Heymann, P.; Winkelmann, G.; Meyer, J.-M., *BioMetals* **1999,** *12* (2), 189-93.

101. De Voss, J. J.; Rutter, K.; Schroeder, B. G.; Su, H.; Zhu, Y.; Barry, C. E., *Proc. Natl. Acad. Sci. U. S. A.* **2000,** *97* (3), 1252-7.

102. Lytton, S. D.; Mester, B.; Dayan, I.; Glickstein, H.; Libman, J.; Shanzer, A.; Cabantchik, Z. I., *Blood* **1993,** *81* (1), 214-21.

103. Mislin, G. L. A.; Schalk, I. J., *Metallomics* **2014,** *6* (3), 408-20.

104. Barry, S. M.; Challis, G. L., *Curr. Opin. Chem. Biol.* **2009,** *13* (2), 205-15.

105. Crosa, J. H.; Walsh, C. T., *Microbiol. Mol. Biol. Rev.* **2002,** *66* (2), 223-49.

106. Guerinot, M. L., *Annu. Rev. Microbiol.* **1994,** *48*, 743-72.

107. Miethke, M.; Marahiel, M. A., *Microbiol. Mol. Biol. Rev.* **2007,** *71* (3), 413-51.

108. Actis, L. A.; Fish, W.; Crosa, J. H.; Kellerman, K.; Ellenberger, S. R.; Hauser, F. M.; Sanders-Loehr, J., *J. Bacteriol.* **1986,** *167* (1), 57-65.

109. Jalal, M. A. F.; Hossain, M. B.; Van der Helm, D.; Sanders-Loehr, J.; Actis, L. A.; Crosa, J. H., *J. Am. Chem. Soc.* **1989,** *111* (1), 292-6.

110. Drechsel, H.; Stephan, H.; Lotz, R.; Haag, H.; Zähner, H.; Hantke, K.; Jung, G., *Liebigs Ann. Recl.* **1995,** *1995* (10), 1727-33.

111. Cox, C. D.; Rinehart, K. L.; Moore, M. L.; Cook, J. C., *Proc. Natl. Acad. Sci. U. S. A.* **1981,** *78* (7), 4256-60.

112. Zunnundzhanov, A.; Bessonova, I. A.; Abdullaev, N. D.; Ogai, D. K., *Chem. Nat. Compd.* **1987,** *23* (4), 461-5.

113. Lee, J. Y.; Moon, S. S.; Hwang, B. K., *Appl. Environ. Microbiol.* **2003,** *69* (4), 2023-31.

114. Youard, Z. A.; Mislin, G. L.; Majcherczyk, P. A.; Schalk, I. J.; Reimmann, C., *J. Biol. Chem.* **2007,** *282* (49), 35546-53.

115. Lin, Z.; Antemano, R. R.; Hughen, R. W.; Tianero, M. D. B.; Peraud, O.; Haygood, M. G.; Concepcion, G. P.; Olivera, B. M.; Light, A.; Schmidt, E. W., *J. Nat. Prod.* **2010,** *73* (11), 1922-6.

116. Shindo, K.; Takenaka, A.; Noguchi, T.; Hayakawa, Y.; Seto, H., *J. Antibiot.* **1989,** *42* (10), 1526-9.

117. Sasaki, O.; Igarashi, Y.; Saito, N.; Furumai, T., *J. Antibiot.* **2002,** *55* (3), 249-55.

118. Igarashi, Y.; Asano, D.; Sawamura, M.; In, Y.; Ishida, T.; Imoto, M., *Org. Lett.* **2016,** *18* (7), 1658-61.

119. Thomas, M. S., *BioMetals* **2007,** *20* (3), 431-52.

120. Castignetti, D., *Curr. Microbiol.* **1997,** *34* (4), 250-7.

121. Seipke, R. F.; Song, L.; Bicz, J.; Laskaris, P.; Yaxley, A. M.; Challis, G. L.; Loria, R., *Microbiology* **2011,** *157* (9), 2681-93.

122. Klumpp, C.; Burger, A.; Mislin, G. L.; Abdallah, M. A., *Bioorg. Med. Chem. Lett.* **2005,** *15* (6), 1721-4.

123. Miller, M. C.; Parkin, S.; Fetherston, J. D.; Perry, R. D.; DeMoll, E., *J. Inorg. Biochem.* **2006,** *100* (9), 1495-500.

124. Budzikiewicz, H., *Fortschr. Chem. Org. Naturst.* **2004,** *87*, 81-237.

125. Meyer, J. M., *Arch. Microbiol.* **2000,** *174* (3), 135-42.

126. Adler, C.; Corbalán, N. S.; Seyedsayamdost, M. R.; Pomares, M. F.; de Cristóbal, R. E.; Clardy, J.; Kolter, R.; Vincent, P. A., *PLoS ONE* **2012,** *7* (10), e46754.

