

Bio-engineering of Antibiotic Enduracidin Biosynthetic Pathways and PreQ₁ Riboswitch

A thesis submitted to the University of Manchester for the degree of
Doctor of Philosophy
in the Faculty of Engineering and Physical Sciences.

2011

Ming-Cheng Wu

**School of Chemistry and
Manchester Interdisciplinary Biocentre**

Table of Contents

Table of Contents	2
Table of Figures	8
Abbreviations	13
Abstract	16
Declaration	17
Copyright Statement	17
Acknowledgements	18
The Author	19
Chapter 1. Introduction	20
1.1. Natural products — Metabolites	21
1.2. The Secondary Metabolite Producing Actinomycetes	23
1.3. Antibiotics	24
1.4. The Target Sites of Antibiotics.....	26
1.4.1. Nucleic Acid Synthesis.....	26
1.4.2. Protein Synthesis	27
1.4.3. Essential Metabolic Pathway.....	27
1.5. Antibiotic Production and Regulation in <i>Streptomyces</i>	33
1.6. Non-Ribosomal Peptide Antibiotic Synthesis	35
1.6.1. The Adenylation (A) Domain.....	39
1.6.2. The Peptidyl Carrier Protein (PCP) Domain	40
1.6.3. The Condensation (C) Domain.....	41
1.6.4. The Thioesterase (TE) Domain	43
1.6.5. The Type II Thioesterase (TE II).....	44
1.6.6. Other Domains in NRPSs	44

1.7.	Tailoring Enzymes.....	46
1.7.1.	Hydroxylation.....	46
1.7.2.	Methylation.....	48
1.7.3.	Halogenation.....	49
1.7.4.	Mannosylation.....	52
1.8.	The Antibiotics A54145, CDA and Daptomycin.....	53
1.8.1.	The Antibiotic CDA.....	54
1.8.2.	The Antibiotic Daptomycin.....	56
1.8.3.	The Antibiotic A54145.....	58
1.9.	The Antibiotics Enduracidin and Ramoplanin.....	59
1.9.1.	The Antibiotic Enduracidin.....	62
1.9.2.	The Antibiotic Ramoplanin.....	66
1.10.	Riboswitches as Gene Regulation Tools.....	70
1.10.1.	Diversity of Riboswitches.....	71
1.10.2.	Riboswitch-Based Applications.....	73
1.11.	Aims of the Project.....	75
1.11.1.	The Biosynthetic Origin of OmAsp.....	75
1.11.2.	The Bio-engineering of the Lipoglycopeptide Antibiotic Ramoplanin by Chlorination and Mannosylation.....	76
1.11.3.	Re-engineering PreQ ₁ Riboswitches.....	76
Chapter 2.	The biosynthetic origins of 3- <i>O</i> -methyl aspartic acid (OmAsp).....	77
2.1.	Introduction.....	77
2.2.	LptL: Non-heme Fe(II)/ α -KG-dependent hydroxylase.....	78
2.2.1.	Bioinformatic Analysis of LptL.....	78
2.2.2.	Heterologous Expression of <i>lptL</i> in <i>E.coli</i>	82
2.2.3.	Hydroxylase Activity Assay.....	83
2.2.4.	Modification timing for formation of OmAsp.....	86

2.2.5.	Summary.....	88
2.3.	LptJ: SyrP-like protein/Hydroxylase.....	89
2.3.1.	Bioinformatic Analysis of LptJ.....	89
2.3.2.	Heterologous expression of <i>lptJ</i> in <i>E.coli</i>	90
2.3.3.	Hydroxylase Activity Assay.....	91
2.3.4.	Summary.....	93
2.4.	LptK: O-methyl transferase (O-MTase).....	94
2.4.1.	Bioinformatic Analysis of LptK.....	94
2.4.2.	Heterologous Expression of <i>lptK</i> in <i>E.coli</i>	95
2.4.3.	O-MTase Activity Assay.....	96
2.4.4.	Summary.....	97
2.5.	<i>In Vivo</i> Studies of the genes <i>lptJ</i> and <i>lptK</i>	98
2.5.1.	Glycerol Inducible Vector pMT3226.....	98
2.5.2.	Non-integrative pIJ86 Vector.....	109
2.6.	Discussion.....	115
Chapter 3.	Bio-engineering of the Lipopeptide Antibiotic Enduracidin by Chlorination and Mannosylation.....	117
3.1.	Introduction.....	117
3.2.	Isolation and Characterisation of Enduracidin from <i>S. fungicidicus</i> ATCC31731 (Strain No. Emt 2-140).....	119
3.3.	Verification of the pMS17 Vector Functionality in <i>S. fungicidicus</i> ...	123
3.4.	Bio-engineering of Enduracidin with Chlorination by the Halogenase Ram20.....	126
3.4.1.	Introduction.....	126
3.4.2.	Bioinformatic Analysis of Ram20.....	127
3.4.3.	Timing of Chlorination.....	129
3.4.4.	Construction of the Transformant Strain SfpMS17-ram20.....	130

3.4.5.	MS analysis of the Fermentation Broth of the SfpMS17-ram20 Strain.....	131
3.4.6.	Summary.....	134
3.5.	Bio-engineering of Enduracidin with Mannosylation by the Mannosyl Transferase (Ram29).....	135
3.5.1.	Introduction.....	135
3.5.2.	Bioinformatic Analysis of Ram29.....	138
3.5.3.	Proposed Mannosylation Mechanism.....	140
3.5.4.	<i>In Vivo</i> Studies of the Ram29.....	141
3.5.5.	<i>In Vitro</i> Studies of Ram29 Activity.....	145
3.5.6.	Summary.....	148
3.6.	Discussion.....	148
Chapter 4.	Re-engineering PreQ ₁ Riboswitches.....	152
4.1.	Introduction.....	152
4.2.	Mutational Analysis of PreQ ₁ Riboswitches.....	156
4.2.1.	Construction of PreQ ₁ Riboswitch Mutants with the Reporter Gene <i>lacZ</i> 156	
4.2.2.	The Effects of Variant PreQ ₁ Riboswitches on the Expression of <i>lacZ</i>	157
4.3.	The Molecular Recognition Determinants for the PreQ ₁ Riboswitch	160
4.3.1.	PreQ ₁ Riboswitch Interactions With Guanine-Derived Molecules.....	160
4.3.2.	Testing PreQ ₁ Analogues With the Variant PreQ ₁ Constructs.....	162
4.4.	Discussion.....	168
Chapter 5.	Conclusions and Future Work.....	170
Chapter 6.	Materials and Methods.....	173
6.1.	General information.....	173
6.2.	Protein Expression Vector Maps.....	176
6.2.1.	pMT3226.....	176

6.2.2.	pIJ86	176
6.2.3.	pMS17 and i-pAV11b	177
6.2.4.	pDG1661	178
6.3.	Media and Buffers	179
6.3.1.	<i>E. coli</i> Culture Media.....	179
6.3.2.	<i>Streptomyces</i> Culture Media.....	179
6.3.3.	General Buffers.....	180
6.4.	Preparation of Chemically Competent <i>E.coli</i> Cells.....	182
6.5.	Preparation of <i>Bacillus</i> Competent Cells	183
6.6.	<i>E. coli</i> Transformation Procedures	183
6.7.	<i>Bacillus</i> Transformation Procedures	183
6.8.	Growth of <i>Streptomyces</i> Cultures.....	184
6.8.1.	Plating out <i>Streptomyces</i> Spore Suspensions	184
6.8.2.	Harvesting of Spores from <i>Streptomyces</i> mycelia.....	184
6.8.3.	Growth of <i>Streptomyces</i> Mycelia in Liquid Media	184
6.9.	Extraction of <i>Streptomyces</i> Genomic DNA	185
6.10.	Cloning of Target Genes (<i>lptL</i> , <i>lptK</i> , <i>lptJ</i> , <i>lptJK</i> , <i>ram20</i> , <i>ram29</i> , <i>ppm1</i> , <i>preQ1</i>)	185
6.10.1.	PCR Amplification Reaction	186
6.10.2.	Primers for the gene <i>lptL</i> , <i>lptK</i> , <i>lptJ</i> , <i>lptJK</i> , <i>ram20</i> , <i>ram29</i> , <i>eGFP</i> and <i>ppm1</i>	187
6.10.3.	Primers for Constructing PreQ ₁ Riboswitch Variants	188
6.11.	DNA Electrophoresis.....	189
6.12.	Extraction of DNA from Agarose Gels	189
6.13.	Digestion, Ligation and Transformation	189
6.14.	Recombinant Protein Expression and Purification.....	190
6.15.	Introduction of DNA into <i>Streptomyces</i> by Conjugation.....	191

6.16.	Catechol Dioxygenase Assay	192
6.17.	Antibiotic Activity Plate Bioassay	192
6.18.	CDA Purification from <i>S. coelicolor</i> Fermentation Broth	193
6.19.	Enduracidin Purification from <i>S. fungicidicus</i> Fermentation Broth ...	193
6.20.	Analysis of CDA and Enduracidin extracts using LC-MS.....	194
6.21.	Tandem Mass Spectrometry (MS).....	195
6.22.	Enduracidin Sequencing	196
6.23.	Bradford Assay (Protein Quantitation).....	196
6.24.	Sodium-Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE)	196
6.25.	Amino Acid Pre-column Derivatisation	197
6.26.	Enzyme Assays.....	198
6.26.1.	Hydroxylase Activity Assay	198
6.26.2.	Methyl Transferase Activity Assay	198
6.26.3.	Alpha-mannosidase Activity Assay.....	198
6.26.4.	Mannosyl Transferase Assay	199
6.27.	Pre-Q ₁ Riboswitch Functional Assay	200
Chapter 7.	References	201

Table of Figures

Figure 1.1. A transcription unit and mRNA transcript structure.	22
Figure 1.2. The developmental life cycle of actinomycetes.	23
Figure 1.3. UDP-MurNAc-pentapeptide biosynthesis.	29
Figure 1.4. Lipid I and Lipid II biosynthesis.	29
Figure 1.5. Transglycosylation of lipid II by transglycosylase.	31
Figure 1.6. The transpeptidation reaction of transpeptidase.	31
Figure 1.7. The peptidoglycan biosynthetic pathway with sites of antibiotic interference highlighted.	32
Figure 1.8. The AbsA two-component signal transduction system in <i>S. coelicolor</i>	34
Figure 1.9. The non-ribosomal peptide biosynthetic pathway.	37
Figure 1.10. Iterative and non-linear NRPS architectures.	38
Figure 1.11. Amino acid adenylation by the A domain of NRPS.	39
Figure 1.12. Activation of the PCP domain by post-translational modification with the ppant moiety.	41
Figure 1.13. Loading of amino acid onto the PCP domain.	41
Figure 1.14. Peptide bond formation by the C domain.	42
Figure 1.15. The peptide product-releasing mechanisms of the TE domain.	43
Figure 1.16. The structure of <i>S</i> -adenosylmethionine (SAM).	45
Figure 1.17. Reactive intermediate in the mechanism of the Fe(II)/ α -KG-dependant hydroxylases.	48
Figure 1.18. Reactive intermediates generated by haloperoxidases.	50
Figure 1.19. The general amino acid sequence assembly of FADH ₂ -dependent halogenases.	51
Figure 1.20. Intermediates in the mechanisms of O ₂ -dependent halogenases.	52
Figure 1.21. Structures of the lipopeptide antibiotics, A54145, CDA and Daptomycin.	53
Figure 1.22. Proposed mode of action for daptomycin.	54
Figure 1.23. Structures of the lipopeptide antibiotic CDA.	55
Figure 1.24. Organisation of the CDA NRPS.	56
Figure 1.25. Structure of the lipopeptide antibiotic daptomycin.	57
Figure 1.26. Organisation of the daptomycin NRPS.	57
Figure 1.27. Structures of the lipopeptide antibiotic A54145.	58
Figure 1.28. Organisation of the A54145 NRPS.	59
Figure 1.29. Structures of the lipoglycopeptide antibiotic ramoplanin A2 and enduracidin A.	60
Figure 1.30. Structures of the lipopeptide antibiotic enduracidin A and B.	62

Figure 1.31. Biosynthetic gene cluster of enduracidin.....	63
Figure 1.32. Organisation of the enduracidin NRPS.....	65
Figure 1.33. Structures of the glycolipopeptide ramoplanin A1, A2 and A3.	66
Figure 1.34. Biosynthetic gene cluster of ramoplanin.	67
Figure 1.35. Organisation of the ramoplanin NRPS.	69
Figure 1.36. Riboswitch mechanisms.	71
Figure 1.37. Thermosensing riboswitches.	73
Figure 2.1. Proposed biosynthetic pathway for OmAsp.	78
Figure 2.2. The LptL amino acid sequence alignment with AsnO and CAS.	80
Figure 2.3. Comparison of AsnO and LptL.	81
Figure 2.4. Agarose and polyacrylamide gel electrophoretic analyses of the gene <i>lptL</i> and protein LptL.....	82
Figure 2.5. RP-HPLC chromatogram of LptL reaction with L-Asn.	84
Figure 2.6. Mass spectra of the products from the LptL enzyme assays.	85
Figure 2.7. RP-HPLC chromatogram of LptL incubation with L-Asp and D-Asn.....	86
Figure 2.8. Proposed biosynthetic pathway for OmAsp.	88
Figure 2.9. The LptJ amino acid sequence alignment with SaSyrP, SvSyrP and TauD.	90
Figure 2.10. The agarose and polyarylamide gel electrophoretic analyses of the gene <i>lptJ</i> and protein LptJ.....	91
Figure 2.11. RP-HPLC spectra of LptJ reaction products.....	92
Figure 2.12. LC-MS analysis of daptomycin from LptJ reaction.	92
Figure 2.13. LC-MS analysis of CDA mixture from LptJ reaction.	93
Figure 2.14. The LptK amino acid sequence alignment with FkbM and RapI.....	95
Figure 2.15. The agarose and polyarylamide gel electrophoretic analyses of the gene <i>lptK</i> and protein LptK.	96
Figure 2.16. RP-HPLC spectra of LptK reaction products.	97
Figure 2.17. Catechol dioxygenase reaction.	99
Figure 2.18. Verification of the vector pMT3226 functionality by expression of XylE.	100
Figure 2.19. Schematic representation of the pMT3226- <i>lptJ</i> and pMT3226- <i>lptJK</i> constructs.....	101
Figure 2.20. Colony screening, catechol dioxygenase assay and PCR analysis of the ScpMT3226- <i>lptJ</i> and ScpMT3226- <i>lptJK</i> strains.	103
Figure 2.21. Bioassays of the ScpMT3226- <i>lptJ</i> and ScpMT3226- <i>lptJK</i> strains.....	104
Figure 2.22. LC-MS analysis of the ScpMT3226 fermentation products.....	106
Figure 2.23. LC-MS analysis of ScpMT3226- <i>lptJ</i> fermentation products.	107
Figure 2.24. LC-MS analysis of ScpMT3226- <i>lptJK</i> fermentation products.....	108

Figure 2.25. Schematic representation of the pIJ86- lptJ and pIJ86-lptJK constructs.	110
Figure 2.26. Colony screening and PCR analyses of the ScpIJ86-lptJ and ScpIJ86-lptJK strains.	110
Figure 2.27. Bioassays of the ScpIJ86-lptJ and ScpIJ86-lptJK strains.	111
Figure 2.28. LC-MS analysis of ScpIJ86 fermentation products.	112
Figure 2.29. LC-MS analysis of ScpIJ86-lptJ fermentation products.	113
Figure 2.30. LC-MS analysis of ScpIJ86-lptJK fermentation products.	114
Figure 3.1. Predicted enduracidin A analogues after modification by the tailoring enzymes Ram20 and Ram29.	118
Figure 3.2. Enduracidin activity test.	120
Figure 3.3. Enduracidin purification procedure.	121
Figure 3.4. RP-HPLC analysis of the <i>S. fungicidicus</i> crude extract.	122
Figure 3.5. LC-MS analysis of enduracidin A and B standards.	122
Figure 3.6. The sequence and PCR amplification of eGFP.	124
Figure 3.7. eGFP expression in SfpMS17-eGFP strain (A-C) and wild-type <i>S. fungicidicus</i> (D-F).	125
Figure 3.8. The upstream sequences of <i>ram20</i> , <i>ram29</i> and <i>eGFP</i>	125
Figure 3.9. Structures of chlorine-containing natural products.	126
Figure 3.10. Ram20 amino acid sequence alignment with known chlorinases.	128
Figure 3.11. PCR amplification and cloning of <i>ram20</i> into pMS17, and screening of the SfpMS17- <i>ram20</i> transformant strains.	131
Figure 3.12. MS analysis of SfpMS17- <i>ram20</i> fermentation products.	132
Figure 3.13. Enduracidin A analogues with different halogenation patterns.	133
Figure 3.14. Structures of the vancomycin-type, teicoplanin-type, and mannopeptimycin (epsilon) antibiotics.	137
Figure 3.15. Ram29 amino acid sequence alignment with Orf3, Dbv20 and Veg22.	139
Figure 3.16. Topology prediction of the protein Ram29.	140
Figure 3.17. A proposed two-component system for mannosylation of the glycopeptides.	141
Figure 3.18. PCR amplification and cloning of <i>ram29</i> into pMS17, and screening of the SfpMS17- <i>ram29</i> transformant strains.	142
Figure 3.19. MS analysis of the SfpMS17- <i>ram29</i> fermentation products.	143
Figure 3.20. Localisation of the glycosyl group on monoglycosylated- enduracidin A by tandem mass spectrometric analysis.	144
Figure 3.21. Monoglycosylated enduracidin A analogue.	144
Figure 3.22. Demannosylation of ramoplanin.	146

Figure 3.23. The mass spectrum for the product from the reaction of ramoplanin aglycon with the SfpMS17-ram29 membrane fraction.	147
Figure 4.1. The queuosine biosynthetic pathway.	152
Figure 4.2. PreQ ₁ riboswitches.	153
Figure 4.3. The preQ ₁ riboswitch crystal structure.	154
Figure 4.4. The preQ ₁ riboswitch gene regulation mechanism.	155
Figure 4.5. Introducing mutation sites into the preQ ₁ riboswitch sequence.	157
Figure 4.6. The <i>lacZ</i> expression profiles of variant preQ ₁ riboswitches.	159
Figure 4.7. The purine analogues tested by Roth et al. [170] for preQ ₁ riboswitch binding affinity.	160
Figure 4.8. The effect of the ligand preQ ₀ on <i>lacZ</i> expression controlled by the variant preQ ₁ riboswitches.	162
Figure 4.9. Synthetic analogues of preQ ₁ and preQ ₀	163
Figure 4.10. The effect of the ligand 2,6-diaminopurine on <i>lacZ</i> expression controlled by variant preQ ₁ riboswitches.	164
Figure 4.11. The effect of the ligand D6 on <i>lacZ</i> expression controlled by variant preQ ₁ riboswitches at 37 °C.	165
Figure 4.12. The effect of the ligand D6 on <i>lacZ</i> expression controlled by variant preQ ₁ riboswitches at 16 °C.	166
Figure 4.13. The effect of the ligand D9 on <i>lacZ</i> expression controlled by variant preQ ₁ riboswitches.	167
Figure 4.14. The dose-response curve and growth inhibition of the compound D9. .	167
Table 1.1. Relevant antibiotics in this study	20
Table 1.2. History of antibiotics in clinical use.	25
Table 1.3. Conserved core motifs of the catalytic domains in peptide synthetases	35
Table 1.4. β -hydroxyl amino acids in peptide antibiotics	47
Table 1.5. The four classes of halogenating enzymes	49
Table 1.6. Amino acid alignment between enduracidin and ramoplanin	60
Table 1.7. Antibiotic activity spectrum	61
Table 1.8. The enduracidin biosynthetic gene cluster	64
Table 1.9. The ramoplanin biosynthetic gene cluster	68
Table 2.1. Identification of LptL function by a Blast protein sequence search	79
Table 2.2. Subgroups of Fe(II)-dependent hydroxylases	80
Table 2.3. The substrate-binding pocket residues for the Asp- and OmAsp-	87
Table 2.4. Identification of LptJ function by a Blast protein sequence search	89
Table 2.5. Identification of LptK function by a Blast protein sequence search	94
Table 2.6. The conserved motifs in MTases	95

Table 3.1. Identification of Ram20 function by a Blast protein sequence search.....	127
Table 3.2. The substrate-binding pocket residues of Hpg-, Dpg- and Chp-activating A domains.....	130
Table 3.3. Identification of Ram29 function by Blast protein sequence searches	138
Table 4.1. <i>lacZ</i> repression ratios observed for the ligand D6 at lower temperatures.	166
Table 6.1. Bacterial strains used in this work.....	173
Table 6.2. Antibiotics used in this work.....	173
Table 6.3. Plasmids used in this work.....	174
Table 6.4. Target genes in this work	175
Table 6.5. Plasmids constructed in this work.....	175
Table 6.6. SDS-PAGE gel preparation.....	182
Table 6.7. PCR reaction components	186
Table 6.8. PCR amplification conditions	186
Table 6.9. List of primers for antibiotic engineering work	187
Table 6.10. List of primers for riboswitch engineering work	188

Word Count: 49788.