127. Noel, S.; Gasser, V.; Pesset, B.; Hoegy, F.; Rognan, D.; Schalk, I. J.; Mislin, G. L. A., *Org. Biomol. Chem.* **2011,** *9* (24), 8288-300.

128. Quadri, L. E. N.; Keating, T. A.; Patel, H. M.; Walsh, C. T., *Biochemistry* **1999,** *38* (45), 14941-54.

129. Patel, H. M.; Walsh, C. T., *Biochemistry* **2001,** *40* (30), 9023-31.

130. Serino, L.; Reimmann, C.; Baur, H.; Beyeler, M.; Visca, P.; Haas, D., *Mol. Gen. Genet.* **1995,** *249* (2), 217-28.

131. Serino, L.; Reimmann, C.; Visca, P.; Beyeler, M.; Chiesa, V. D.; Haas, D., *J. Bacteriol.* **1997,** *179* (1), 248-57.

132. Reimmann, C.; Patel, H. M.; Serino, L.; Barone, M.; Walsh, C. T.; Haas, D., *J. Bacteriol.* **2001,** *183* (3), 813-20.

133. Reimmann, C.; Patel, H. M.; Walsh, C. T.; Haas, D., *J. Bacteriol.* **2004,** *186* (19), 6367-73.

134. Kotowska, M.; Pawlik, K., *Appl. Microbiol. Biotechnol.* **2014,** *98* (18), 7735-46.

135. Youard, Z. A.; Wenner, N.; Reimmann, C., *BioMetals* **2011,** *24* (3), 513-22.

136. Inahashi, Y.; Zhou, S.; Bibb, M. J.; Song, L.; Al-Bassam, M. M.; Bibb, M. J.; Challis, G. L., *Chem. Sci.* **2016**, submitted for publication.

137. Fontecave, M.; Atta, M.; Mulliez, E., *Trends Biochem. Sci.* **2004,** *29* (5), 243-9.

138. Dong, C.; Huang, F.; Deng, H.; Schaffrath, C.; Spencer, J. B.; O'Hagan, D.; Naismith, J. H., *Nature* **2004,** *427* (6974), 561-5.

139. Deng, H.; Cobb, S. L.; McEwan, A. R.; McGlinchey, R. P.; Naismith, J. H.; O'Hagan, D.; Robinson, D. A.; Spencer, J. B., *Angew. Chem. Int. Ed. (English)* **2006,** *45* (5), 759-62.

140. Lin, H., *Bioorg. Chem.* **2011,** *39* (5-6), 161-70.

141. Bauerle, M. R.; Schwalm, E. L.; Booker, S. J., *J. Biol. Chem.* **2015,** *290* (7), 3995-4002.

142. Broderick, J. B.; Duffus, B. R.; Duschene, K. S.; Shepard, E. M., *Chem. Rev.* **2014,** *114* (8), 4229-317.

143. Zhang, Q.; van der Donk, W. A.; Liu, W., *Acc. Chem. Res.* **2012,** *45* (4), 555-64.

144. Hu, Y.; Ribbe, M. W., *Curr. Opin. Chem. Biol.* **2016,** *31*, 188-94.

145. Benítez-Páez, A.; Villarroya, M.; Armengod, M.-E., *RNA* **2012,** *18* (10), 1783-95.

146. Giessing, A. M. B.; Jensen, S. S.; Rasmussen, A.; Hansen, L. H.; Gondela, A.; Long, K.; Vester, B.; Kirpekar, F., *RNA* **2009,** *15* (2), 327-36.

147. Allen, K. D.; Xu, H.; White, R. H., *J. Bacteriol.* **2014,** *196* (18), 3315-23.

148. Zhou, S.; Alkhalaf, L.; de los Santos, E. L. C.; Challis, G. L., *Curr. Opin. Chem. Biol.* **2016,** *35*, 73-79.

149. Huang, C.; Huang, F.; Moison, E.; Guo, J.; Jian, X.; Duan, X.; Deng, Z.; Leadlay, P. F.; Sun, Y., *Chem. Biol.* **2015,** *22* (2), 251-61.

150. Kuzuyama, T.; Seki, T.; Dairi, T.; Hidaka, T.; Seto, H., *J. Antibiot.* **1995,** *48* (10), 1191-3.

151. Marous, D. R.; Lloyd, E. P.; Buller, A. R.; Moshos, K. A.; Grove, T. L.; Blaszczyk, A. J.; Booker, S. J.; Townsend, C. A., *Proc. Natl. Acad. Sci. U. S. A.* **2015,** *112* (33), 10354-8.

152. Rachid, S.; Scharfe, M.; Blöcker, H.; Weissman, K. J.; Müller, R., *Chem. Biol.* **2009,** *16* (1), 70-81.

153. Benjdia, A.; Pierre, S.; Gherasim, C.; Guillot, A.; Carmona, M.; Amara, P.; Banerjee, R.; Berteau, O., *Nat. Commun.* **2015,** *6*, 8377.

154. Blaszczyk, A. J.; Silakov, A.; Zhang, B.; Maiocco, S. J.; Lanz, N. D.; Kelly, W. L.; Elliott, S. J.; Krebs, C.; Booker, S. J., *J. Am. Chem. Soc.* **2016,** *138* (10), 3416-26.

155. Huo, L.; Rachid, S.; Stadler, M.; Wenzel, S. C.; Muller, R., *Chem. Biol.* **2012,** *19* (10), 1278-87.

156. Gomez-Escribano, J. P.; Song, L.; Bibb, M. J.; Challis, G. L., *Chem. Sci.* **2012,** *3* (12), 3522-5.

157. Crone, W. J. K.; Leeper, F. J.; Truman, A. W., *Chem. Sci.* **2012,** *3* (12), 3516-21.

158. Freeman, M. F.; Gurgui, C.; Helf, M. J.; Morinaka, B. I.; Uria, A. R.; Oldham, N. J.; Sahl, H.-G.; Matsunaga, S.; Piel, J., *Science* **2012,** *338* (6105), 387-90.

159. Werner, W. J.; Allen, K. D.; Hu, K.; Helms, G. L.; Chen, B. S.; Wang, S. C., *Biochemistry* **2011,** *50* (42), 8986-8.

160. Ahlert, J.; Shepard, E.; Lomovskaya, N.; Zazopoulos, E.; Staffa, A.; Bachmann, B. O.; Huang, K.; Fonstein, L.; Czisny, A.; Whitwam, R. E.; Farnet, C. M.; Thorson, J. S., *Science* **2002,** *297* (5584), 1173-6.

161. Lohman, J. R.; Huang, S. X.; Horsman, G. P.; Dilfer, P. E.; Huang, T.; Chen, Y.; Wendt-Pienkowski, E.; Shen, B., *Mol. Biosyst.* **2013,** *9* (3), 478-91.

162. Westrich, L.; Heide, L.; Li, S. M., *ChemBioChem* **2003,** *4* (8), 768-73.

163. Welander, P. V.; Coleman, M. L.; Sessions, A. L.; Summons, R. E.; Newman, D. K., *Proc. Natl. Acad. Sci. U. S. A.* **2010,** *107* (19), 8537-42.

164. Noike, M.; Matsui, T.; Ooya, K.; Sasaki, I.; Ohtaki, S.; Hamano, Y.; Maruyama, C.; Ishikawa, J.; Satoh, Y.; Ito, H.; Morita, H.; Dairi, T., *Nat. Chem. Biol.* **2015,** *11* (1), 71-6.

165. Ostash, B.; Doud, E. H.; Lin, C.; Ostash, I.; Perlstein, D. L.; Fuse, S.; Wolpert, M.; Kahne, D.; Walker, S., *Biochemistry* **2009,** *48* (37), 8830-41.

166. Gilbert, I. H.; Ginty, M.; O'Neill, J. A.; Simpson, T. J.; Staunton, J.; Willis, C. L., *Bioorg. Med. Chem. Lett.* **1995,** *5* (15), 1587-90.

167. Georgiades, S. N.; Clardy, J., *Bioorg. Med. Chem. Lett.* **2008,** *18* (10), 3117-21.

168. Morin, J. B.; Adams, K. L.; Sello, J. K., *Org. Biomol. Chem.* **2012,** *10* (8), 1517-20.

169. Sakuda, S.; Tanaka, S.; Mizuno, K.; Sukcharoen, O.; Nihira, T.; Yamada, Y., *J. Chem. Soc., Perkin Trans. 1* **1993,** (19), 2309-15.

170. Hellman, L. M.; Fried, M. G., *Nat. Protoc.* **2007,** *2* (8), 1849-61.

171. Laniel, M. A.; Beliveau, A.; Guerin, S. L., *Methods Mol. Biol.* **2001,** *148*, 13-30.

172. Stanton, B. C.; Nielsen, A. A. K.; Tamsir, A.; Clancy, K.; Peterson, T.; Voigt, C. A., *Nat. Chem. Biol.* **2014,** *10* (2), 99-105.