Abbreviations

ACP	Acyl Carrier Protein
A domain	Adenylation domain
AMP	Antimicrobial peptide
ATP	Adenosine-5'-triphosphate
α -KG	α -ketoglutarate
aTC	anhydro-tetracycline
bp	base pair
BSA	bovine serum albumine
CDA	Calcium Dependent Antibiotic
C domain	Condensation domain
CoA	Coenzyme A
Da	dalton
DMSO	Dimethylsulfoxide
<i>E. coli</i>	<i>Escherichia coli</i>
E domain	Epimerase domain
Gtf	Glycosyl transferase
FPLC	Fast Protein Liquid Chromatography
HPLC	High Pressure Liquid Chromatography
AsnO	Asn oxygenase
IPTG	Isopropyl β -D-1-thiogalactopyranoside
Kan	Kanamycin
LC-MS	Liquid Chromatography-Mass Spectrometry
MTase	Methyl transferase
MRSA	Methicillin resistant <i>Staphylococcus aureus</i>
NRP	Non-Ribosomal Peptide
NRPS	Non-Ribosomal Peptide Synthetase
ONPG	<i>O</i> -nitrophenyl- β -galactoside
OD _x	Optical density at x nm
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PCP domain	Peptidyl Carrier Protein domain
PCR	Polymerase chain reaction
PK	Polyketide
PITC	Phenylisothiocyanate
PLP	Pyridoxal 5'-Phosphate
Ppant	4'-phosphopantetheine

PPTase	Phosphopantetheinyl transferase
RP-HPLC	Reverse phase-High Pressure Liquid Chromatography
SAH	<i>S</i> -Adenosyl-L-homocysteine
SAM	<i>S</i> -adenosyl-methionine
SD sequence	Shine-Dalgarno sequence
SDS	Sodium dodecylsulphate
TE domain	Thioesterase domain
UTR	Untranslated region
UV	Ultraviolet
VRE	Vancomycin-resistant <i>Enterococcus faecium</i>

Amino Acids:

Amino acids	One letter designation	Three letters designation
Alanine	A	Ala
Arginine	R	Arg
Asparagine	N	Asn
Aspartic acid	D	Asp
Cysteine	C	Cys
Glutamine	Q	Gln
Glutamic acid	E	Glu
Glycine	G	Gly
Histidine	H	His
Isoleucine	I	Ile
Leucine	L	Leu
Lysine	K	Lys
Methionine	M	Met
Phenylalanine	F	Phe
Proline	P	Pro
Serine	S	Ser
Threonine	T	Thr
Tryptophan	W	Trp
Tyrosine	Y	Tyr
Valine	V	Val

Non-proteinogenic amino acids

Citrulline	Cit
3,5-dichloro-4-hydroxyl phenylglycine	Dpg
Enduracididine	End
Hydroxyl asparagine	hAsn
Hydroxyl phenylglycine	Hpg
Kynurenine	Kyn
3-methyl glutamic acid	3mGlu
3- <i>O</i> -methyl aspartic acid	OmAsp
Ornithine	Orn
phospho-hydroxyl asparagine	phAsn
Sarcosine	Sar

Three letters designation**Nucleotides:**

Adenine	A
Cytosine	C
Guanine	G
Thymine	T
Uracil	U

Abstract

The University of Manchester

Ming-Cheng Wu

Doctor of Philosophy (Ph. D.)

Bio-engineering of Antibiotic Enduracidin Biosynthetic Pathway and PreQ₁ Riboswitch

September 2011

Non-ribosomally synthesised natural products derived mainly from bacteria and fungi act as important therapeutic agents. Due to their complex structures it is difficult to chemically synthesise such compounds, therefore the engineering of biosynthetic and biocatalytic pathways that are vital for their production are the aims of this thesis.

The first project involved the study of the biosynthesis of 3-*O*-methyl aspartic acid (OmAsp) in the antibiotic A54145. We demonstrated that LptL functions as an asparagine hydroxylase. We also predicted that LptJ and LptK are involved in the biosynthesis of OmAsp, although we did not find evidence of this non-standard amino acid in the antibiotic CDA when we over-expressed both proteins in *Streptomyces coelicolor*.

The second project involved the study of the chlorination and mannosylation of hydroxyphenyl glycine (Hpg) residues in ramoplanin. We over-expressed the putative chlorinase and mannosyl transferase separately in the enduracidin-producer, in which the production of enduracidin has been characterised. We then found that trichlorinated and mannosylated enduracidin analogues were produced.

The final project involved the re-engineering of a preQ₁ riboswitch. We successfully created ten preQ₁ riboswitch mutants fused to the reporter gene *lacZ* in *Bacillus subtilis*, all of which showed different levels of β -galactosidase activity. Subsequently, we found that the preQ₁ C17U mutant riboswitch can respond specifically to the synthetic ligands 5-(aminomethyl)furo[2,3-*d*]pyrimidine-2,4-diamine (**D6**) and 2,4-diamino-7H-pyrrolo[2,3-*d*]pyrimidine-5-carbonitrile (**D9**) to control gene expression in a dose dependent manner.

The results described here show the successful production of variant antibiotics by bio-engineering and the use of an engineered preQ₁ riboswitch as a tool for regulating gene expression.

Declaration

I declare that no portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright Statement

- i. The author of this thesis (including any appendices and/or schedules to this thesis) owns any copyright in it (the “Copyright”) and he has given The University of Manchester the right to use such Copyright for any administrative, promotional, educational and/or teaching purposes.
- ii. Copies of this thesis, either in full or in extracts, may be made **only** in accordance with the regulations of the John Rylands University Library of Manchester. Details of these regulations may be obtained from the Librarian. This page must form part of any such copies made.
- iii. The ownership of any patents, designs, trade marks and any and all other intellectual property rights except for the Copyright (the “Intellectual Property Rights”) and any reproductions of copyright works, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property Rights and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property Rights and/or Reproductions.
- iv. Further information on the conditions under which disclosure, publication and exploitation of this thesis, the Copyright and any Intellectual Property Rights and/or Reproductions described in it may take place is available from the Head of School of Chemistry.

Acknowledgements

I would like to deeply thank my supervisor Prof. Jason Micklefield for his tireless guidance, continual encouragement and constant support. I also thank him for giving me the opportunity to study for a PhD in his group.

I would also like to thank all the past and present members of the Micklefield group. In the antibiotic engineering work, I sincerely appreciate the technical help with Mass Spectrometric analysis from Laura Nunns and Matthew Styles. In the riboswitch work, I would like to thank Dr. Neil Dixon for his kind advice and encouragement, Dr. Chris Robinson for thesis proof-reading and Phillip Lowe for riboswitch ligand synthesis. A special thanks to Brian Law, who spent much time with me on my thesis correction. Also thanks to Happy-man Chinnan Karthikeyan, his smile accompanied and encouraged me everyday. Thanks also to Dr. Lu Shin Wong for thesis proof-reading and his encouragement and advice on my work, and other members Dr. Mitchell Jacobs, Dr. John Duncan, Matthew Wilding, Christopher Chesters, Michael Power, Mark Goodall, Sarah Shepherd and Ross Lewin for your company and help. It was a pleasure to work with all of you.

My thanks go to Dr. Fiona Flett for her practical advice and knowledge on *Streptomyces* biology and Dr. Sheona P. Drummond for helping me with fluorescence microscopy. Thank you also to the School of Chemistry for my overseas scholarship.

Finally words cannot express my deep gratitude and appreciation for my girlfriend and family members. My girlfriend Yu-Ting Lin sacrificed her sleeping time everyday in Taiwan to talk with me and has stood by me on my pursuit of a PhD. My parents and brother have provided constant support and encouragement. Without you I would not have been able to achieve this work, so thank you.

Ming-Cheng Wu

Manchester

September 2011

The Author

The author completed his Master degree in Insect Physiology and Biochemistry from the National Chung-Hsing University, Taiwan in July 2004. After working as a research assistant for Prof. Hueih-Ming Chen and Kuang-Hui Lu, the author undertook a PhD under the supervision of Prof. Jason Micklefield from September 2008 to September 2011.

Chapter 1. Introduction

Microbial metabolites are important and valuable resources for humans. Many of these metabolites have been discovered, manipulated and developed as therapeutics. Antibiotics, one class of metabolites, play an important role in the defence against bacterial infection in the medical system. However abuse from overprescription has encouraged the rise of resistant pathogens which are a serious threat to human health. For example, in 2010 a bacterium containing a new metallo- β -lactamase was found to be resistant to all known antibiotics [1]. Therefore there is an urgent need to develop new classes of antibiotics or engineer antibiotic synthesis/biosynthesis pathways to produce variant antibiotics. This work is concerned with five antibiotics (Table 1.1) which are biosynthesised by non-ribosomal peptide synthetases (NRPSs). Three of them, A54145 [2], Calcium Dependant Antibiotic (CDA) [3] and daptomycin [4], are produced by *Streptomyces fradiae*, *Streptomyces coelicolor* and *Streptomyces roseosporus* respectively, and share similar structures (Chapter 2). The other two, enduracidin [5] and ramoplanin [6] are produced by *Streptomyces fungicidicus* and *Actinoplanes* strain ATCC33076 respectively, and also share similar structures with each other (Chapter 3). Daptomycin (Cubicin; Cubist Pharmaceuticals), a new class of natural antibiotic, was approved by the U.S. Food and Drug Administration (FDA) in 2003 for the treatment of infections caused by Gram-positive bacteria [7]. Ramoplanin is currently in Phase III clinical trials and also has a broad-spectrum activity against Gram-positive bacteria [6].

Table 1.1. Relevant antibiotics in this study

Antibiotic	Origin	Peptide size (amino acids)	Reference
CDA	<i>S. coelicolor</i>	11	Hojati <i>et al.</i> (2002) [3]
A54145	<i>S. fradiae</i>	13	Miao <i>et al.</i> (2006) [2]
Daptomycin	<i>S. roseosporus</i>	13	Miao <i>et al.</i> (2005) [4]
Ramoplanin	<i>Actinoplanes</i> strain ATCC33076	17	McCafferty <i>et al.</i> (2002) [6]
Enduracidin	<i>S. fungicidicus</i>	17	Yin and Zabriskie (2006) [5]

In an effort to improve the regulation of recombinant gene expression, and provide potential applications to biosynthetic pathways, the development of artificial riboswitches was also explored. Here, the work on riboswitches aims to re-engineer a transcriptional “OFF” preQ₁ riboswitch to be regulated by synthetic non-natural molecules (Chapter 4).

1.1. Natural products — Metabolites

Natural products from microbes and plants have long been of great interest to bio-industry and medicine [8]. Such natural products are the metabolites from cellular metabolism and can be subdivided into two major categories- **primary metabolites** and **secondary metabolites** [9]. Primary metabolites such as amino acids, nucleotides, vitamins, carbohydrates and fatty acids can be assembled into complex products that are essential for cellular growth, development and reproduction. Additionally, the by-products of primary metabolism such as ethanol, acetone and butanol are important industrial commodities. Secondary metabolites are compounds that are non-essential for the growth of organisms, but which aid in the survival of organisms in terms of competing with other species and combating stresses from their natural environments. The secondary metabolites have specific biological activities including antibiotic, immunosuppressant and antitumor effects, and play major roles in modern therapy [8, 10]. Based on structure, they can be classified into six classes: polyketides and fatty acids, terpenoids and steroids, phenylpropanoids, alkaloids, non-proteinogenic amino acids and peptides [9].

In the cell, the mechanisms which control metabolism are vital. They allow conservation of energy and enable organisms to grow in environments with limited resources. Generally, the regulation mechanisms can be engaged at the gene or protein level. At the gene level, transcription and translation are mostly governed by special genetic elements. For example, in DNA, **promoter** sequences which normally lie at the -10 and -35 positions upstream of the transcription start site, play an important role in transcription regulation by the binding and recognition of cellular transcription factors, RNA polymerases and nucleosomes [11] (Figure 1.1). A typical bacterial RNA transcript is composed of three regions: the 5'-untranslated region (5'-UTR), the

protein coding region that begins with the start codon (typically AUG) and ends with the stop codon (UAA, UGA or UAG) and the 3'-untranslated region (3'-UTR). The **Shine-Dalgarno (SD)** sequence which is generally located approximately 8 base-pairs upstream of the start codon in the 5'-UTR of the RNA transcript, is a purine-rich region which helps to recruit the ribosome for initiation of protein synthesis [12] (Figure 1.1). Additionally in some cases, specific elements in the 5'-UTR of the RNA transcript have been found to bind metabolites and affect gene expression; these are known as **riboswitches** [13] (Figure 1.1). At the protein level, there are many kinds of substrates, cofactors and modifications which can govern protein activity.

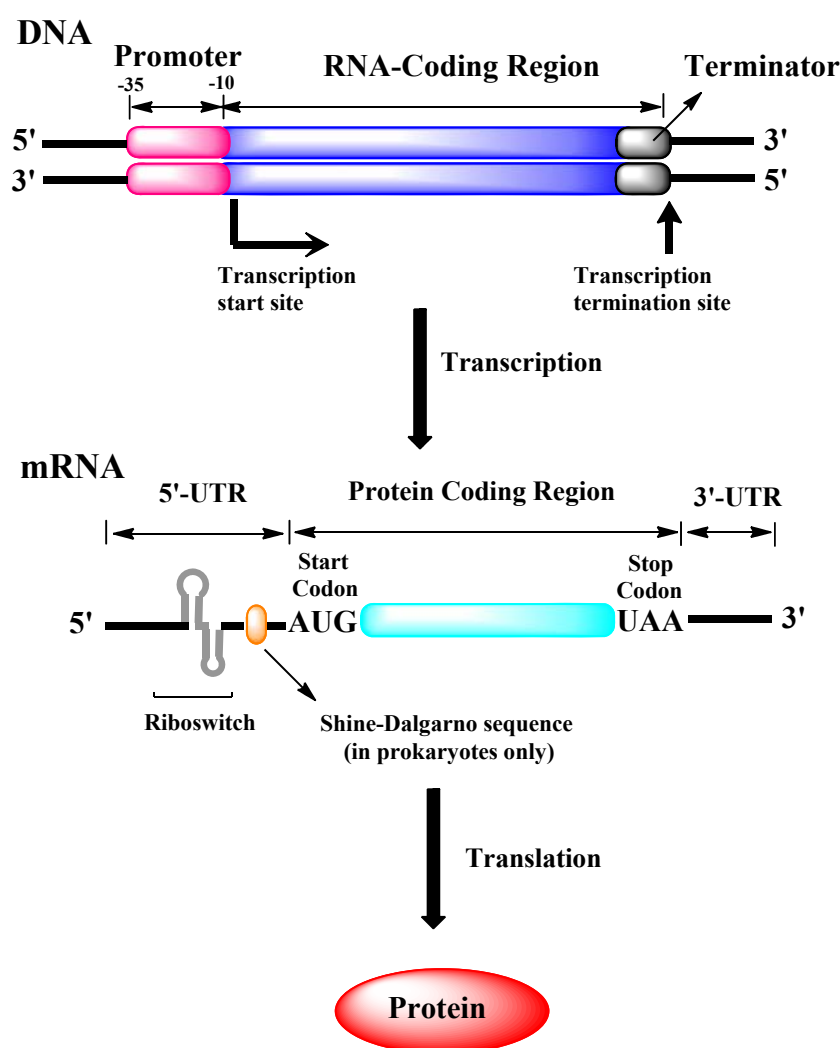


Figure 1.1. A transcription unit and mRNA transcript structure.

A transcription unit is composed of a promoter, RNA-coding region and terminator. The RNA transcript is synthesised by an RNA polymerase recruited at the promoter, and is complementary to the sequence of the template DNA strand between the transcription start site and the transcription termination site.

1.2. The Secondary Metabolite Producing Actinomycetes

The actinomycetes are a group of Gram-positive bacteria with high GC content genomes, diverse morphological characteristics, complex life cycles and diverse ecology [14]. Their life cycle characteristics are similar to that of filamentous fungi in many ways, for instance, they can produce spores that allow them to disperse widely in the environment and they form mycelia with filamentous structures [14] (Figure 1.2). Most actinomycetes are non-pathogenic and play a key role in soil biodegradation *via* the excretion of enzymes into the soil, resulting in denitrification and nutrient recycling [14, 15]. For example, the genus *Rhodococcus* has been studied intensively over the past two decades and is the most promising organism for the biodegradation of nitro aromatic compounds and organic nitriles — the most widespread environmental pollutants [16]. Additionally actinomycetes are well known producers of biologically active secondary metabolites including a great number of antibiotics. So far two-thirds of the known antibiotics have been contributed by actinomycete species, most of which (approximately 80 %) are produced from the genus *Streptomyces* [14].

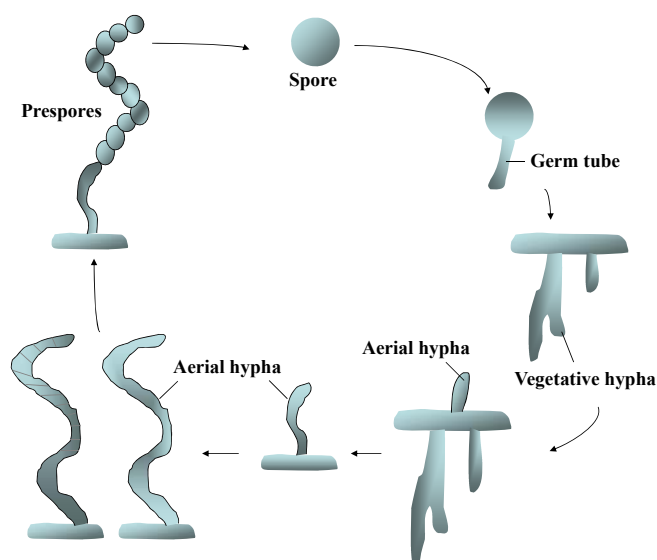


Figure 1.2. The developmental life cycle of actinomycetes.

When spores encounter a suitable environment, they germinate by growing germ tubes to form vegetative hyphae that normally penetrate into the substrates. In the presence of nutrient depletion or other physiological stresses, they start to grow into the air to form aerial hyphae. Subsequently, aerial hyphae form long chains of prespores which mature into spores and are dispersed into the environment.

1.3. Antibiotics

Following the discovery of penicillin in 1929 a systematic search for antibiotics was launched, and during the period from 1940 to 1960, the golden age of antibiotics, at least 10 antibiotics were introduced into clinical use [17] (Table 1.2). The introduction of antibiotics for medicinal applications has provided effective control of infectious diseases, resulting in a decline in death rates. However, the overuse of antibiotics, along with the fact that no new structural classes of antibiotics have been introduced into clinical use during the 30-year interval between 1970 and 2000, has seen the occurrence of resistant bacteria [17-19]. A recent report suggests that bacteria containing a new transmissible genetic element conferring resistance to all known antibacterial agents, have been spreading globally in 2010 [1]. This bacterium was first found in New Delhi and contained a novel metallo- β -lactamase, so named the **New Delhi Metallo- β -lactamase 1 (NDM-1)**. The term NDM-1 does not refer to a single bacterial species but to all bacteria which contain a transmissible genetic element encoding metallo- β -lactamase 1 [1]. Moreover, in past decades the discovery of new antibiotics from natural products has been slowed due to the high incidence of the re-discovery of known natural products [20]. Therefore, exploitation of pathways for the discovery of new antibiotics or the biosynthesis of new variant antibiotics is essential.

Table 1.2. History of antibiotics in clinical use

Year	Antibiotic	Class	Target site
1929	Penicillin	β -lactam	Cell wall synthesis
1932	Sulfapyridine	Sulfonamide	Metabolite synthesis
1944	Streptomycin	Aminoglycoside	Protein synthesis
1945	Cephalosporin	β -lactam	Cell wall synthesis
1947	Chloramphenicol	Phenylpropanoid	Protein synthesis
1948	Chlortetracycline	Tetracycline	Protein synthesis
1950	Erythromycin	Macrolide	Protein synthesis
1955	Vancomycin	Glycopeptide	Cell wall synthesis
1955	Virginiamycin	Streptogramin	Protein synthesis
1955	Amphotericin	Polyene	Cell membrane
1955	Lincomycin	Lincosamide	Protein synthesis
1959	Rifamycin	Ansamycin	RNA polymerase
1962	Nalidixic acid	Quinolone	DNA replication
1969	Fosfomycin	Phosphonate	Cell wall synthesis
2000	Linezolid	Oxazolidinone	Protein synthesis
2003	Daptomycin	Lipopeptide	Cell wall

Note: All antibiotics are from natural products except sulfapyridine, nalidixic acid and linezolid which are synthetic products. Adapted and reproduced from Walsh and Wright (2005) [17].

Approximately 1500 microbial genomic sequences, including that of *S. coelicolor*, have been published and the sequencing of about 5000 bacterial genomes is in progress (<http://www.ncbi.nlm.nih.gov/genomes/lproks.cgi>, accessed June 2011). They provide the foundations for genomic mining and understanding of the biosynthetic machinery for microbial natural products, opening up new routes for the discovery of natural products [21-26]. Depending on structural diversity, antibiotics can be classified into several groups: peptide antibiotics (ribosomal and non-ribosomal), polyketides, aminoglycosides and other small compounds [27]. The following sections will focus mainly on the non-ribosomal peptide antibiotics.

1.4. The Target Sites of Antibiotics

The target sites for desirable new antibiotics should be essential for bacterial survival and have no counterpart in eukaryotic cells, for example, steps in the bacterial-specific synthesis of nucleic acids, proteins, essential metabolites and the bacterial cell wall [17, 28]. However the characteristics of bacteria, such as their short generation time and high mutation rate, drive the selection of bacteria with acquired defence mechanisms. They adopt strategies such as enzymatic degradation of the antibiotics, structural alteration of the antibiotic target sites, or increased efflux of antibiotics by transporter systems to produce resistance [17, 29]. In addition, horizontal gene transfer allows acquired resistance to be transferred between various species (NDM-1 cases) [1, 17, 28, 30].

The antibiotic target sites can be changed by altering the antibiotic structure. For example, the original vancomycin target site is the transpeptidation step of peptidoglycan synthesis which can be changed to the transglycosylation step by adding a chlorobiphenyl group to the disaccharide moiety of vancomycin [31]. Therefore it is crucial to understand the mechanism of action for antibiotics and their associated resistance mechanisms, to develop new and promising antibiotics and reduce the probability of antibiotic resistance. Examples of target sites and representative antibiotics are outlined above in Table 1.2.

1.4.1. Nucleic Acid Synthesis

The processes and molecular machineries for DNA replication and RNA synthesis in prokaryotes are significantly different from those of eukaryotes. For example, in DNA replication, prokaryotes employ a single origin of replication rather than the multiple replication origin strategies adopted by eukaryotes [32, 33]. Additionally, enzymes for DNA replication in prokaryotes and eukaryotes are completely different. One example is gyrase, which is primarily present in prokaryotes and is thought to control DNA topology during DNA replication and transcription. The quinolone antibiotic targets gyrase and thereby interferes with the prokaryotic DNA replication and transcription

processes [34]. Regarding RNA synthesis, antibiotics can bind to the β -subunit of prokaryotic RNA polymerase, leading to inhibition of this enzyme. For example rifampin, an important drug for the treatment of tuberculosis adopts this mode of action [35].

1.4.2. Protein Synthesis

There are many classes of antibiotics developed to target bacterial protein synthesis, a consequence of the differences between the prokaryotic 70S ribosome (comprising 30S and 50S subunits) and the eukaryotic 80S ribosome (comprising 40S and 60S subunits) structures. The phenylpropanoid, macrolide, lincosamide, streptogramin and oxazolidinone antibiotic classes can bind to the prokaryotic 50S ribosomal subunit, interfering with protein translation [29]. The aminoglycoside class of antibiotics, such as neomycin, kanamycin and spectinomycin, interferes with the decoding of mRNA by binding to the A site of the prokaryotic 30S ribosomal subunit [36] whilst the tetracycline class of antibiotics prevents the docking of amino-acylated tRNA by binding to the 30S ribosomal subunit [37]. Thus all of these antibiotic classes kill bacteria by blocking the bacterial protein synthetic pathway.