173. Mori, K., *Tetrahedron Lett.* **1981,** *22* (35), 3431-2.

174. Posner, G. H.; Weitzberg, M.; Jew, S.-S., *Synth. Commun.* **1987,** *17* (6), 611-20.

175. Kinoshita, T.; Hirano, M., *J. Heterocycl. Chem.* **1992,** *29* (4), 1025-6.

176. Parsons, P. J.; Lacrouts, P.; Buss, A. D., *J. Chem. Soc., Chem. Commun.* **1995,** (4), 437-8.

177. Crawforth, J. M.; Fawcett, J.; Rawlings, B. J., *J. Chem. Soc., Perkin Trans. 1* **1998,** (10), 1721-6.

178. Chavan, S. P.; Pasupathy, K.; Shivasankar, K., *Synth. Commun.* **2004,** *34* (3), 397-404.

179. Fernández-Martínez, L. T.; Borsetto, C.; Gomez-Escribano, J. P.; Bibb, M. J.; Al-Bassam, M. M.; Chandra, G.; Bibb, M. J., *Antimicrob. Agents Chemother.* **2014,** *58* (12), 7441-50.

180. Ankenbauer, R. G.; Toyokuni, T.; Staley, A.; Rinehart Jr, K. L.; Cox, C. D., *J. Bacteriol.* **1988,** *170* (11), 5344-51.

181. Rinehart, K. L.; Staley, A. L.; Wilson, S. R.; Ankenbauer, R. G.; Cox, C. D., *J. Org. Chem.* **1995,** *60* (9), 2786-91.

182. Zamri, A.; Abdallah, M. A., *Tetrahedron* **2000,** *56* (2), 249-56.

183. Dunetz, J. R.; Magano, J.; Weisenburger, G. A., *Org. Process Res. Dev.* **2016,** *20* (2), 140-77.

184. Al-Warhi, T. I.; AL-Hazimi, H. M. A.; El-Faham, A.; Albericio, F., *Molecules* **2010,** *15* (12), 9403-17.

185. Allen, K. D.; Wang, S. C., *Arch. Biochem. Biophys.* **2014,** *543*, 67-73.

186. Kuzuyama, T.; Hidaka, T.; Kamigiri, K.; Imai, S.; Seto, H., *J. Antibiot.* **1992,** *45* (11), 1812-4.

187. Woodyer, R. D.; Li, G.; Zhao, H.; van der Donk, W. A., *Chem. Commun.* **2007,** (4), 359-61.

188. Hammerschmidt, F., *Liebigs Ann. Chem.* **1992,** *1992* (6), 553-7.

189. Hammerschmidt, F.; Kaehlig, H., *J. Org. Chem.* **1991,** *56* (7), 2364-70.

190. Kellenberger, J. L. PhD thesis. ETH Zurich, 1997.

191. Aberhart, D. J.; Lin, L. J.; Chu, J. Y.-R., *J. Chem. Soc., Perkin Trans. 1* **1975,** (23), 2517-23.

192. Young, D. W.; Sen, P. K.; Morecombe, D. J., *Eur. J. Biochem.* **1977,** *75* (1), 133-47.

193. Borcsok, E.; Abeles, R. H., *Arch. Biochem. Biophys.* **1982,** *213* (2), 695-707.

194. Axelsson, B. S.; O'Toole, K. J.; Spencer, P. A.; Young, D. W., *J. Chem. Soc., Perkin Trans. 1* **1994,** (7), 807-15.

195. Ruszczycky, M. W.; Anderson, V. E., *Bioorg. Chem.* **2004,** *32* (1), 51-61.

196. Ramer, S. E.; Moore, R. N.; Vederas, J. C., *Can. J. Chem.* **1986,** *64* (4), 706-13.

197. Slieker, L.; Benkovic, S. J., *J. Labelled Compd. Radiopharm.* **1982,** *19* (5), 647-57.

198. Baldwin, J. E.; Adlington, R. M.; Robinson, N. G.; Ting, H.-H., *J. Chem. Soc., Chem. Commun.* **1986,** (5), 409-11.

199. Strauss, E.; Begley, T. P., *Bioorg. Med. Chem. Lett.* **2003,** *13* (3), 339-42.

200. Maeda, Y.; Tago, K.; Eguchi, T.; Kakinuma, K., *Biosci., Biotechnol., Biochem.* **1996,** *60* (8), 1248-54.

201. Oba, M.; Iwasaki, A.; Hitokawa, H.; Ikegame, T.; Banba, H.; Ura, K.; Takamura, T.; Nishiyama, K., *Tetrahedron: Asymmetry* **2006,** *17* (12), 1890-4.

202. Jurczak, J.; Golebiowski, A., *Chem. Rev.* **1989,** *89* (1), 149-64.

203. Eliot, A. C.; Kirsch, J. F., *Annu. Rev. Biochem* **2004,** *73*, 383-415.

204. Sambrook, J.; Russell, D. W., *Molecular cloning: a laboratory manual 3rd edition*. Cold Spring Harbor Laboratory Press: 2001.