1.4.3. Essential Metabolic Pathway

A well characterised essential metabolic pathway targeted by antibiotics is folic acid synthesis. Folic acid is a vital cofactor in both prokaryotic and eukaryotic cells for the biosynthesis of a diverse range of cellular metabolites such as purines, thymidine, methionine and other important molecules [38, 39]. The significant distinction between prokaryotes and mammalian cells is that bacteria synthesise their own folic acid, whereas mammalian cells must obtain it from their diet. This makes the folic acid synthetic pathway a good target site for antibiotics. For example, the antibiotic sulfonamide disrupts the biosynthesis of folic acid by inhibiting the conversion of *p*-aminobenzoic acid to dihydropteroate, resulting in the inhibition of nucleic acid and protein synthesis and interruption of bacterial growth [40].

1.4.4. Bacterial Cell Wall

The bacterial cell wall, composed mainly of a polymeric mesh of peptidoglycan (also called murein), is the most distinct component from animal cells [41]. Therefore it is a good candidate target site for antibiotics. Antibiotics target the bacterial cell wall by two mechanisms: one is disruption of the structure and function of the cell wall, whilst the other is inhibition of cell wall biosynthesis. For example, the antibiotic daptomycin binds to the surface of the cell wall, resulting in the formation of pores which cause depolarization of the cell [42] (see section 1.8). The antibiotics fosfomicin [43], enduracidin [44], ramoplanin [44], mannopeptimycin [45], penicillin [46] and moenomycin [47] interfere with different steps of bacterial cell wall biosynthesis.

The cell wall peptidoglycan is biosynthesised in three stages: (1) synthesis of the UDP- *N*-acetylmuramic acid (MurNAc) pentapeptide, (2) formation of lipid I and lipid II, and (3) cross-linking of lipid II [48, 49]. The first and second stages of peptidoglycan biosynthesis occur in the cytoplasm whilst the cross-linking of lipid II takes place on the cell membrane surface [50]. During the first stage, six enzymes (MurA-F) are involved in the synthesis of UDP-MurNAc pentapeptide [50] (Figure 1.3). MurA is responsible for transferring a pyruvyl group to UDP-*N*-acetylglucosamine (GlcNAc) to form the 3-enolpyruvyl ether derivative. Then a reduction is performed by MurB using NADPH to form UDP-MurNAc. Next, ATP-dependent ligases (MurC, D, E and F) are responsible for the respective additions of L-Ala, D-Glu, L-Lys and the dipeptide D-alanyl-D-alanine to UDP-MurNAc, forming the UDP-MurNAc pentapeptide. The antibiotic fosfomicin interferes with MurA activity, resulting in an inhibition of cell wall biosynthesis [43].

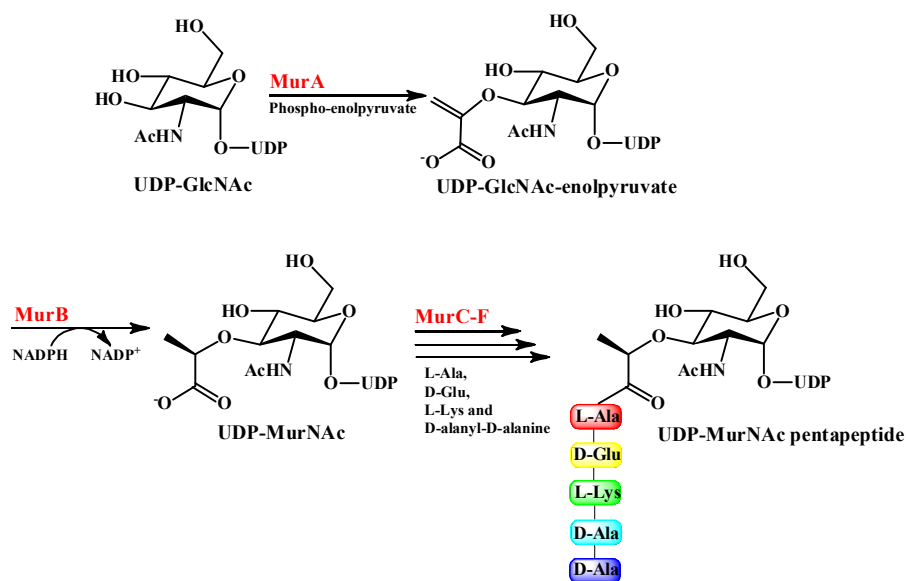


Figure 1.3. UDP-MurNAc-pentapeptide biosynthesis.

UDP-GlcNAc-enolpyruvate is synthesised by MurA, following by a reduction carried out by MurB to form UDP-MurNAc. Then a sequential addition of five amino acids to UDP-MurNAc is carried out by four enzymes (MurC-F) to form UDP-MurNAc pentapeptide.

During the second stage, two enzymes, MraY and MurG are responsible for the formation of lipid I and II respectively [48-50] (Figure 1.4). The phospho-MurNAc pentapeptide moiety is transferred from the UDP-MurNAc pentapeptide to the membrane acceptor undecaprenyl phosphate by the transferase MraY, yielding lipid I. The transferase MurG then adds GlcNAc from UDP-GlcNAc to the MurNAc moiety of lipid I, forming lipid II. At this second stage, antibiotics such as tunicamycin and liposidomycin can bind to the active site of MraY to inhibit the biosynthesis of lipid I [51].

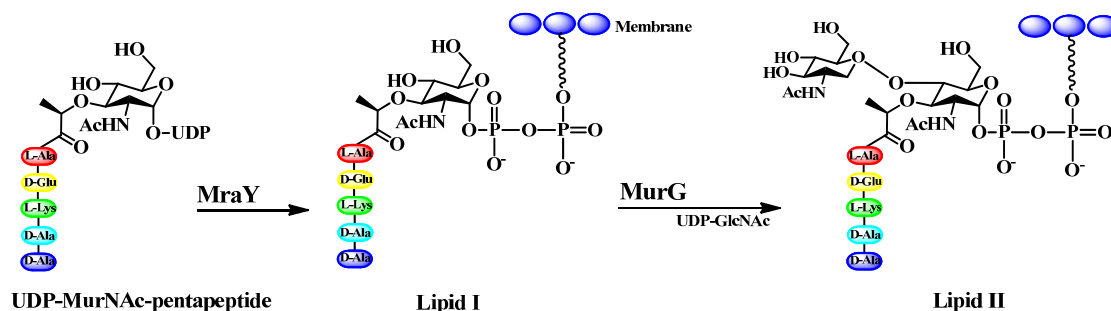


Figure 1.4. Lipid I and Lipid II biosynthesis.

The lipid II is then translocated from the cytoplasm to the cell membrane surface and used as the substrate for the polymerisation at the third stage of cell wall synthesis [49]. Transglycosylation and transpeptidation of lipid II occur on the cell surface to form the peptidoglycan [49]. Transglycosylation, which is catalysed by glycosyltransferases, forms a glycosidic bond between the GlcNAc moiety of lipid II and the MurNAc-pentapeptide moiety of another lipid II [48] (Figure 1.5). Transpeptidation is carried out by transpeptidases which can form a new amide bond between the ϵ -amine of the lysine residue in the pentapeptide and the amide carbonyl of the alanine on a neighboring pentapeptide [48] (Figure 1.6). The antibiotics ramoplanin and enduracidin can bind to the lipid II, which are the substrates for the transglycosylase and transpeptidase, and inhibit the biosynthesis of cell wall indirectly [44]. The antibiotics penicillin and moenomycin can bind to transpeptidase and transglycosylase respectively, resulting in an inhibition of peptidoglycan formation. Cell wall biosynthesis is illustrated in Figure 1.7; the pathway begins with the formation of lipid I from UDP-MurNAc-pentapeptide and continues to the polymerised and cross-linked peptidoglycan.

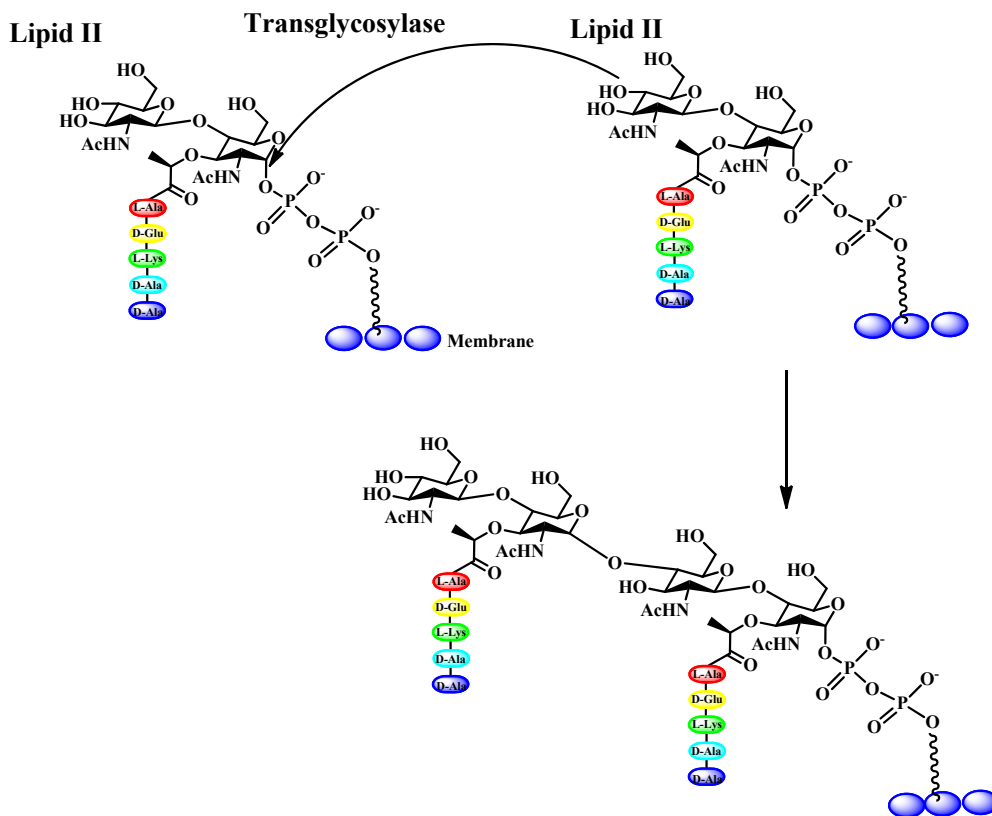


Figure 1.5. Transglycosylation of lipid II by transglycosylase.

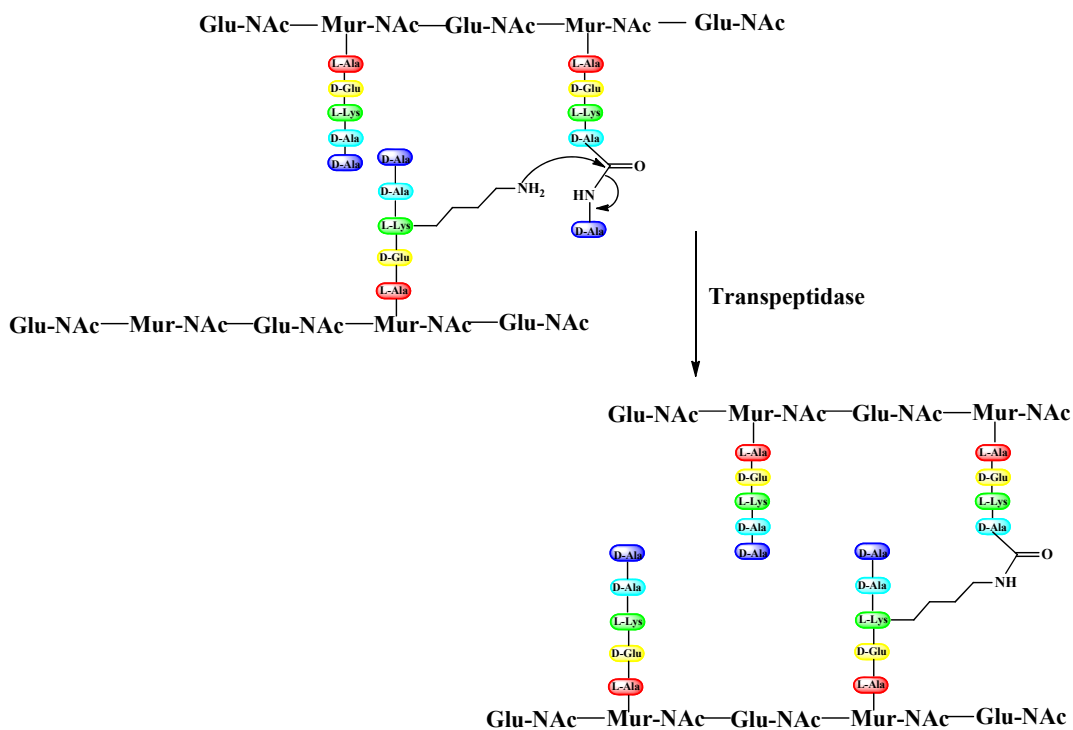


Figure 1.6. The transpeptidation reaction of transpeptidase.

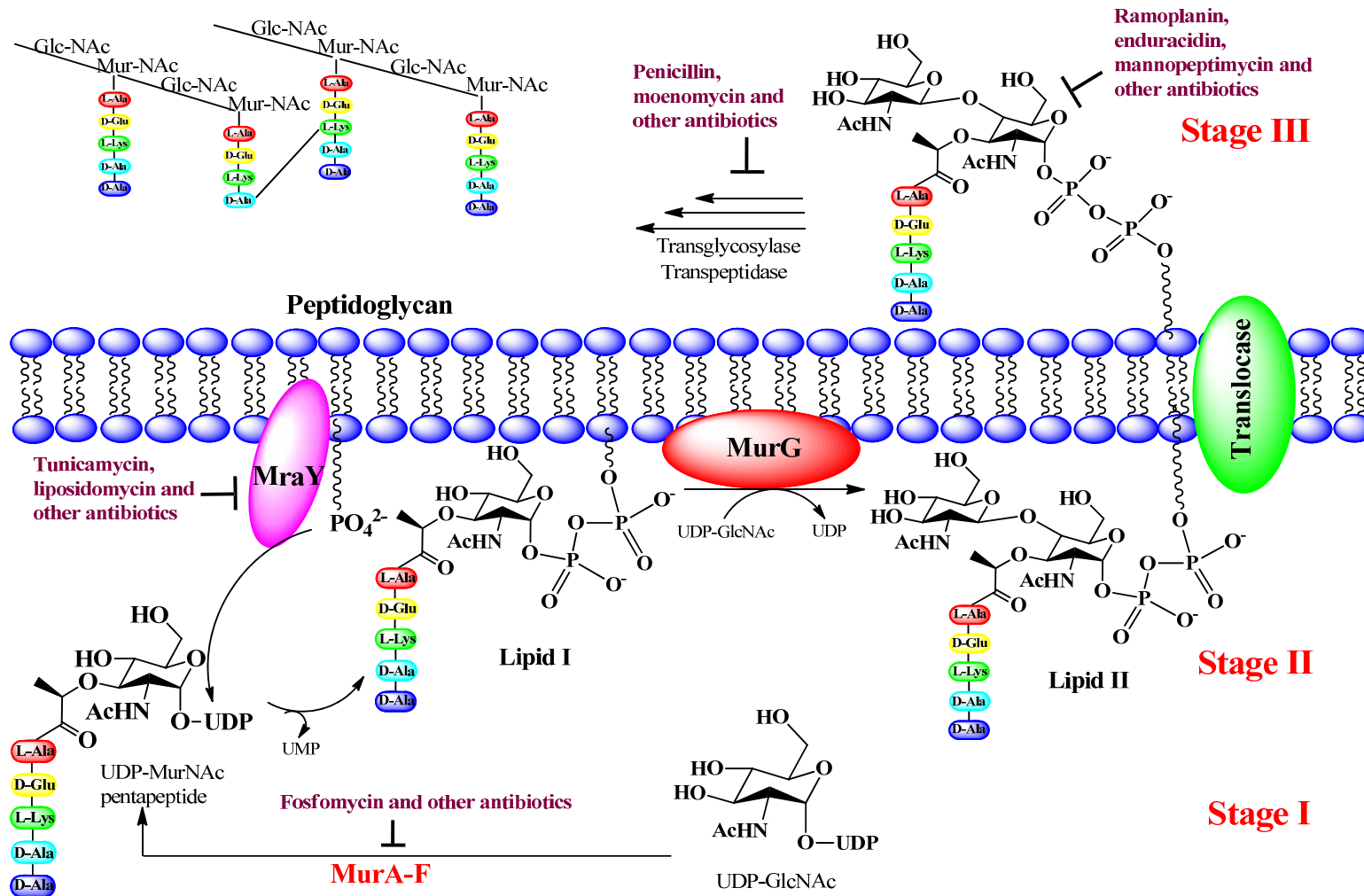


Figure 1.7. The peptidoglycan biosynthetic pathway with sites of antibiotic interference highlighted.

1.5. Antibiotic Production and Regulation in *Streptomyces*

Antibiotics from *Streptomyces* are chromosomally-determined, stemming from gene clusters that encode for the proteins required for synthesis, modification, regulation, resistance and transport of the antibiotic. For example, the antibiotic CDA is derived from 40 open reading frames (ORFs) of a biosynthetic gene cluster [3]. The onset of antibiotic biosynthesis is controlled by regulatory proteins called transcriptional activators, which bind to promoter elements [52]. Some of the best-characterised regulatory proteins are the *Streptomyces* Antibiotic Regulatory Proteins (SARPs), which can be classified into two categories: **pathway-specific regulators** that regulate production of a specific antibiotic; and **pleiotropic regulators** that regulate the production of more than one antibiotic [52, 53]. For example in *S. coelicolor*, the pathway-specific regulators CdaR, ActII-ORF4 and RedZ activate the production of the antibiotics CDA, actinorhodin and undecylprodigiosin respectively [54-57]. All of these pathway-specific regulators may be regulated by a pleiotropic regulator, for example the two-component system AbsA, comprised of the sensor kinase AbsA1 and the response regulator AbsA2 [52, 53]. AbsA1 is activated by environmental signals and negatively regulates genes *via* phosphorylation of AbsA2. Phosphorylated AbsA2 interacts with the promoter of pathway-specific regulator genes to repress the expression of activators such as CdaR, ActII-ORF4 and RedZ, resulting in inhibition of antibiotic production [53, 57] (Figure 1.8).

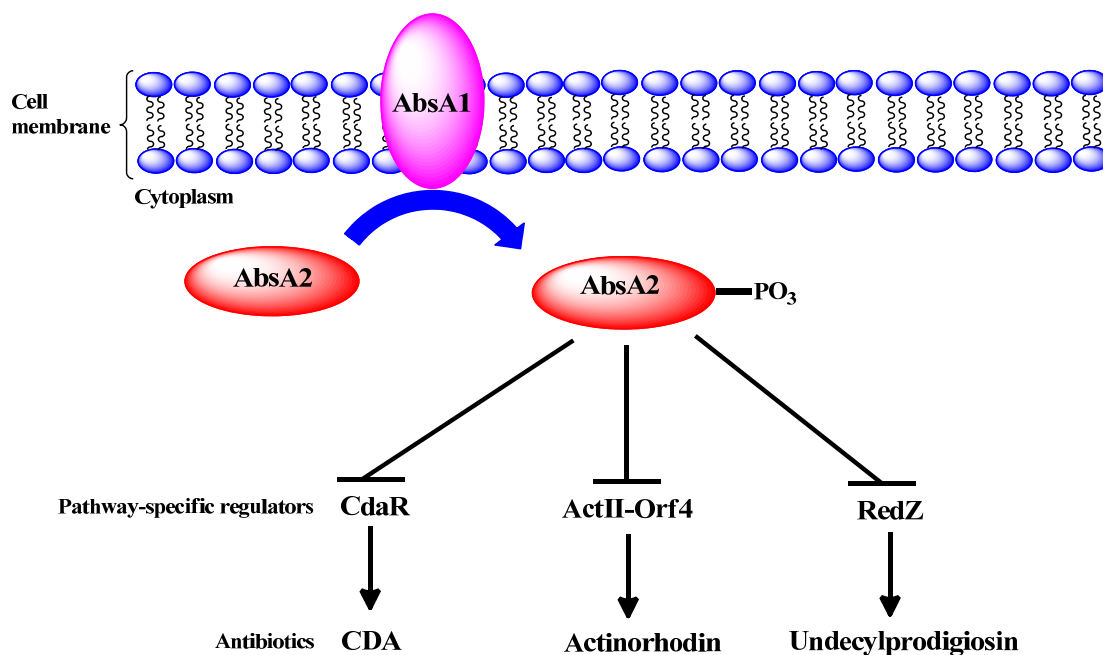


Figure 1.8. The AbsA two-component signal transduction system in *S. coelicolor*. The sensor kinase AbsA1 phosphorylates AbsA2. Phosphorylated AbsA2 represses transcription of *cdaR*, *actII-ORF4* and *redZ*, preventing the production of the antibiotics CDA, actinorhodin and undecylprodigiosin, respectively. Adapted from McKenzie and Nodwell (2007) [57].

Antibiotic production in *Streptomyces spp.* is growth phase-dependent and can be influenced by physiological factors such as growth rate, imbalances in metabolism and other physiological stresses [14]. Additionally, environmental factors such as the sources of carbon, nitrogen and phosphate also play important roles in antibiotic fermentation; they can trigger the initiation of signaling cascades which ultimately lead to antibiotic production [14, 58]. Taking the carbon source as an example, a regulatory mechanism termed “carbon catabolite regulation”, guides the bacterial metabolism pathway [58, 59]. When a preferential carbon source such as glucose is present in the medium, the cell catabolises it and directs the primary metabolic pathways to produce energy for growth. Simultaneously, the secondary metabolic pathways are repressed. Antibiotics such as actinomycin, kanamycin and neomycin have been confirmed to have their production repressed by the presence of glucose [59]. For this reason polysaccharides (*e.g.* starch), oligosaccharides (*e.g.* lactose) and oil (*e.g.* soybean oil) are preferable for the fermentation production of antibiotics [58].

1.6. Non-Ribosomal Peptide Antibiotic Synthesis

Non-ribosomal peptide (NRP) antibiotics are synthesised by large multi-functional enzymes (0.15-1.5 MDa) termed non-ribosomal peptide synthetases (NRPSs), which are mainly found in bacteria and fungi [60]. NRPSs exhibit a modular organisation where each module of the NRPS is responsible for the incorporation of a single amino acid into a peptide product chain [60]. The modules can be further sub-divided into domains, *i.e.* the Adenylation (A) domain, Peptidyl Carrier Protein (PCP) domain and Condensation (C) domain, with the final module normally containing a peptide releasing Thioesterase (TE) domain [60, 61]. Each domain can be identified by conserved protein sequence motifs [62] (Table 1.3). Each domain, except the PCP domain, represents an enzymatic unit that catalyses an individual step during non-ribosomal peptide synthesis [62, 63].

Table 1.3. Conserved core motifs of the catalytic domains in peptide synthetases

Domain	Core	Consensus sequence
Adenylation	A1	L(TS)YxEL
	A2	LKAGxAYL(VL)P(LI)D
	A3	LAYxxYTSG(ST)TGxPKG
	A4	FDxS
	A5	NxYGPTE
	A6	GELxIxGxG(VL)ARGYL
	A7	Y(RK)TGDL
	A8	GRxDxQVKIRGxRIELGEIE
	A9	LPxYM(IV)P
	A10	NGK(VL)DR
Peptidyl Carrier Protein	PCP	DxFFxxLGG(HD)S(LI)
Condensation	C1	SxAQxR(LM)(WY)XI
	C2	RHExLRTxF
	C3	MHHxISDG(WV)S
	C4	YxD(FY)AVW
	C5	(IV)GxFVNT(QL)(CA)xR
	C6	(HN)QD(YV)PFE
	C7	RDxSRNPL
Thioesterase	TE	G(HY)SxG

Note: 'x' stands for any amino acid residue. Adapted from Marahiel *et al.* (1997) [62].

A hypothetical NRP synthesis by an NRPS with three modules is illustrated in Figure 1.9 and details of the processes taking place at each domain are given in following sections [63]. Before the NRPS starts peptide synthesis, it needs to be activated by the addition of a thiol moiety, *i.e.* the 4'-phosphopantetheine (ppant) moiety from coenzyme A (CoA), to the PCP domains, transforming the enzyme complex into the active *holo*-NRPS (Figure 1.9, A₁. priming). If the PCP domains are misprimed with the acyl-S-ppant moiety from using acyl-CoA as a substrate (Figure 1.9, A₂. mispriming), a proofreading type II thioesterase (TE II) enzyme will re-generate an active *holo*-NRPS from the misprimed NRPS. Meanwhile the A domains each identify and activate a specific amino acid substrate. The thiol moiety of each *holo*-PCP domain is then loaded with the correct substrate amino acid through the A domain (Figure 1.9, B. loading). Formation of the first peptide bond between the amino acid bound to the PCP domain of the first module and the amino acid of the adjacent module is catalysed by the C domain, starting the direct synthesis along the protein template. This process is termed "initiation" (Figure 1.9, C). The peptide chain is further elongated as it is passed along each of the downstream modules, until the linear product is bound to the PCP domain of the last module (Figure 1.9, D. elongation). The TE domain is generally located at the last module and frequently catalyses cyclisation of the linear peptide product and release from the NRPS (Figure 1.9, E. termination and F. cyclisation). The peptide product can be further modified by glycosylation and lipidation, as well as other modifications, resulting in a diverse range of peptide products (Figure 1.9, G. further modifications).

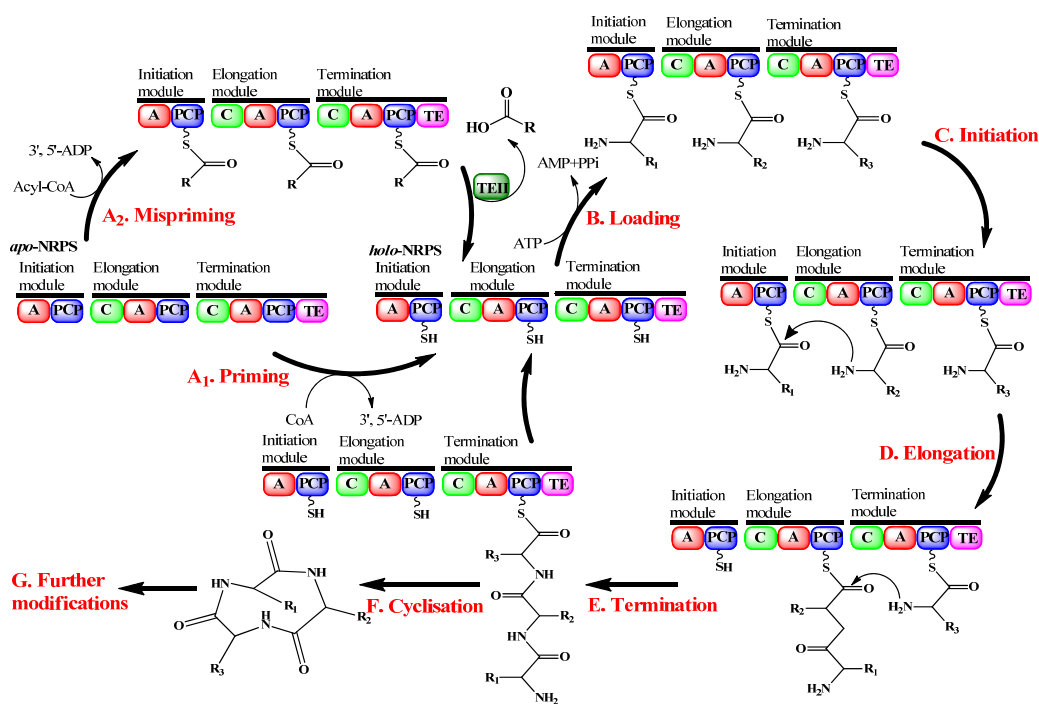


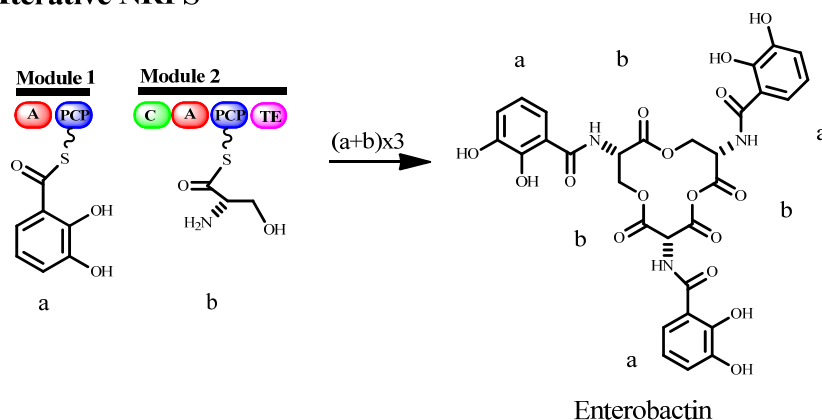
Figure 1.9. The non-ribosomal peptide biosynthetic pathway.

This biosynthetic pathway is illustrated by steps A-G. Details of each process are described in the text. Adapted from Schwarzer *et al.* (2003) [63].

The modular architectures of NRPSs can be classified into three types: linear, iterative and non-linear NRPS [64]. Linear NRPSs consist of a serial organisation of modules, as illustrated in Figure 1.9 with each module responsible for the specific incorporation of one amino acid into the peptide product backbone [62]. Generally each module consists of three domains, *i.e.* the C, A and PCP domains in the order of C-A-PCP [62]. In most cases, the initiation module lacks the C domain, and the termination module has a TE domain to release the peptide product from the assembly line. The antibiotics studied in this thesis (A54145 [2], daptomycin [4], CDA [3], enduracidin [5] and ramoplanin [6]) are examples of linear NRPS products, and their NRPS organisations are given in sections 1.8 and 1.9. Iterative NRPSs produce peptides by using modules or domains more than once, as seen in the biosynthesis of the siderophore enterobactin, which is a cyclic trimer of dihydroxybenzoyl-serine units [64] (Figure 1.10). The third type is the non-linear NRPS, which has an unusual arrangement of the core domains C, A, and PCP, and each peptide product has their own special assembly line. As shown in Figure 1.10, the vibriobactin NRPS contains two modules, and the second module presents a special domain arrangement “C-Cy-Cy-A-C-PCP-C” which differs from the standard arrangement (C-A-PCP) by

an additional two C domains and two Cy domains (Cy domain is responsible for heterocyclic ring formation) [65].

Iterative NRPS



Non-linear NRPS

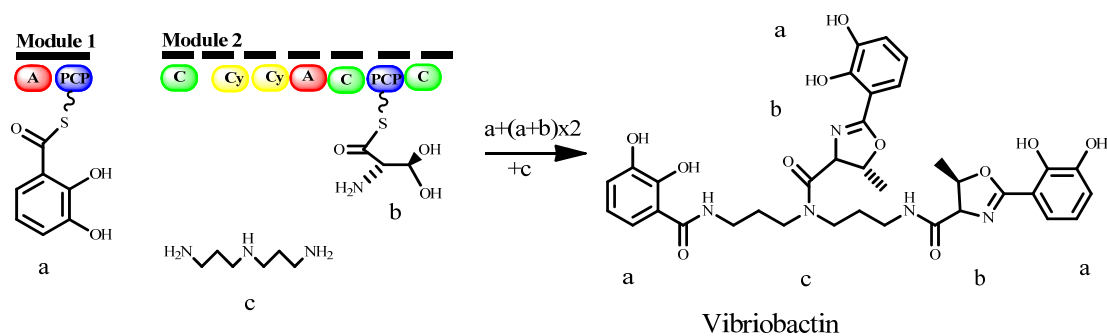


Figure 1.10. Iterative and non-linear NRPS architectures.

The enterobactin NRPS is a representative of an iterative NRPS. This NRPS assembles a dihydroxybenzoyl unit (a) and serine (b) together three times to form a cyclic trimer of dihydroxybenzoyl-serine units, enterobactin. The vibriobactin NRPS is an example of a non-linear NRPS. This NRPS assembles dihydroxybenzoyl units (a), hydroxyl-serine (b) and triamine norspermidine (c) together in an unusual way to form vibriobactin. The Cy domain in module 2 is responsible for heterocyclic ring formation. Adapted from Mootz *et al.* (2002) [64].

1.6.1. The Adenylation (A) Domain

As implied by its name, the A domain of an NRPS carries out the adenylation reaction using ATP. It consists of approximately 550 amino acids and shares a highly conserved structure with the adenylation-forming enzyme, luciferase [66, 67]. This domain can recognise a specific amino acid from a substrate pool and activate this amino acid by reacting it with ATP to form the aminoacyl-adenylate [62] (Figure 1.11). This activation mode is functionally equivalent to the reaction performed by aminoacyl-tRNA synthetase in the ribosomal synthesis system; however, there is no similarity of sequence or structure between them [68].

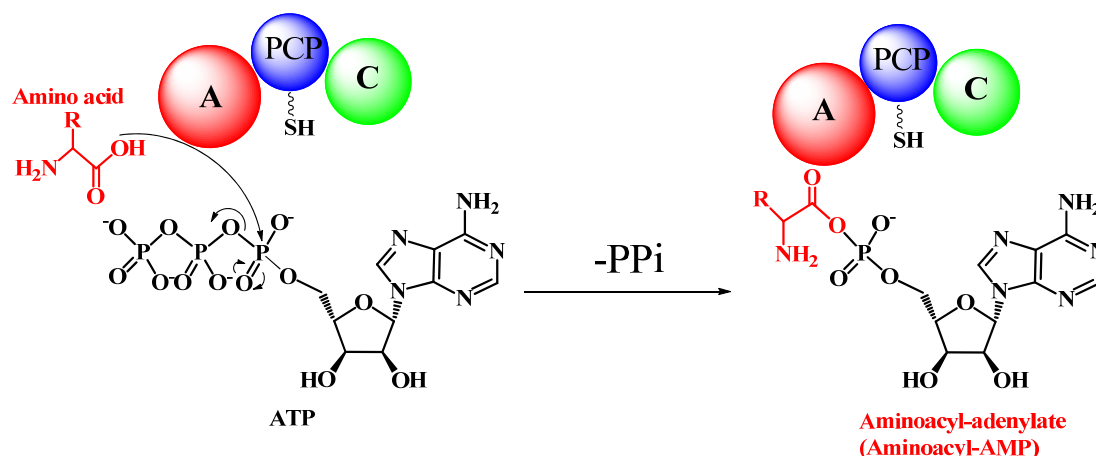


Figure 1.11. Amino acid adenylation by the A domain of NRPS.

The A domain can recognise a specific amino acid and react it with ATP to form the aminoacyl-adenylate. Adapted from Marahiel *et al.* (1997) [62].

Conti *et al.* (1997) solved the three dimensional crystal structure of the A domain of gramicidin *S* synthetase module A (GrsA) in complex with its amino acid substrate Phe and identified ten amino acids (positions 235, 236, 239, 278, 299, 301, 322, 330, 331 and 517) as the substrate binding pocket, which exists in the sequence between the conserved motifs A4 and A5 [69] (see Table 1.3). Subsequently Stachelhaus *et al.* (1999) extracted the sequences between the conserved motifs A4 and A5 from 160 A domains and aligned them with the same region of GrsA [70]. The substrate binding pocket residues were then compared and revealed that the substrate specificity can be determined, that is, A domains activate the same substrate when they have same substrate binding residues. Therefore the ten amino acid residues which form the

substrate binding pocket can be interpreted as the “code” of NRPSs and were named the signature sequence or specificity-conferring codes [69, 70].

By virtue of the information from the signature sequences of known NRPSs, it is possible to perform a reliable prediction of the amino acid substrates for newly identified NRPS sequences [71]. It also provides an alternative way to create novel NRPS products, by changing the selectivity of the A domain [71]. For example, Eppelmann *et al.* (2002) successfully altered the substrate specificity of the A domain of the surfactin synthetase A by site-directed mutagenesis [72]. By changing the single amino acid Lys²³⁹ to Gln²³⁹ in the L-Glu activating initiation module C-A_{Glu}-PCP, the substrate specificity could be altered from L-Glu to L-Gln [72].

1.6.2. The Peptidyl Carrier Protein (PCP) Domain

The PCP domain, which is the only NRPS domain without autonomous catalytic activity, consists of 100 amino acids and is located downstream of the A domain [62, 73]. Its functionality is achieved by post-translational modification with a ppant moiety. This moiety is transferred from CoA to the side-chain hydroxyl group of a conserved Ser residue within the consensus motif “DxFFxLGG(H/D)S(L/I)” of the PCP domain by a ppant transferase, converting the *apo*-PCP (inactive form) to *holo*-PCP (active form) (Figure 1.12) [74-76]. The ppant group of each PCP domain is about 20 Å in length and can serve as a swinging arm to bind the activated amino acid from the A domain, *via* a thio-ester bond, and transport the intermediate to the next catalytic centre, *i.e.* the C domain [75, 76] (Figure 1.13).

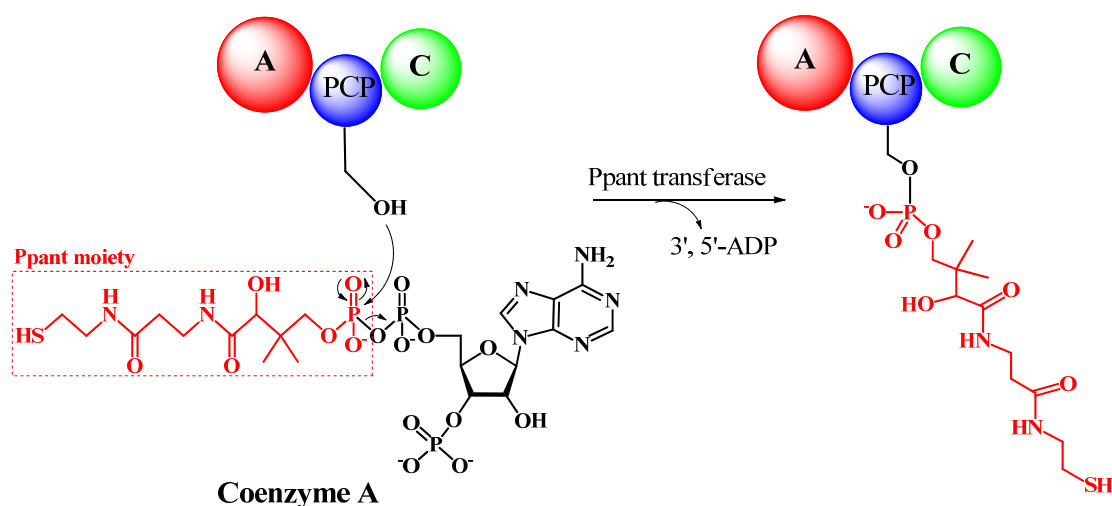


Figure 1.12. Activation of the PCP domain by post-translational modification with the ppant moiety.

The PCP domain is activated as *holo*-PCP by post-translational modification with a ppant moiety, which is transferred from coenzyme A by ppant transferase. Adapted from Marahiel *et al.* (1997) [62].

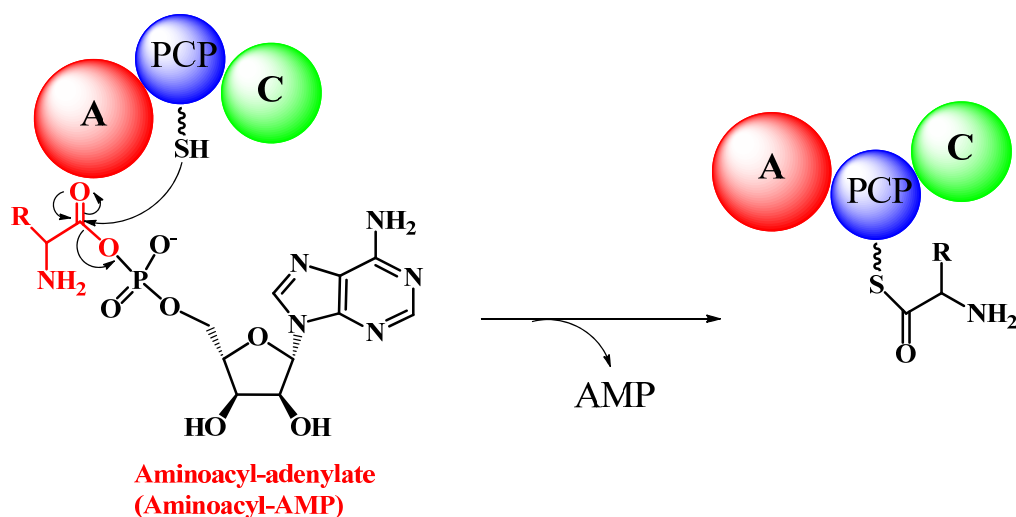


Figure 1.13. Loading of amino acid onto the PCP domain.

The A domain binds an amino acid which is then attached to the PCP domain by the thiol functional group. Adapted from Marahiel *et al.* (1997) [62].

1.6.3. The Condensation (C) Domain

The C domain, approximately 450 amino acids in size, has been confirmed as possessing the catalytic role in peptide bond formation by mutagenesis experiments [73]. Stachelhaus *et al.* (1998) further pointed out that the second His residue in the

conserved C3 motif “MHHxISDG(WV)S” of the C domain (Table 1.3) plays an important catalytic role in peptide bond formation [77]. Donor and acceptor sites exist in C domains that can accept upstream aminoacyl- or peptidyl-S-PCP moieties and downstream aminoacyl-S-PCP moieties respectively [78] (Figure 1.14). When aminoacyl- or peptidyl-S-PCP moieties occupy the donor and acceptor sites, the C domain catalyses peptide bond formation between them. Belshaw *et al.* (1999) showed that the acceptor site has substrate selectivity by studying the first condensation domain of tyrocidine synthetase [78]. This indicates that the C domain also represents a selectivity filter during non-ribosomal peptide synthesis.

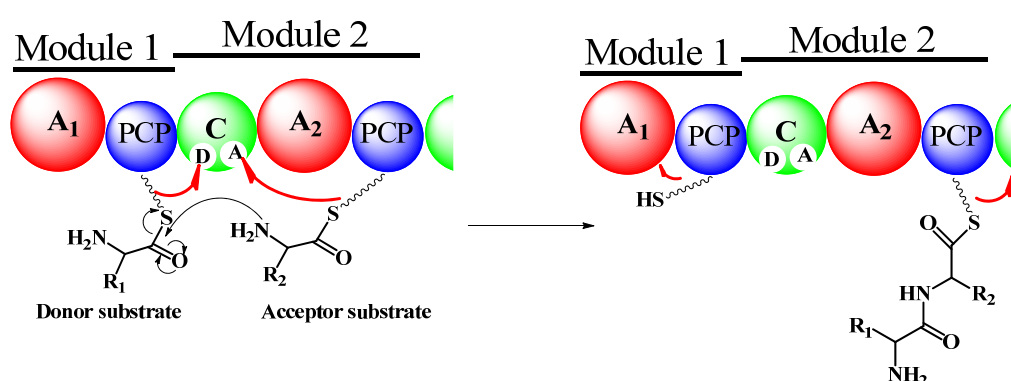


Figure 1.14. Peptide bond formation by the C domain.

A detailed description of peptide bond formation is given in the main text. The red arrows show the swinging directions of the ppant moieties. Adapted from Schwarzer *et al.* (2003) [63].

The ppant moiety of the PCP domain plays a crucial role in transporting amino acids from the A domain to the C domain for peptide bond formation. Schwarzer *et al.* (2002) proposed a model explaining the interactions between the ppant moiety and other domains [63]. In the unbound substrate state, the ppant arm has an enhanced affinity towards the A domain, where it can accept the activated amino acid as an aminoacyl-S-ppant moiety (Figure 1.14). The aminoacyl-S-ppant moieties of PCP domains of module 1 and 2 then have enhanced affinity towards the donor and acceptor sites on the C domain respectively (Figure 1.14). The swinging arms move to the donor and acceptor sites, where the aminoacyl-S-ppant moiety remains until formation of a peptide bond. After peptide bond formation, the affinity of the peptidyl-S-ppant moiety changes towards the donor site of the C domain of the next module.

1.6.4. The Thioesterase (TE) Domain

The TE domain, usually located at the C-terminal end of the last module of the NRPS assembly line, serves as a peptide chain termination module and catalyses the release of the peptide products [60]. This domain is approximately 250 amino acids long and contains a conserved sequence motif “G(HY)SxG” (Table 1.3), which can also be found in other thioesterases and acyltransferases [60]. The Ser residue in this conserved motif can bind the peptidyl chain by nucleophilic attack on the PCP-peptidyl thioester to generate a covalent acyl-enzyme intermediate [79]. The covalently-attached intermediate can be released by hydrolysis as a linear product, *e.g.* the tripeptide precursor of the β -lactam antibiotics, or released by the nucleophilic attack of an internal side chain (amino or hydroxyl group) to produce a cyclic product (lactam or lactone ring), *e.g.* enduracidin [62, 79-81] (Figure 1.15). It has been shown that the TE domain exhibits substrate specificity towards the terminal amino acids of each end of the peptide [82]. For example, in tyrocidine biosynthesis, the TE domain cannot release the peptide product if either of the amino acids at the N- or C-terminus of the peptide are changed [82]. However, there is no effect on the TE domain activity if the internal amino acids of the peptide are altered [82].

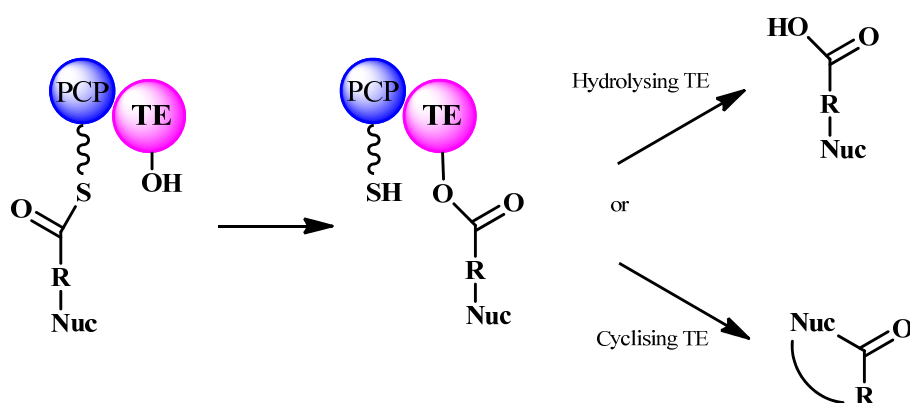


Figure 1.15. The peptide product-releasing mechanisms of the TE domain.

The peptidyl chain is transferred from the phosphopantetheine tether (–SH) of the PCP domain to an active site Ser (–OH) of the TE, resulted in the formation of an acyl-enzyme intermediate. The intermediate may be broken down by hydrolysis to give a linear product or by the attack of an intramolecular nucleophile on the acyl-enzyme ester bond, producing a cyclic product. “Nuc” and “R” stand for nucleophilic group and the peptidyl chain respectively. Adapted from Kohli and Christopher (2002) [79].

Another alternative termination strategy for the NRPS to release peptides can occur at a reductase or condensation domain, which is located at the C-terminus of the NRPS when the TE domain is absent [81]. The reductase domain can catalyse the reduction of the carboxyl group to release an aldehyde product through an NAD(P)H-coupled reaction [81]. The condensation domain uses amide bond formation as a peptide termination strategy by catalysing inter- or intramolecular nucleophilic attack on the thioester bond, resulting in the release of the peptide [81].

1.6.5. The Type II Thioesterase (TE II)

The TE II has a proofreading activity on the PCP domains to ensure NRPS functionality. In the cell, there is a large pool of CoA metabolites, such as free CoA, acyl-CoA and other CoA thioesters, due to its role as an essential cofactor in metabolism [83]. As seen in Figure 1.9, ppant transferase transfers the thiol moiety (ppant) from CoA to the PCP domain, resulting in the formation of an active *holo*-NRPS. However it has been demonstrated that ppant transferase can transfer not only CoASH, but also various acyl-CoA derivatives like acetyl-CoA to target sites [84]. It is likely that PCP domains can be misprimed with acylated ppant cofactors, resulting in an inactive NRPS. Schwarzer *et al.* (2002) characterised two TE II enzymes that are associated with the biosynthesis of the antibiotics surfactin and bacitracin. The results showed that TE II can re-generate misprimed NRPS templates by catalysing the hydrolysis of the inhibitory acyl groups, supporting a role of TE II in NRPS proofreading [85].

1.6.6. Other Domains in NRPSs

In addition to the essential domains, *i.e.* the A, PCP and C domains, some NRPSs contain other modification domains such as epimerization (E) and methylation (MT) domains. These modification domains can contribute to the production of D-amino acids from L-amino acids, or cause methylation on amino acids, resulting in a vast structural diversity of non-ribosomal peptide products [61, 62, 86, 87].

The conversion of L-amino acids to D-amino acids occurs in two ways. The first is *via* enzyme catalysis by the E domain which is approximately 450 amino acids in length and located between the PCP domain and the C domain of the next module [88]. In gramicidin *S* synthetase, a E domain has been found in the first module which catalyses the de- and re-protonation of the α -carbon of the aminoacyl-*S*-PCP moiety, changing the configuration of the amino acid to produce a D-aminoacyl-*S*-PCP intermediate [88]. This catalysis occurs after adenylation, but before condensation, at the thioester-bound stage [88, 89]. Only the correct enantiomer can be incorporated into the growing peptide chain, due to the enantioselectivity of the donor site of the downstream C domain [78, 90]. Alternatively, the conversion can be catalysed by an external racemase. For example, a racemase which is involved in the biosynthesis of cyclosporin in the fungus *Tolypocladium niveum*, has been purified and confirmed to catalyse the transformation of L-Ala into D-Ala [91]. Then D-alanine can then be activated by a D-Ala-specific A domain and incorporated into the cyclosporine peptide chain [91].

The MT domain, which usually consists of 420 amino acids, is located between the A and PCP domains. The sequence of MT domain has high homology to other *S*-adenosylmethionine-dependent methyl transferases (SAM-dependent MTases). They transfer methyl group from SAM (Figure 1.16) to substrate amino acids before peptide bond formation [62, 63, 80]. For example, seven of the eleven amino acids of cyclosporin are methylated by MT domains whilst they are bound to PCP domains as aminoacyl-*S*-PCP intermediates [92]. An MT domain can also be found in the A54145 antibiotic synthetase, and it is postulated that this domain can catalyse the methylation of glycine to form the non-proteinogenic amino acid sarcosine (Sar) [2]. In addition to the MT domain, an external MTase can also be responsible for this modification; more detail on external MTases will be given in section 1.7.2.

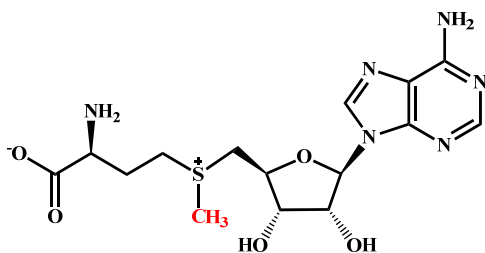


Figure 1.16. The structure of *S*-adenosylmethionine (SAM).

1.7. Tailoring Enzymes

There are a number of tailoring enzymes whose genes are usually located nearby or within the NRPS gene cluster. They can carry out modifications, such as methylation, hydroxylation, halogenation and phosphorylation as well as glycosylation and lipidation. These modifications determine many key peptide product properties, such as their activity, solubility and structural stability. Examples of how modifications affect peptide product properties are given in chapters 2 and 3. These modifications can occur before, during, or after peptide assembly [61, 86]. As a result of such modification mechanisms, NRPSs can produce many diverse peptide product structures with varied properties. Therefore it is important to explore and understand the function of tailoring enzymes in the production of potential peptide products with desirable properties, such as good activity and solubility. The following sections will introduce four types of modifications that we are investigating, *i.e.* hydroxylation, methylation, halogenation and mannosylation.

1.7.1. Hydroxylation

Hydroxylation at the β -carbon of an amino acid is a common modification found in many peptide antibiotics. Ten different proteinogenic amino acids have been found to be modified by β -hydroxylation in antibiotics [93] (Table 1.4). This is an important modification due to the fact that the hydroxyl group can be a site for further modification, such as glycosylation [94], methylation [2], phosphorylation [95], oxidation [96] and macrolactonization [6], further diversifying the potential peptide product structures. To date, two types of enzyme which introduce oxygen onto the un-activated carbon centre of amino acid methyl groups have been discovered: **non-heme, Fe(II)/ α -ketoglutarate (α -KG)-dependant hydroxylases and cytochrome P450 monooxygenases.**

Table 1.4. β -hydroxyl amino acids in peptide antibiotics

β-OH amino acid	Antibiotic
β -OH-Tyr	chloroeremomycin [97], vancomycin [98], teicoplanin [94], novobiocin [99], coumermycin A ₁ [100]
β -OH-Phe	katanosin B [101]
β -OH-Gln	echinocandins [102]
β -OH-Asn	katanosin B [101], A54145 [2], CDA [103], ramoplanin [6]
β -OH-Asp	plusbacin A ₃ [104], syringomycin [105]
β -OH-His	bleomycin [106], nikkomycin [96]
β -OH-Pro	plusbacin A ₃ [104], echinocandins [102]
β -OH-Leu	katanosin B [101]
β -OH-Val	aureobasidin [107]
β -OH-Arg	streptothricin [108]

Adapted from Chen *et al.* (2001) [93].

The non-heme Fe(II)/ α -KG-dependant hydroxylases are a superfamily of enzymes that catalyse a wide range of oxidative reactions including hydroxylations, epoxidations, desaturations and other types of oxidative modifications [109]. These reactions cover a wide range of substrates and play important roles in protein modification, repair of nucleotides, biosynthesis of antibiotics, lipid metabolism and other pathways [109, 110]. Most of these enzymes need α -KG as a co-substrate to complete their reactions. For example the Asn oxygenase (AsnO) from the CDA biosynthetic gene cluster belongs to this type of hydroxylase superfamily and can hydroxylate L-Asn in the presence of iron and α -KG [111]. In the catalytic centre, iron can be bound by a conserved motif, the “2-His-1-carboxylate facial triad”, which consists of two histidyl groups and one glutaryl or asparagyl residue and exists in all of this superfamily of enzymes [109, 110] (Figure 1.17). Meanwhile α -KG can also chelate iron *via* its C-1 carboxylate and C-2 ketone groups, and maintains additional binding stabilization through interaction between the C-5 carboxylate group and the neighboring side chains of the enzyme [109] (Figure 1.17). The iron centre can then bind O₂ and change its oxidation state, promoting the introduction of a hydroxyl group into the substrates [109].

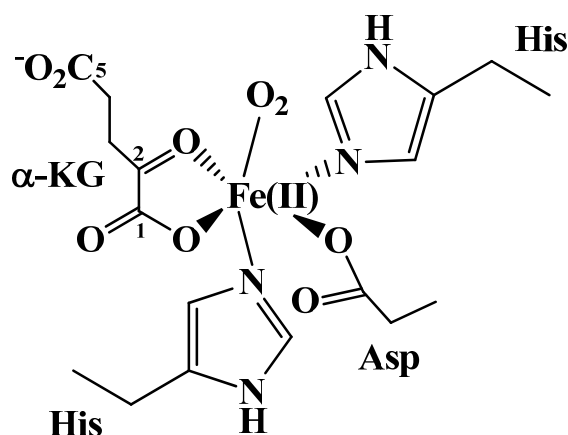


Figure 1.17. Reactive intermediate in the mechanism of the Fe(II)/ α -KG-dependant hydroxylases.

Cytochrome P450 monooxygenases belong to a superfamily of heme-containing enzymes. Their functionality is similar to that of Fe(II)/ α -KG-dependant hydroxylases and they can also catalyse a wide range of oxidative reactions such as hydroxylations, epoxidations and ring formation using oxygen and a corresponding electron transfer system, but they require the cofactor NAD(P)H [112]. In peptide antibiotic biosynthesis, these monooxygenases have been found to be involved in the hydroxylation of Tyr during the biosynthesis of novobiocin [99], and in ring formation during the biosynthesis of vancomycin-type antibiotics [113].

1.7.2. Methylation

Methylations are carried out by a large group of MTases and are involved in many critical physiological functions, such as gene regulation, protein repair and signal transduction. SAM-dependent MTases are the major group amongst the MTases [114]. They transfer a methyl group from the substrate SAM to a wide range of target molecules such as nucleotides (DNA and RNA), lipids, proteins and small compounds, by targeting carbon, nitrogen, oxygen, sulphur or even halide atoms in the substrate [114, 115]. So far, around 140 members of SAM-MTase family have been identified in the UniProtKB/Swiss-Prot database [116, 117].

In the biosynthesis of non-ribosomal peptide antibiotics, methylations can occur on oxygen, nitrogen and carbon atoms, and are catalysed by discrete O-, N-, and

C-MTases, or by MTase domains embedded within an NRPS module [86]. Examples include 3-methyl Glu (3mGlu) and β -methyl Phe found in the antibiotics CDA and mannopeptimycin, and whose methyl groups are transferred from SAM by the C-MTases GlnT and MppJ respectively [118, 119].

1.7.3. Halogenation

Thousands of halogenated natural products exist in nature, with structural classes ranging from simple phenolic and aliphatic compounds to complex peptides with a wide variety of biological activities, for example antibiotic and antitumor activities [120]. Halogenating enzymes have been found in a wide range of organisms and can be grouped into two major types: (1) the less substrate-specific haloperoxidases which utilise hydrogen peroxide and (2) the highly substrate-specific halogenases which require dioxygen for enzyme activity [120] (Table 1.5).

Table 1.5. The four classes of halogenating enzymes

Type of halogenase	Redox cofactors	Oxidants	Substrates	Proposed forms of activated halogen
Haloperoxidases	Heme Fe(II)	H ₂ O ₂	Aromatic and electron rich substrates	X ⁺
	Vanadium			
O₂-dependent Halogenases	FADH ₂	O ₂	Aromatic and electron rich substrates	X ⁺
	Non-heme Fe(II)	O ₂	α -KG, aliphatic unactivated carbon	X [•]

1.7.3.1. Haloperoxidases

Haloperoxidases can be classified into heme-dependent and vanadium-dependent halogenases according to the presence of the respective cofactors heme or vanadium in their catalytic centres [121]. They can utilize hydrogen peroxide to convert a halide

ion (X^-) to a hypohalite ($-OX$) intermediate, which can react as “ X^+ ” equivalents with electron-rich substrates [121] (Figure 1.18). However, unlike heme-dependent halogenases, the vanadium-dependent halogenases do not change the oxidation state of their metal centre during the catalytic process [121, 122] (Figure 1.18).

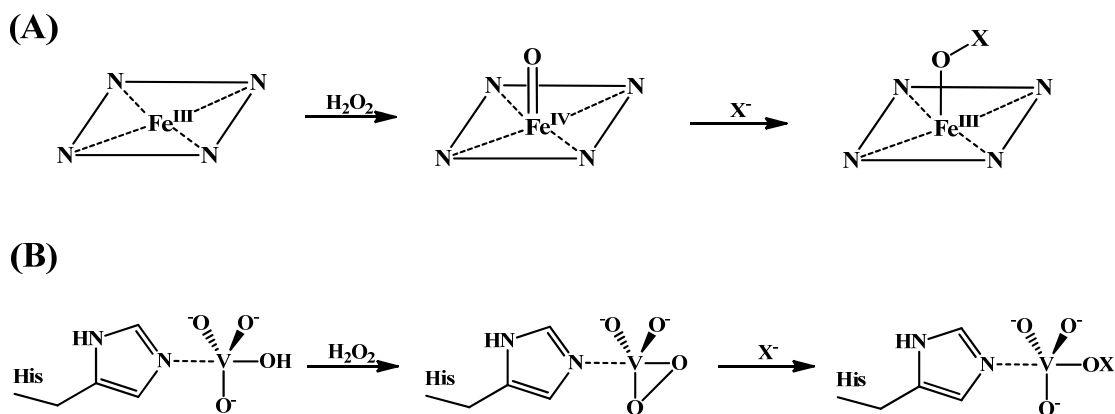


Figure 1.18. Reactive intermediates generated by haloperoxidases.

The intermediates in the mechanisms of heme- (A) and vanadium- (B) dependent haloperoxidases. Adapted from Vaillancourt *et al.* (2006) [121].

1.7.3.2. O₂-dependent halogenases

O₂-dependent halogenases can be classified into FADH₂-dependent and non-heme Fe(II)/ α -KG-dependant halogenases, based on the utilisation of the cofactors FADH₂ and Fe(II) [123]. FADH₂-dependent halogenase genes are found in gene clusters coding for the biosynthesis of natural products containing chlorinated electron-rich aromatic side chains, *e.g.* Trp, pyrrole and/or phenol derivatives [123]. They contain two conserved motifs, one is the GGGxxG motif which is located at the N-terminal for FAD-binding, and the other is the GWTWxIP motif which can be found in the middle of halogenase sequence [124] (Figure 1.19). The Trp residues in the second motif are assumed to prevent the halide substrate from binding to the flavin cofactor, so that the enzymes cannot act as a flavin-dependent monooxygenases [125]. The *in vitro* activity of this type of halogenase has been set up using the tryptophan halogenase PrnA from the pyrrolnitrin biosynthetic gene cluster [126]. It is a two-component enzyme that consists of a flavin reduction system and a halogenation system. Flavin reduction is carried out by flavin reductase that reduces FAD to FADH₂, which is then oxidized by the halogenase with O₂ to give FAD-OOH. It has

been proposed that Cl^- reacts with FAD-OOH to generate HOCl which can then act as the chlorinating agent, providing an electrophilic Cl^+ to directly chlorinate the electron-rich aromatic side chains (Figure 1.20A) [125, 127].



Figure 1.19. The general amino acid sequence assembly of FADH₂-dependent halogenases.

“x” represents a variable amino acid residue.

Non-heme Fe(II)/ α -KG-dependant halogenases are homologous to the non-heme Fe(II)/ α -KG-dependant hydroxylases. They also use an iron catalytic centre, which can be bound by two His residues and cosubstrate α -KG, for the chlorination of methyl groups on substrates [121, 123] (Figure 1.20B). Initial insight into their catalytic mechanism was derived from investigations on SyrB2, which chlorinates the γ -methyl group of L-Thr in the lipodepsipeptide syringomycin E, produced by *Pseudomonas syringae* [128]. Chlorination of L-Thr occurs only when tethered to the PCP domain of syringomycin E synthetase [128]. In the SyrB2 crystal structure, the conserved 2-His-1-carboxylate facial triad motif, which usually exists in non-heme Fe(II)/ α -KG-dependant hydroxylases, can also be identified in this enzyme’s catalytic centre. However only two His residues (His-116 and His-235) are observed to bind the iron, and the Asp or Glu carboxylate residue has been replaced with an Ala¹¹⁸ residue [129]. It is thought that the replacement with Ala provides enough space for a chloride ion to enter the catalytic centre [129]. The catalytic mechanism is very similar to that of Fe(II)/ α -KG-dependant hydroxylases, and uses the iron centre to generate reactive species such as Cl^\bullet and R-CH_2^\bullet , resulting in the introduction of a chloride ion into the substrate [121] (Figure 1.20B).

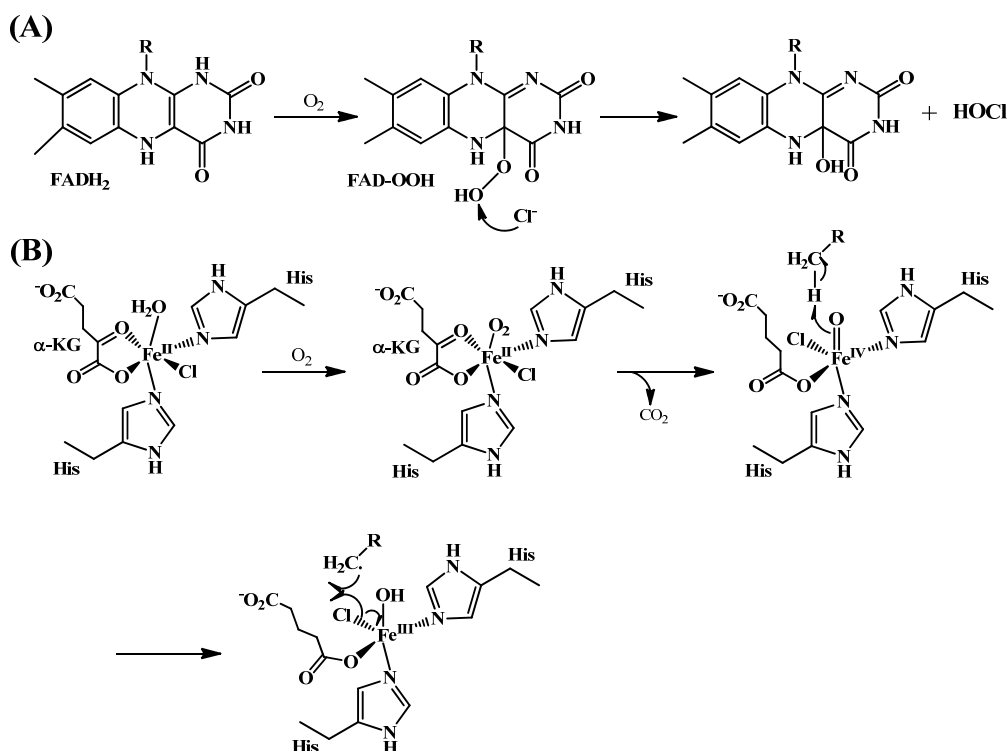


Figure 1.20. Intermediates in the mechanisms of O₂-dependent halogenases. FADH₂-dependent halogenases (A) and non-heme Fe(II)/ α -KG-dependant halogenases (B). Adapted from Vaillancourt *et al.* (2006) [121].

1.7.4. Mannosylation

Glycosylation is a common modification on proteins, nucleotides, secondary metabolites and other small molecules, and is widespread amongst prokaryotes and eukaryotes. It plays multiple important biological roles, such as in cell wall synthesis by bacteria, fungi and yeast, and in the modification of proteins for cellular localisation and ligand interaction [130]. Such glycosylations are primarily carried out by glycosyl transferases (Gtfs) and involve the transfer of a sugar moiety from a nucleotide sugar or a lipid phosphate sugar (for example a polyprenyl phosphoryl sugar) to the target molecule. We are interested in the mannosylation of non-ribosomal peptide antibiotics. Although mannosylation has been studied extensively in cell wall synthesis for bacteria, fungi and yeast, as well as in protein modification; the mannosylation of antibiotics such as teicoplanin, mannopeptimycin and ramoplanin has not yet been characterised. In this study we describe our investigation into the mannosylation of ramoplanin. The details of this study and the importance of glycosylation will be outlined in Chapter 3.

1.8. The Antibiotics A54145, CDA and Daptomycin

The antibiotics A54145, CDA and daptomycin, produced from *S. fradiae*, *S. coelicolor* and *S. roseosporus* respectively, are Ca^{2+} -dependent cyclic lipodepsipeptides which share many structural features [2-4]. All are ten-residue lactone rings, which are linked by an ester bond from Ile/Val-13 (A54145), Trp-11 (CDA) or Kynurenine-13 (Kyn-13) (Daptomycin) to Thr-2 or 4, and with an exocyclic amino acid tail containing one or three amino acids (Figure 1.21). The N-termini of the exocyclic amino acids are generally coupled to fatty acids that are important for their antibacterial activity [131, 132]. All three antibiotics contain three conserved residues, Asp, Asp/3-*O*-methyl Asp (OmAsp) and Glu/3mGlu, which are important for antimicrobial activity. For example, Kopp *et al.* (2006) found that the antibiotic A54145 had reduced antibiotic activity as a result of the single deletion of Asp-7 or Asp-9, or from the deletion of both amino acids [133]. In a CDA study, the replacement of Asp-7 with Asn caused loss of CDA activity [134]. Similarly, Nguyen *et al.* (2006) showed that the activity of daptomycin which contains 3mGlu is 4-fold higher than that of Glu-containing analogues [135].

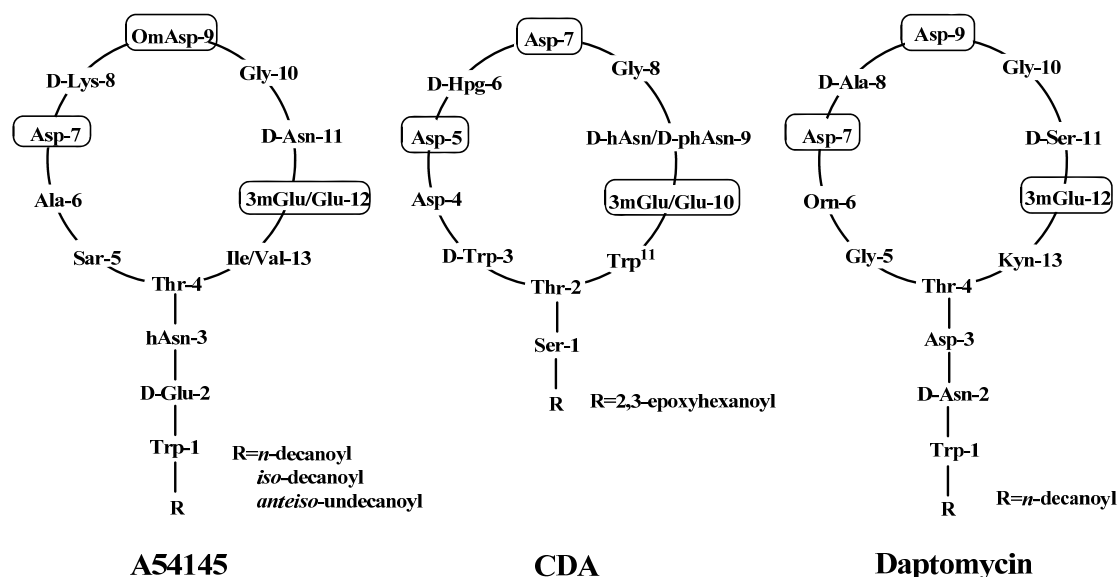


Figure 1.21. Structures of the lipopeptide antibiotics, A54145, CDA and Daptomycin.

Abbreviations for unusual amino acids: hAsn for hydroxyl Asn, OmAsp for 3-*O*-methyl Asp, 3mGlu for 3-methyl Glu, phAsn for phospho-hydroxyl Asn, Hpg for hydroxyl-phenylglycine, Orn for ornithine, Kyn for kynurenine, Sar for sarcosine.

Silverman *et al.* (2003) proposed a multistep model for the mode of action of daptomycin by investigating the depolarization of the bacterial cell membrane of *Staphylococcus aureus* [42]. As shown in Figure 1.22, the long fatty acid chain of daptomycin is proposed to insert into the cytoplasmic membrane. With the binding of calcium ions, daptomycin monomers aggregate together and penetrate deeper into membrane, resulting in the creation of a membrane channel. These channels trigger potassium ion release, leading to membrane depolarization and the disruption of cellular activity.

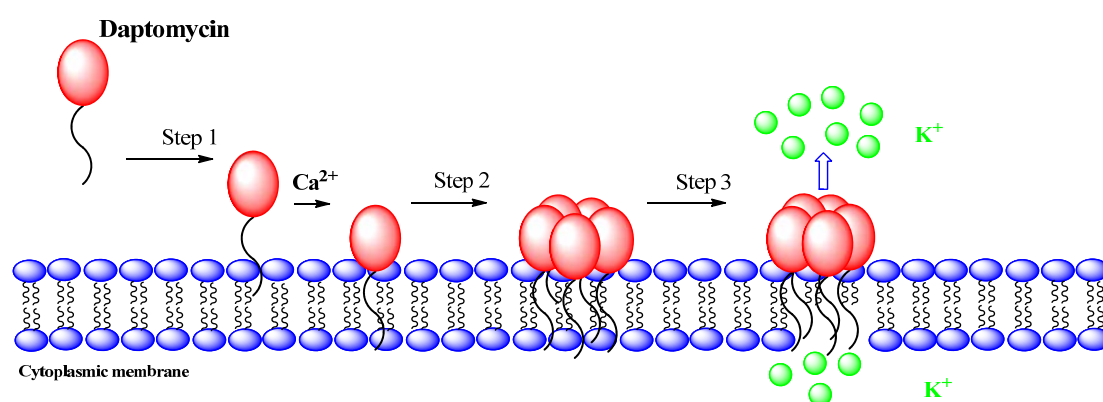


Figure 1.22. Proposed mode of action for daptomycin.

Step 1: Daptomycin inserts into the cytoplasmic membrane. Step 2: Oligomerisation of daptomycin is triggered by the binding of calcium ions. Step 3: Intracellular ions are released from the channels, leading to disruption of the cell membrane potential. Adapted from Silverman *et al.* (2003) [42].

1.8.1. The Antibiotic CDA

In 1978 an antibiotic produced by *S. coelicolor* was found and shown to inhibit the growth of Gram-positive bacteria in the presence of calcium. Accordingly it was given the name of **C**alcium-**D**ependent **A**ntibiotic (CDA) [136]. It was first isolated by Hopwood and Wright in 1983 and confirmed to be a chromosomally-determined antibiotic [137]. The chemical structure of CDA was solved by Kempter *et al.*, in 1997 and comprises of 11 amino acids including three non-proteinogenic amino acids: D-hydroxyl phenylglycine (D-Hpg), D-hydroxyl Asn (D-hAsn)/phospho-hydroxyl Asn (phAsn) and 3mGlu [103] (Figure 1.23). Six CDA variants can be produced by *S. coelicolor* depending on the type of nitrogen source present in the culture medium

[103]. For example, fermentation with yeast extract yielded CDA1b and CDA2b (with relative monoisotopic molecular masses (M_r) of 1562.5 and 1576.5 respectively), whilst fermentation using soy peptone and meat extract gave CDA3a, CDA3b, CDA4a, and CDA4b ($M_r = 1480.5, 1482.5, 1494.5$ and 1496.5 , respectively). In 2002, the sequencing of the *S. coelicolor* genome was completed and it was shown that the CDA gene cluster comprises 82 kb of the genome (approximately 1 % of the total *S. coelicolor* genome) and encodes over 40 ORFs [3, 138]. In this biosynthetic gene cluster, three genes that encode three large NRPS enzymes, designated CDA1, CDA2 and CDA3 containing six, three and two modules respectively, are responsible for the assembly of the CDA peptide (Figure 1.24). Additionally, other genes in the CDA biosynthetic gene cluster encode proteins or enzymes responsible for the regulation of CDA production and the transport, resistance to and modification of CDA [3].

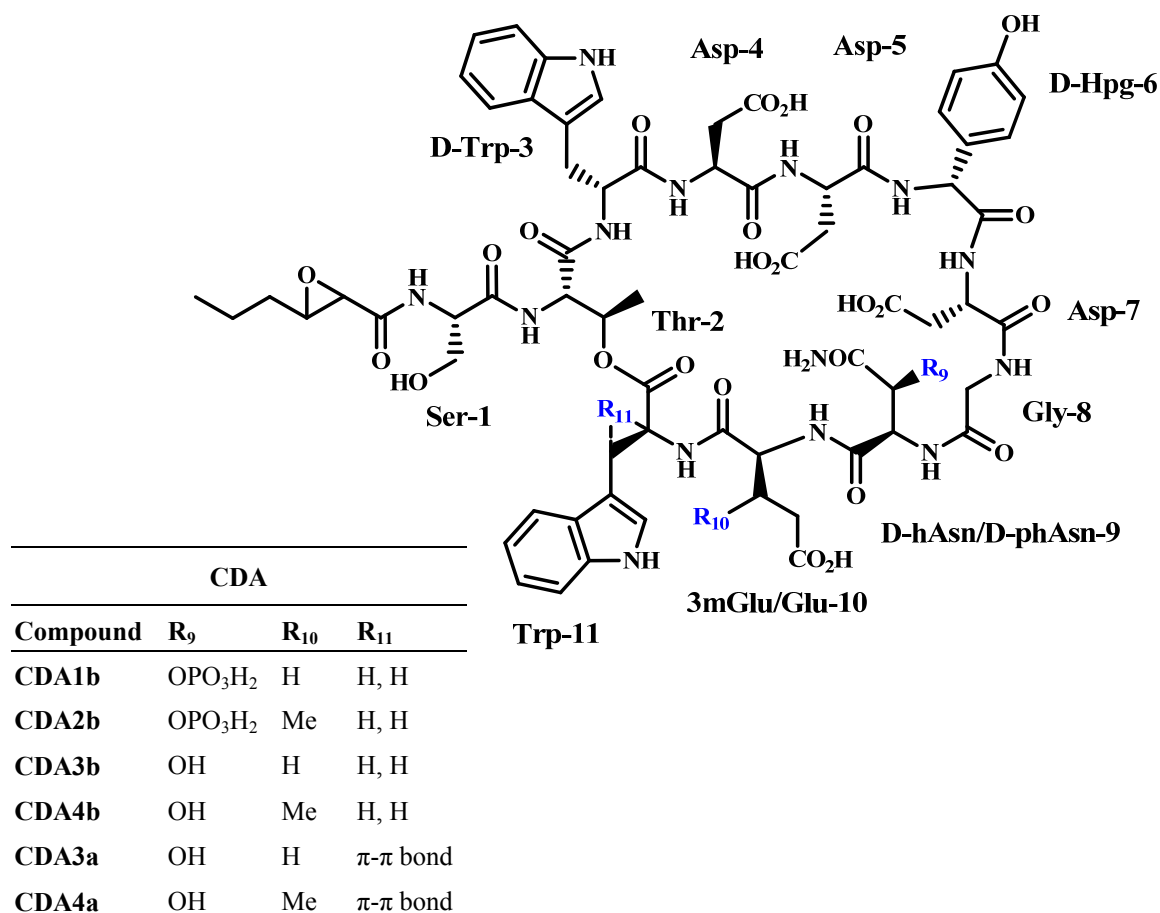


Figure 1.23. Structures of the lipopeptide antibiotic CDA.

The CDA structure with three of the variable amino acid residues highlighted in blue: R₉, R₁₀ and R₁₁. The table indicates which groups are present at the three positions in each of the six CDA variants.

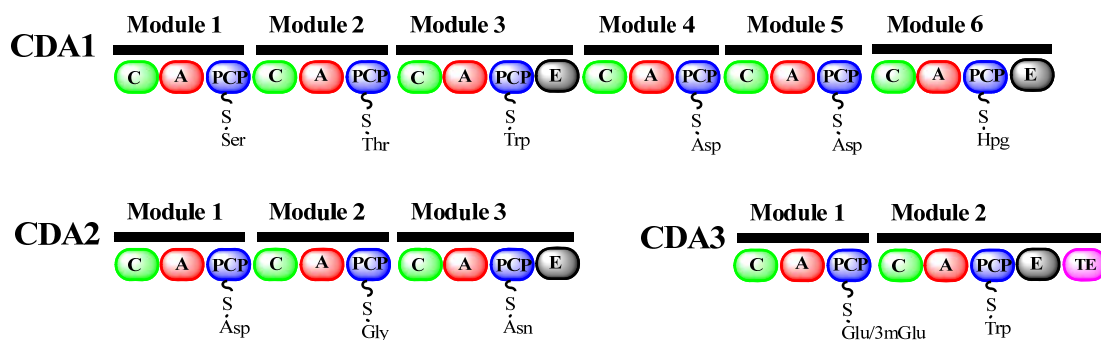


Figure 1.24. Organisation of the CDA NRPS.

1.8.2. The Antibiotic Daptomycin

In 2003, daptomycin (Cubicin; Cubist Pharmaceuticals) was approved by the U.S. FDA for the treatment of infections caused by Gram-positive bacteria [7]. It was the first new class of naturally-derived antibiotic to reach the clinic for over three decades. It has shown potent activity against many important Gram-positive pathogens, including methicillin-resistant *S. aureus* (MRSA), vancomycin-resistant *S. aureus* (VRSA) and vancomycin-resistant *Enterococcus faecalis* (VREF) [139].

A complex mixture of antibiotics (A21978C) with activity against Gram-positive bacteria was found from fermentation of *S. roseosporus* by Eli Lilly Company [139, 140]. It is a mixture of lipopeptides with the same cyclic peptide core but with different fatty acid side-chain attachments [141]. Subsequently, it was found that a compound which had a decanoic acid side-chain possessed a strong antibiotic activity, and was subsequently named daptomycin [141] (Figure 1.25). Daptomycin consists of 13 amino acids including three non-proteinogenic amino acids (Kyn, ornithine (Orn) and 3mGlu), and can be produced from the fermentation of *S. roseosporus* with feeding of *n*-decanoic acid [141].

The daptomycin biosynthetic gene cluster (*dpt*) covers 128 kb and encodes over 60 ORFs. Miao *et al.* (2005) and Penn *et al.* (2006) cloned the *dpt* gene cluster into the heterologous host *S. lividans* and successfully produced daptomycin, validating that the 128 kb gene cluster encodes for daptomycin production [4, 142]. Three daptomycin synthetases are encoded in the *dpt* gene cluster, DptA, DptB and DptC,

which are responsible for the assembly of the 13-residue daptomycin peptide backbone (Figure 1.26) [4]. Other genes encoded in the *dpt* gene cluster are responsible for the regulation of daptomycin production and the transport, resistance to and modification of daptomycin [4].

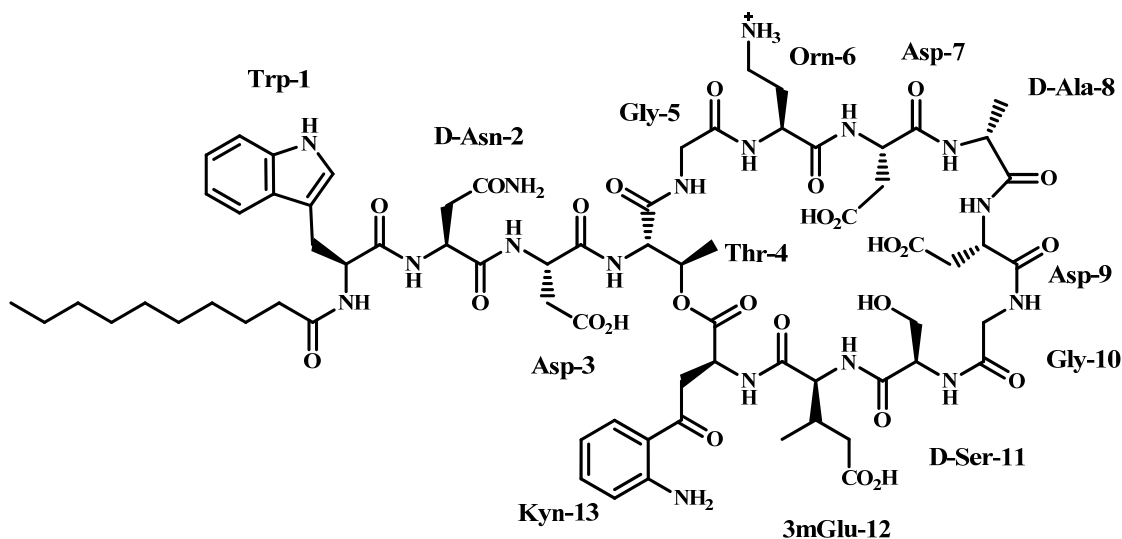


Figure 1.25. Structure of the lipopeptide antibiotic daptomycin.

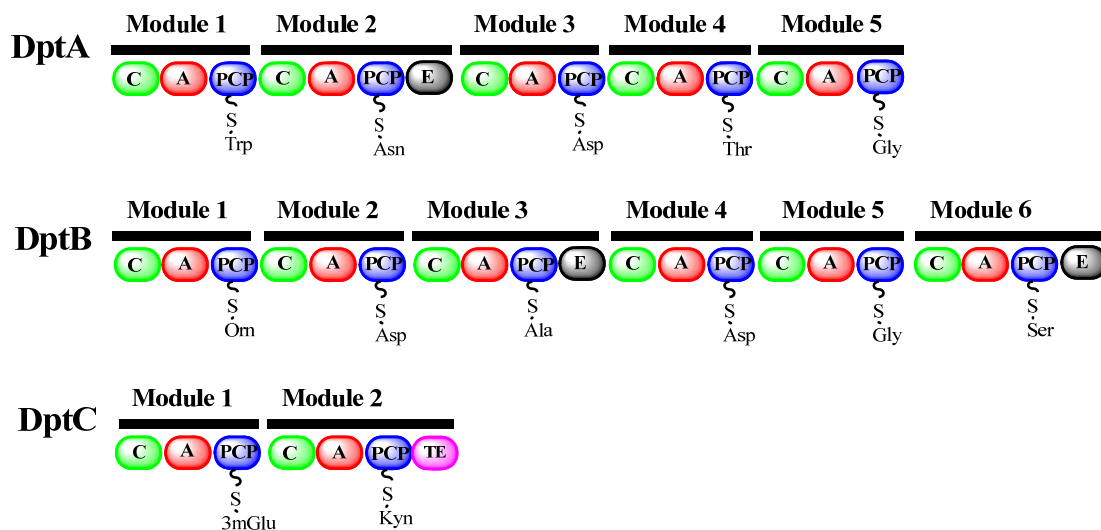


Figure 1.26. Organisation of the daptomycin NRPS.

1.8.3. The Antibiotic A54145

In 1990, eight forms of the lipopeptide A54145 (Figure 1.27) were discovered, isolated and characterised from the bacterium *S. fradiae* NRRL18159 by Boeck and co-workers, and which were found to be active against Gram-positive bacteria [131, 143, 144]. In 2006, the A54145 biosynthetic gene cluster was determined and shown to encompass a 127 kb region, with 61 ORFs [2]. Through sequence analysis, 4 ORFs were found to encode LptA, LptB, LptC and LptD synthetases for assembly of the 13-residue A54145 peptide backbone (Figure 1.28). Other accessory genes encode proteins that are proposed to be involved in peptide tailoring, export, regulation and resistance. This antibiotic contains four non-proteinogenic amino acids: hAsn, Sar, OmAsp and 3mGlu. In this study we were interested in the biosynthesis of the non-standard amino acid OmAsp, details of which are presented in Chapter 2.

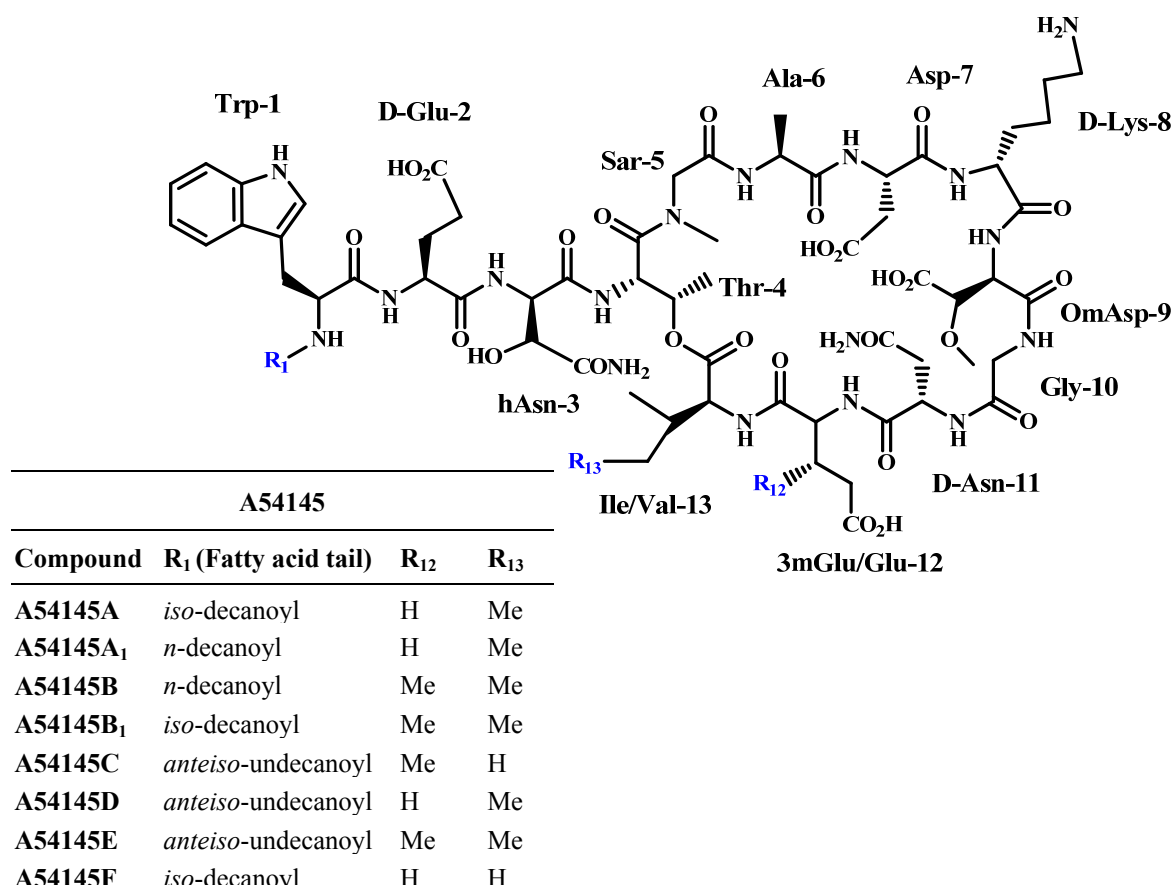


Figure 1.27. Structures of the lipopeptide antibiotic A54145.

The A54145 structure with variable fatty acid tails and two of the variable amino acid residues highlighted in blue: R₁, R₁₂ and R₁₃. The table indicates which fatty acid chain and chemical groups are present in each of the eight A54145 variants.

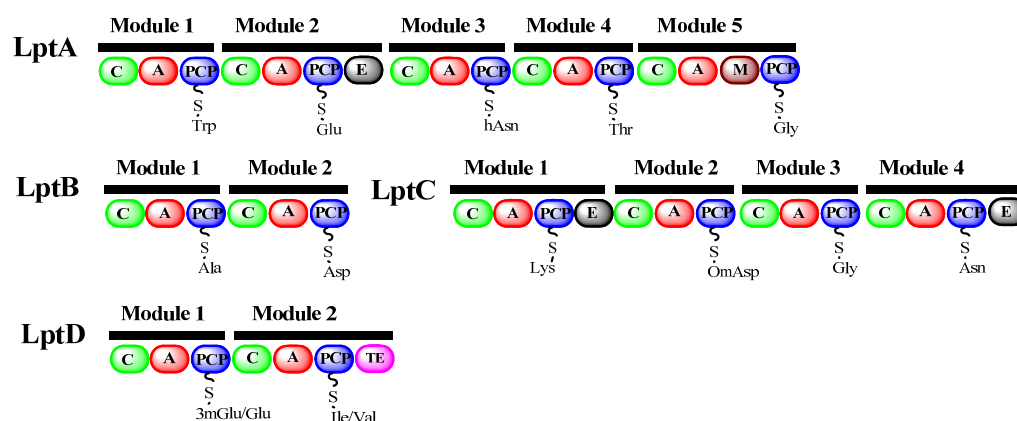


Figure 1.28. Organisation of the A54145 NRPS.

1.9. The Antibiotics Enduracidin and Ramoplanin

Enduracidin and ramoplanin are seventeen amino acid lipodepsipeptides produced by the soil bacteria *S. fungicidicus* ATCC21013 (Strain No.B5477) [145, 146] and *Actinoplanes* strain ATCC33076 [147] respectively. Both antibiotics share similar primary and secondary structures (Figure 1.29). For example, the amino acid stretches Hpg-3 to Thr-8 and Gly-14 and Ala-16 of both antibiotics are identical, and the NMR structures of both antibiotics solved by Castiglione *et al.* (2005) showed that they have similar β -hairpin arrangements [6, 148]. However there are significant different arrangements of the side-chains between enduracidin and ramoplanin, *e.g.* six out of seventeen residues are different amino acids (Table 1.6). Additionally, three Hpg residues (Hpg-11, Hpg-13 and Hpg-17) in both antibiotics are also different from one another: Hpg-11 is unmodified in enduracidin but is dimannosylated in ramoplanin; Hpg-13 is dichlorinated to form 3, 5-dichloro-Hpg (Dpg) in enduracidin but is unmodified in ramoplanin; and Hpg-17 is unmodified in enduracidin but is monochlorinated to form 3-chloro-Hpg (Chp) in ramoplanin (Table 1.6). In ramoplanin, the aromatic residues 3, 9, and 17 can form a hydrophobic core which presumably stabilises the ramoplanin structure [148]. In contrast, the residue Citrulline-9 (Cit-9) in enduracidin substitutes the hydrophobic side chain of the Phe in ramoplanin, resulting in the loss of the hydrophobic core [148].

Table 1.6. Amino acid alignment between enduracidin and ramoplanin

Residues	1	2	3	4	5	6	7	8	9
Enduracidin	Asp	Thr	D-Hpg	D-Orn	D-aThr	Hpg	D-Hpg	aThr	Cit
Ramoplanin	Asn	hAsn	D-Hpg	D-Orn	D-aThr	Hpg	D-Hpg	aThr	Phe

Residues	10	11	12	13	14	15	16	17
Enduracidin	D-End	Hpg	D-Ser	Dpg	Gly	End	D-Ala	Hpg
Ramoplanin	D-Orn	Hpg	D-aThr	Hpg	Gly	Leu	D-Ala	Chp

Note: The six different amino acids between enduracidin and ramoplanin are labelled in red. Hpg-11, Hpg-13 and Hpg-17 residues which have different modifications are labelled in blue. The dimannosyl group is present on residue Hpg-11 of ramoplanin. The residues End-10 and End-15 in enduracidin represent the non-proteinogenic amino acid enduracididine.

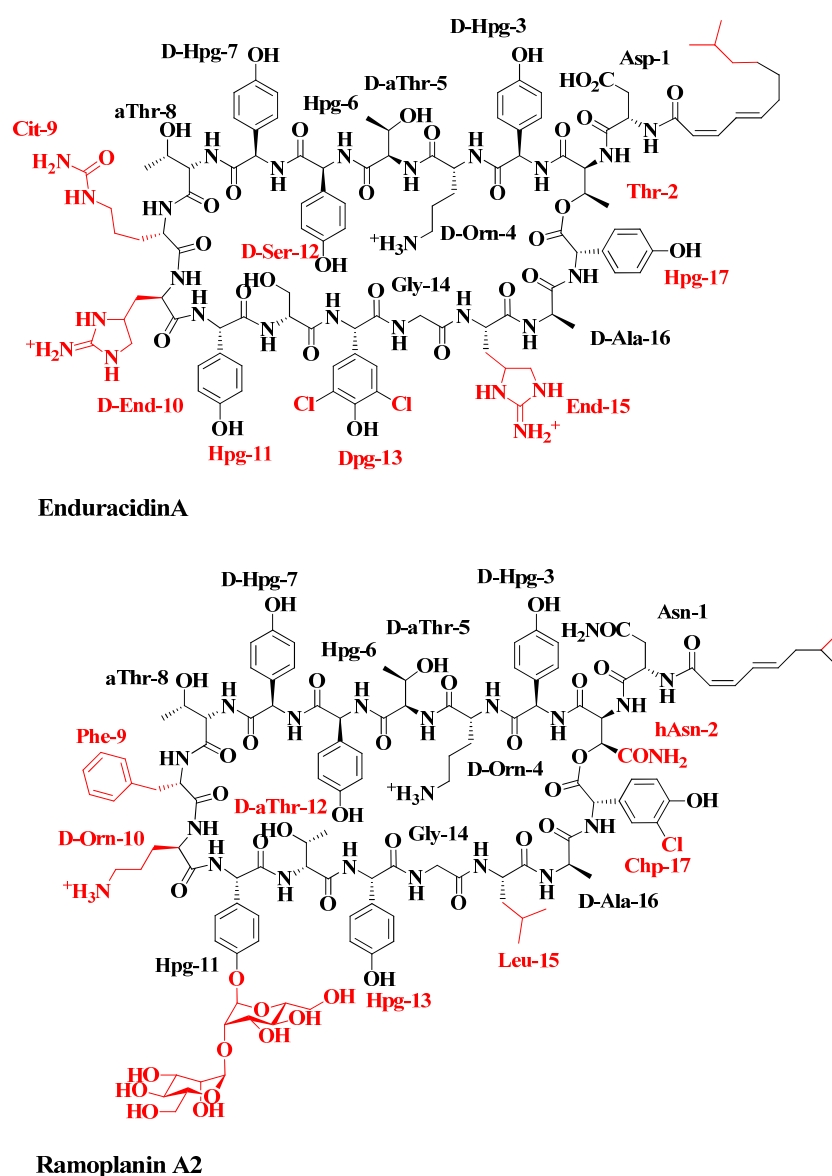


Figure 1.29. Structures of the lipoglycopeptide antibiotic ramoplanin A2 and enduracidin A.

The differences between ramoplanin and enduracidin are labelled in red.

Both antibiotics have excellent activity against life-threatening pathogens such as *S. aureus* and Vancomycin-resistant *Enterococci* (VRE) in comparison to other antibiotics such as vancomycin [149-152] (Table 1.7). The mode of action of both antibiotics is proposed to be binding of the transglycosylase substrate lipid II, leading to interference with the transglycosylation step of peptidoglycan formation and inhibition of bacterial cell wall biosynthesis [44]. They are promising antibiotics as the conserved target site results in difficulty for bacteria to evolve new structures for lipid II. So far no resistant strains in nature have been reported although McCafferty and co-workers produced a ramoplanin-resistant *S. aureus* strain after serial passages in the presence of ramoplanin-containing medium in 2010 [153].

Table 1.7. Antibiotic activity spectrum

Antibiotics	Important strains	MIC* (mg/l)
	<i>S. aureus</i>	0.03-1.5
Ramoplanin [149, 150]	<i>E. faecalis/faecium</i> Vancomycin-resistant <i>Enterococci</i> (VRE) <i>Streptococcus pneumoniae</i>	0.06-1 0.1-1.5 0.03-0.12
	<i>S. aureus</i>	0.5-2
Vancomycin [151]	<i>E. faecalis/faecium</i> Vancomycin-resistant <i>Enterococci</i> (VRE) <i>S. pneumoniae</i>	0.5-8 Not published 0.12-0.5
	<i>S. aureus</i>	0.78
Enduracidin [152]	<i>E. faecalis/faecium</i> Vancomycin-resistant <i>Enterococci</i> (VRE) <i>S. pneumoniae</i>	Not published Not published Not published

* Note: MIC represents Minimum Inhibitory Concentration which inhibits the visible growth of microorganism after overnight incubation.

1.9.1. The Antibiotic Enduracidin

Enduracidin was first isolated and characterised by Takeda Chemical Industries in 1968 [145, 146] (Figure 1.30). The peptide was separated into two major forms, enduracidin A and B, which differ by one carbon in the length of the attached lipid chain [154, 155]. Structurally, it contains several unusual amino acids including Hpg, Dpg, Orn, Cit, and enduracididine (End). Sixteen of seventeen amino acids form a macrocycle connected by a lactone linkage from the β -hydroxyl group of Thr-2 to the carboxyl group of Hpg-17 [155]. It has been shown to possess a broad spectrum of activity against Gram-positive bacteria including *S. aureus*, *Bacillus subtilis*, and *Bacillus cereus* and has been used as a poultry feed additive for many years [145]. Additionally, it has been shown to be effective in humans against MRSA urinary tract and skin infections [156].

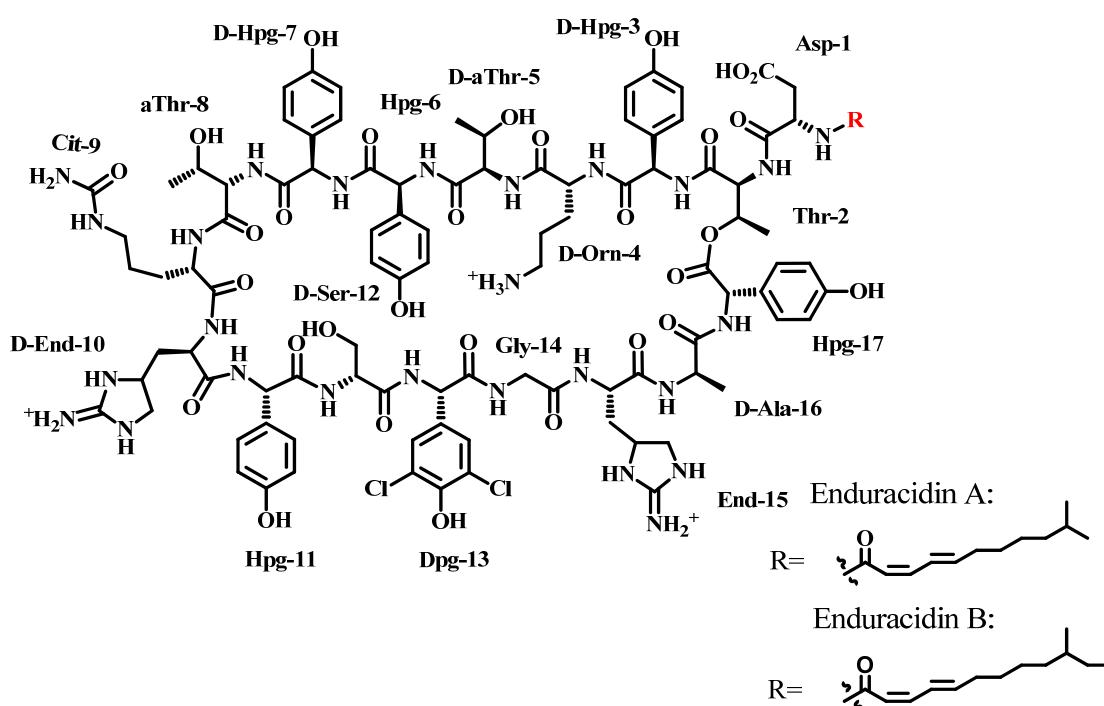


Figure 1.30. Structures of the lipopeptide antibiotic enduracidin A and B. Enduracidin structure with variable fatty acid tails (labelled by R) highlighted in red.

The biosynthetic gene cluster for enduracidin was cloned and sequenced from *S. fungicidicus* ATCC21013 [5]. It is an 84 kb gene cluster located within a 116 kb genetic locus and contains 25 ORFs (*orf22* to *orf46*) (Figure 1.31). As shown in Table 1.8, the region includes genes required for the formation of the lipid side chain, the formation of non-proteinogenic amino acids, assembly of the peptide backbone, enduracidin export and the regulation of biosynthesis.

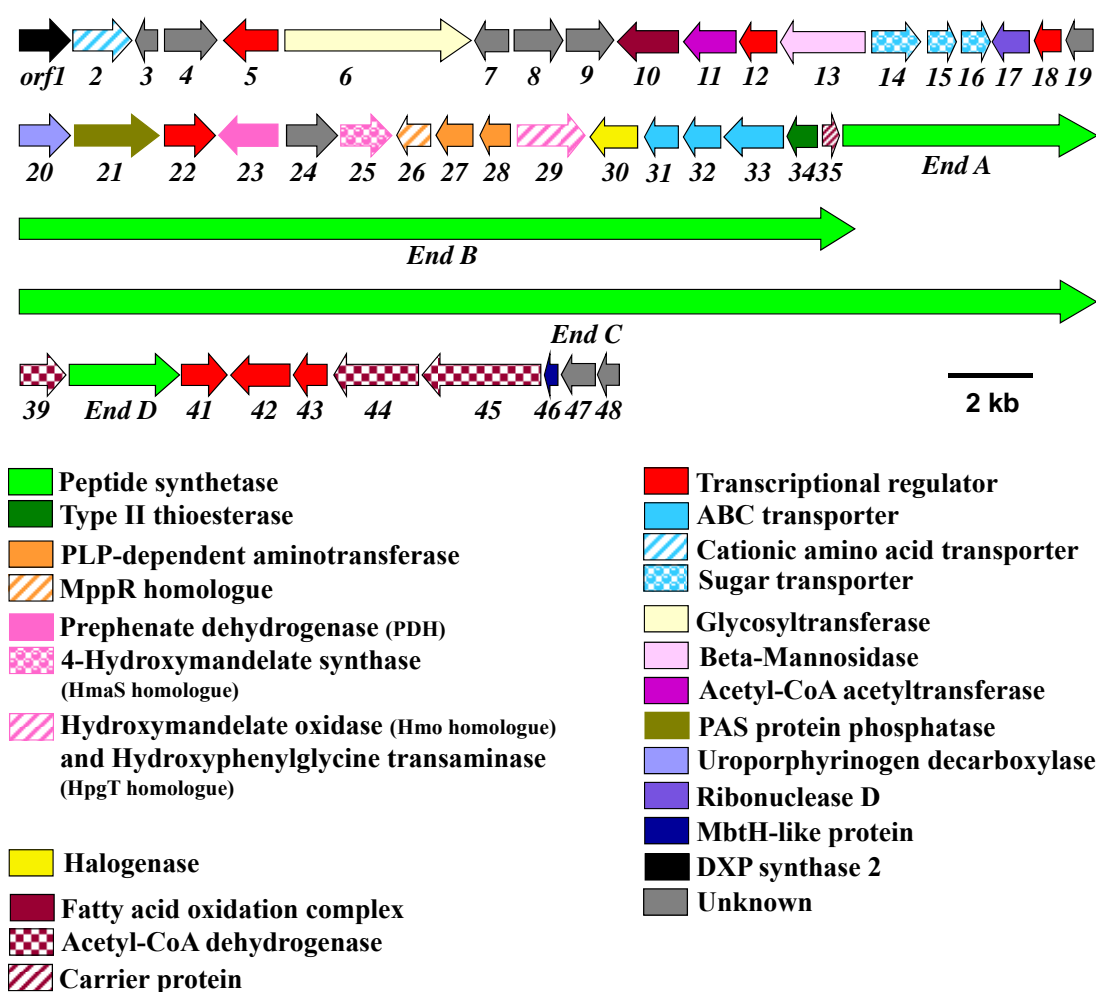


Figure 1.31. Biosynthetic gene cluster of enduracidin.

The *end* gene cluster contains 25 ORFs from *orf22* to *orf46*, whilst the flanking region contains another 23 ORFs (*orf1* to *orf21* and *orf47* and *orf48*).

Table 1.8. The enduracidin biosynthetic gene cluster

	ORF	Gene symbol	Size (a.a.)	Proposed Function	Note
Assembly of the peptide backbone	35		90	Carrier protein (ACP/PCP)	
	36	<i>endA</i>	2101	Peptide synthetase (NRPS)	
	37	<i>endB</i>	6943	Peptide synthetase (NRPS)	
	38	<i>endC</i>	8986	Peptide synthetase (NRPS)	
	40	<i>endD</i>	859	Peptide synthetase (NRPS)	
	34		275	Type II thioesterase	
Non-proteinogenic amino acid formation	30		504	Halogenase	For Dpg
	23		362	Prephenate dehydrogenase	For Hpg
	25		356	4-Hydroxyphenylpyruvate dioxygenase	
	29		808	FMN-dependent α -hydroxyacid dehydrogenase (HmaO homologue) and PLP-dependent class I and II aminotransferase (HpgT homologue)	
	26	<i>endR</i>	279	MppR homologue	For End
	27	<i>endQ</i>	419	PLP-dependent aminotransferase (MppQ homologue)	
	28	<i>endP</i>	293	PLP-dependent aminotransferase (MppP homologue)	
Lipid formation	39		274	Acyl-CoA dehydrogenase/reductase	
	44		625	Acyl-CoA dehydrogenase	
	45		1177	Acyl-CoA ligase/dehydrogenase fusion protein	
Export	31		341	Transmembrane transport protein	
	32		307	ABC transporter	
	33		651	ABC transporter	
Regulation	22		328	StrR-like regulatory protein	
	41		223	Two-component response regulator	
	42		370	Two-component system sensor kinase	
	43		181	Two-component system sensor kinase	
Unknown	24		311		
	46		71		

Adapted from Yin and Zabriskie (2006) [5].

Assembly of the 17-residue enduracidin peptide backbone is catalysed by four synthetases designated End A, End B, End C and End D [5]. As shown in Figure 1.32, the organisation of the modules and domains in these four synthetases follows the linear NRPS modular architecture principle [62, 64]. The most striking feature is that the 6th module of End B that should incorporate Thr-8 lacks an A domain. This module is presumed to be loaded *in trans* by the single module protein End D which is predicted to activate Thr. Another interesting feature is that enduracidin contains seven D-amino acids but no epimerization (E) domains and no epimerase/racemase genes were found in its NRPS gene cluster. Balibar *et al.* (2005) demonstrated that the D-amino acids in the biosurfactant arthroactin are generated by a C domain which has the dual functions of condensation and epimerization [157]. Sequence analysis revealed that C domains that follow the modules corresponding to the D-amino acids of enduracidin belong to a unique subset of C domains. Such C domains probably also have the dual functions of condensation and epimerization [5]. Finally, the putative chlorinase End30 is predicted to chlorinate Hpg tethered to the PCP domain of module 4 of End C.

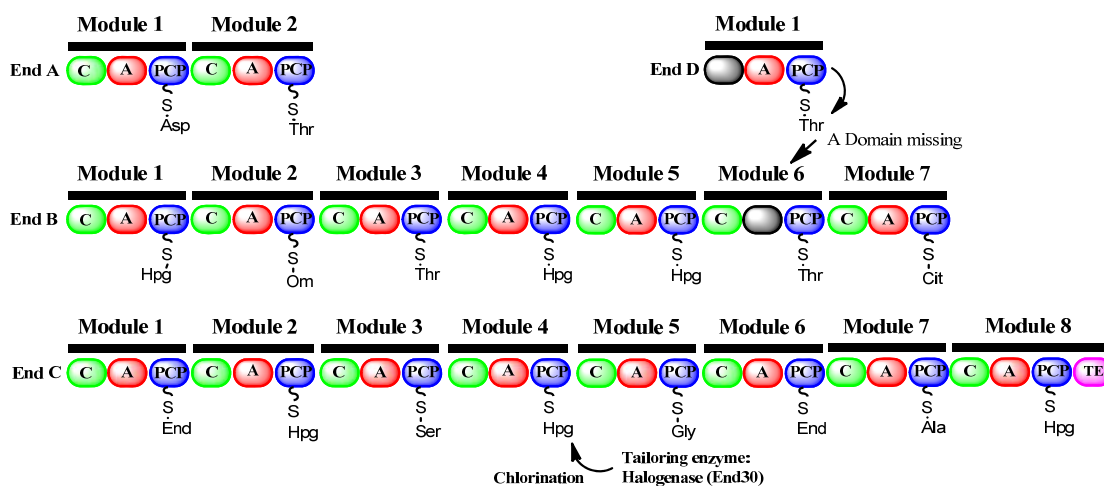


Figure 1.32. Organisation of the enduracidin NRPS.

1.9.2. The Antibiotic Ramoplanin

Ramoplanin was discovered in 1984 by the company of Biosearch Italia in a screening program with the aim of identifying novel antibiotic activities acting against biosynthesis of the bacterial cell wall [147]. Three major forms of ramoplanin were identified as A1, A2 and A3, which differ only in the length of the *N*-terminal acyl chain (Figure 1.33). Additionally three other components in the fermentation broth were identified later and designated as A'1, A'2 and A'3, which differ from A1, A2 and A3 by the absence of one mannose unit from the glycosidic group [158].

Ramoplanin has a very similar structure to enduracidin and also contains the same abundance of Hpg residues in both D and L configurations, in addition ramoplanin also contains several other unusual amino acids including Chp, Orn and hAsn. Sixteen of the seventeen amino acids form a macrocyclic ring by a lactone linkage between the β -hydroxyl group of hAsn-2 and the carboxyl group of Chp-17.

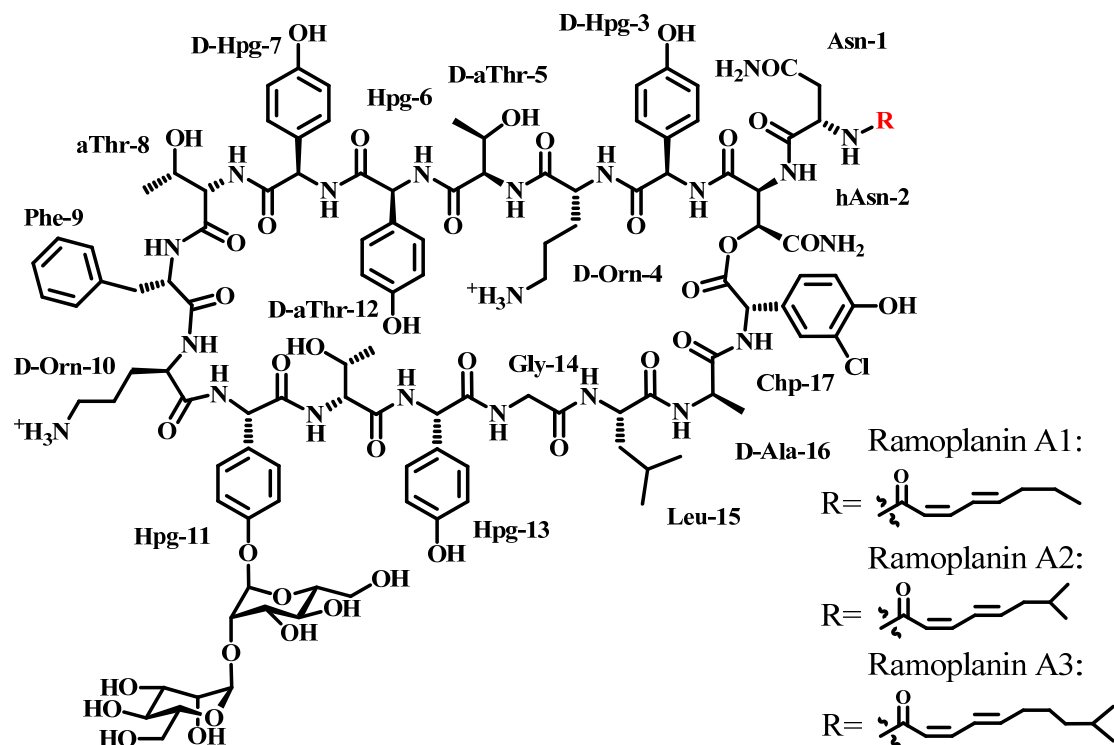


Figure 1.33. Structures of the glycolipopeptide ramoplanin A1, A2 and A3.

Ramoplanin structure with three variable fatty acid tails (labelled by R) highlighted in red.

Nanotherapeutics, Inc. acquired the patent rights of ramoplanin from Oscient Pharmaceuticals Corporation in 2010. This antibiotic has been further developed by Nanotherapeutics as an oral antibiotic drug against *Clostridium difficile*-associated disease (CDAD) and is entering Phase III clinical trials. It also shows a broad spectrum of activity against Gram-positive bacteria including the genera of *Enterococci*, *Staphylococci*, *Bacilli*, *Streptococci* and *Listeria* [6].

The 88.5 kb biosynthetic gene cluster of ramoplanin was cloned and sequenced from *Actinoplanes* ATCC33076 [159]. It contains 33 ORFs which are responsible for ramoplanin biosynthesis, including the formation of the lipid side-chain, the synthesis of non-proteinogenic amino acids, assembly of the peptide backbone, antibiotic export/resistance, and the regulation of antibiotic production (Figure 1.34, Table 1.9).

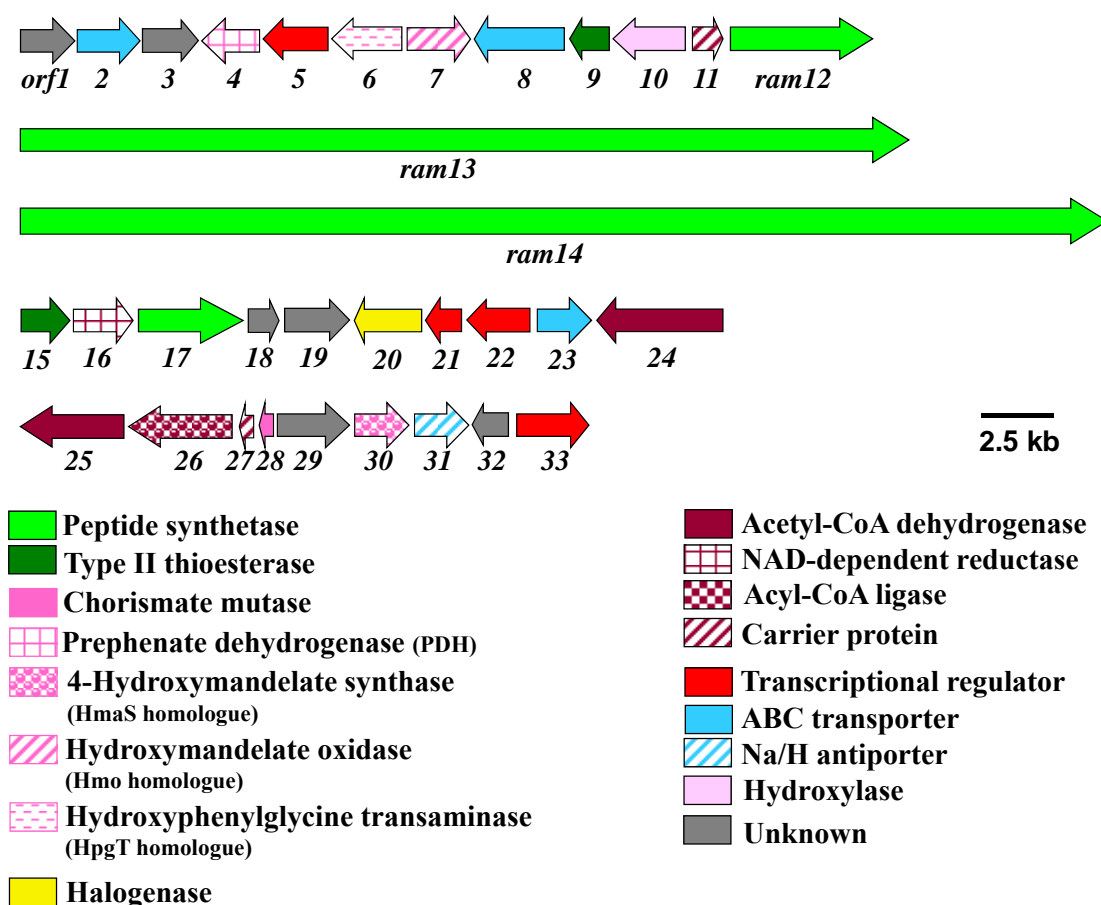


Figure 1.34. Biosynthetic gene cluster of ramoplanin.
The *ram* gene cluster contains 33 ORFs.

Table 1.9. The ramoplanin biosynthetic gene cluster

	ORF	Gene symbol	Size (a.a.)	Proposed Function	Note
Assembly of the peptide backbone	11	<i>ram11</i>	90	Carrier protein (ACP/PCP)	
	12	<i>ram12</i>	1051	Peptide synthetase (NRPS)	
	13	<i>ram13</i>	6893	Peptide synthetase (NRPS)	
	14	<i>ram14</i>	8695	Peptide synthetase (NRPS)	
	15	<i>ram15</i>	234	Type II thioesterase	
	17	<i>ram17</i>	891	Peptide synthetase (NRPS)	
Non-proteinogenic amino acid formation	10		529	Hydroxylase	For hAsn
	20	<i>ram20</i>	491	Halogenase	For Chp
	29	<i>ram29</i>	619	Transmembrane protein	
	4		283	Prephenate dehydrogenase	For Hpg
	6		444	Aminotransferase	
	7		356	Hydroxymandelate oxidase	
	28		94	Chorismate mutase	
	30		355	Hydroxymandelate synthase	
Lipid formation	9		271	Type II thioesterase	
	16		274	NAD dependent ACP reductase	
	24		553	Acyl-CoA dehydrogenase	
	25		585	Acyl-CoA dehydrogenase	
	26	<i>ram26</i>	587	Acyl-CoA ligase	
	27		75	Carrier protein (ACP/PCP)	
Antibiotic export and resistance	2		204	ABC transporter	
	8		640	ABC transporter	
	23		309	ABC transporter	
	31		429	Na/H antiporter	
Regulation of antibiotic production	5		336	Regulatory protein	
	21		217	Two-component response regulator	
	22		403	Two-component response regulator	
	23		309	ABC transporter	
	33		309	Regulator	
Unknown	1		333		
	3		321		
	18		187		
	19		415		
	32		189		

Adapted from McCafferty *et al.* (2002) [6].

Four peptide synthetases, designated Ram12, Ram13, Ram14 and Ram17, are responsible for the assembly of the ramoplanin peptide core (Figure 1.35) [6]. The organisation of the modules and domains in these four synthetases follows the linear NRPS modular architecture principle [62, 64]. Ram12 consists of a single module and is suggested to function iteratively, activating and condensing two Asn residues to generate a dipeptide. It is predicted that the second Asn is subsequently hydroxylated by Ram10 when Asn is tethered to the PCP domain. Ramoplanin shares two unusual features with enduracidin. Firstly the 6th module of Ram13 lacks an A domain which should incorporate Thr-8. This module is presumed to be loaded *in trans* by the single module protein Ram17, which is predicted to activate Thr. Secondly ramoplanin contains an abundance of D-amino acids but there are no epimerization (E) domains and no epimerases/racemases encoded in its NRPS gene cluster. So it is again presumed that C domains corresponding to the D-amino acids contain the dual functions of condensation and epimerization. Additionally the putative chlorinase Ram20 is predicted to chlorinate Hpg whilst it is tethered to the PCP domain of module 8 of Ram14.

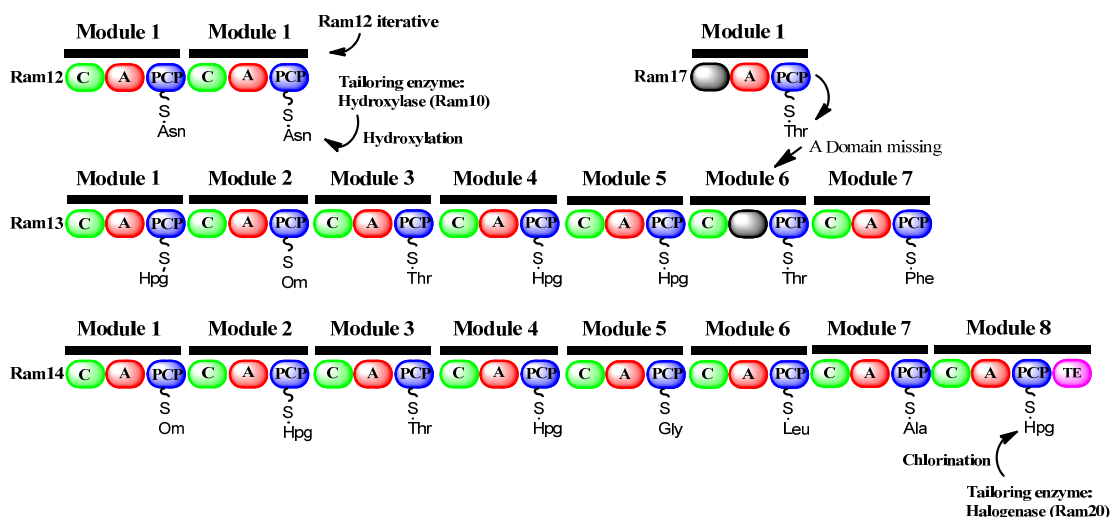


Figure 1.35. Organisation of the ramoplanin NRPS.

1.10. Riboswitches as Gene Regulation Tools

Riboswitches may be considered as metabolite-sensing RNAs [160]. They are RNA elements, generally located upstream of genes in the 5'-UTR, which respond to physiological signals to control gene expression by *cis*-acting regulation [13, 160, 161]. These elements are widespread across the bacterial kingdom, and have also recently been identified within introns and in the 3'-UTRs of eukaryotic transcripts [162]. Structurally, riboswitches consist of two domains: the **aptamer domain** and the **expression platform** [13, 163] (Figure 1.36). The aptamer domain is responsible for recognition and binding of the metabolite ligands. The expression platform contains the components required for the regulation of gene expression, such as an intrinsic transcriptional terminator or an anti-SD sequence in a translational repressor stem, which govern gene transcription and translation respectively. Thus most riboswitches can be generally classified into two groups, transcriptionally-acting riboswitches and translationally-acting riboswitches [163]. Upon binding of the ligand, the aptamer domain dictates the state of the expression platform as to whether it is “ON” or “OFF”. In the “ON” state, transcription or translation proceeds; whilst in the “OFF” state, transcription or translation is repressed.

In the case of transcriptionally-acting riboswitches, the expression platforms often form anti-terminator structures in the “ON” state (Figure 1.36). In the “OFF” state, the expression platforms form an intrinsic transcriptional terminator -- a stable GC-rich stem loop structure, normally followed by 6-10 consecutive uridines -- which stalls the RNA polymerase and promotes its dissociation from the DNA template thereby halting transcription [164, 165]. In the translationally-acting riboswitches, the expression platform can control the accessibility of the SD sequence, a short purine-rich element that recruits the 30S ribosomal subunits to the mRNA near the start codon, to allow or prevent the initiation of mRNA translation in the “ON” or “OFF” states respectively [13] (Figure 1.36).

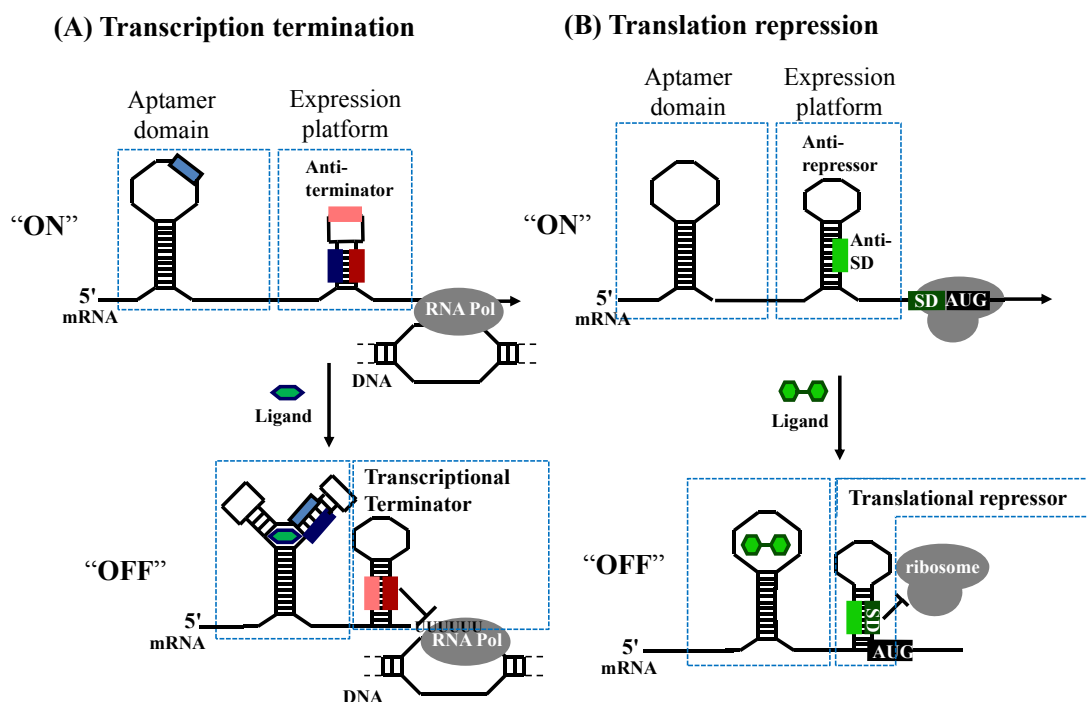




Figure 1.36. Riboswitch mechanisms.

Schematic diagram of (A) a transcriptionally-acting riboswitch, and (B) a translationally-acting riboswitch. In the absence of ligand, the expression platform forms an anti-terminator loop (A) or sequesters the anti-SD sequence (B) enabling the RNA polymerase or the ribosome to proceed. In the presence of ligand ( or ), the aptamer domain changes its structural configuration, altering the structure of the expression platform leading to the formation of a transcriptional terminator (A) or a translational repressor (B), preventing transcription or translation of the gene respectively. Note: examples also exist of ligand-binding turning riboswitch regulated genes “ON”, through both transcriptional and translational mechanisms.

1.10.1. Diversity of Riboswitches

There are at least 20 distinct classes of naturally-occurring riboswitches, recognising a broad range of small molecule ligands such as: **tRNA** [166]; **purine nucleobases** (adenine [167], guanine [168], 2'-deoxyguanosine [169] and queuosine precursor (preQ₁) [170]); **amino acids** (L-Gln [171], L-Gly [172], L-Lys [173] and L-Arg [174]); **aminosugars** (glucosamine-6-phosphate [175]); **secondary messenger molecules** (cyclic di-GMP [176]); **metal ions** (Mg²⁺) [177]; and other **enzyme cofactors** such as vitamin B12 [178], thiamine pyrophosphate (TPP) [179], flavin mononucleotide (FMN) [180], SAM [181], S-adenosylhomocysteine (SAH) [182], and molybdenum cofactor [183]. Due to their need to respond to diverse ligands, the aptamer domains of riboswitches have evolved diverse structures. For example, the smallest riboswitch

is composed of 34 nucleotides (nt) and binds preQ₁ to regulate the biosynthesis of queuosine in response to this queuosine precursor [170]. Some of the largest and most complex riboswitches are the amino acid-responsive riboswitches (Lys and Gly), with aptamer domains that are composed of 200 nt or more, and which directly sense amino acids concentrations to control their flux through metabolism in bacteria [184]. SAM-binding riboswitches display the greatest structural diversity and have been grouped into five classes: the SAM-I (S-box) [181], SAM-II [185], SAM-III (SMK-box) [186], SAM-IV [187] and SAM-V riboswitches [188]. All five classes of SAM-riboswitches recognise the same ligand SAM, but with different aptamer structures. Another interesting group of riboswitches is the purine- and purine-derivative-sensing riboswitches. They share conserved secondary structure elements but have different ligand specificities to discriminate between similar ligands, typically by recognition through canonical Watson-Crick base pair such as A-U or G-C [189].

In addition to responding to tRNA, metal ions and metabolites, riboswitches can also sense temperature (Figure 1.37). One such thermosensing riboswitch, named the repressor of heat-shock gene expression (ROSE) element, has been characterised [190]. It contains a simple helical domain which sequesters the SD sequence and AUG start codon to repress translation at normal cellular temperature. However higher temperatures destabilise the secondary structure in the SD region to allow ribosome binding and subsequent expression of heat-shock gene products [190]. Another thermosensor riboswitch is involved in virulence of the food-borne pathogen *Listeria monocytogenes* [191]. Its thermoregulation determines the expression of *prfA*, a transcriptional activator that regulates expression of invasion proteins. At 30 °C this riboswitch switches the expression of PrfA protein off by sequestering the SD sequence, resulting in low virulence. However, when *L. monocytogenes* enters a mammalian host with a core temperature of 37 °C, the riboswitch releases the SD sequence, resulting in the activation of PrfA protein expression and subsequent pathogenicity.

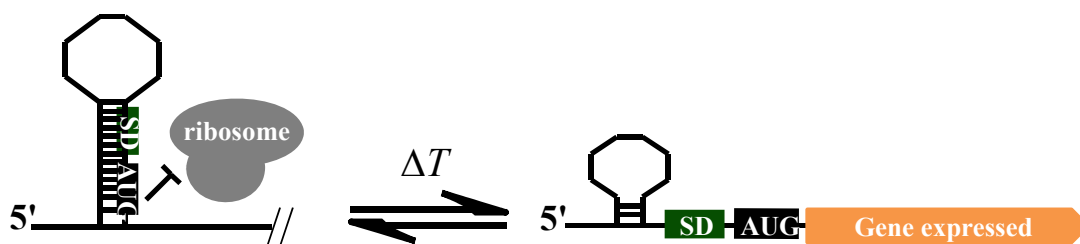


Figure 1.37. Thermosensing riboswitches.

The SD sequence and the AUG start codon are sequestered at lower temperatures, resulting in inhibition of gene translation. However they are released at higher temperatures, allowing for the initiation of mRNA translation.

1.10.2. Riboswitch-Based Applications

1.10.2.1. Riboswitches as Synthetic Biology Tools

The aptamer and expression platform domains of riboswitches are modular, thus potentially allowing for modification of the ligand-binding specificity of the aptamer domain without affecting the expression platform functionality. For example, in the aptamer core of purine riboswitches, a highly conserved nucleotide plays a critical role in ligand recognition by Watson-Crick base-pairing. Such nucleotides include the cytidine in guanine riboswitches and uridine in adenine riboswitches [189]. It has been demonstrated that replacing this conserved nucleotide in the wild-type guanine and adenine riboswitches can switch the specificity to the opposite purine ligand (*i.e.* A to G or G to A ligand specificity changes) [167]. Thus riboswitches present good candidates which can be engineered and used as novel gene-expression tools. Additionally they can be used as biosensors to monitor the presence of metabolites in solution by combining them with reporter genes, such as the genes for green fluorescent protein, luciferase or β -galactosidase [163].

1.10.2.2. Riboswitches as Anti-Bacterial Targets

Due to the conserved ligand-binding capacity of the aptamer domain and the ligand specificity, it is possible to design ligand analogues that bacteria cannot metabolize, targeting riboswitches that regulate the expression of genes essential for bacterial

survival or virulence [192]. For example, in *S. aureus* one possible target gene is *metK*, which encodes a SAM synthetase predicted to be essential for both survival and virulence [193, 194]. This gene is regulated by the class I SAM-binding riboswitch [181, 195]. Therefore the design of SAM-metabolite analogues which can bind to the riboswitch provides the possibility of inhibiting the growth and virulence of *S. aureus* by shutting down essential gene expression. Another example is the drug pyrithiamine (PT), an analogue of thiamine, which can inhibit the growth of bacteria and fungi [196, 197]. Like thiamine, it is readily phosphorylated to pyrithiamine pyrophosphate (PTPP) upon entering the cell. Owing to the similarity in structure between PTPP and thiamine pyrophosphate (TPP), it was found that PTPP could also bind to several TPP riboswitches *in vitro* [198]. From this, it has been postulated that the mode of action of PT *in vivo* is through its binding to TPP riboswitches and the resulting repression of one or more TPP riboswitch-regulated genes. Riboswitches are therefore potential target sites for the development of antibiotic compounds.

1.11. Aims of the Project

Antibiotic modifications such as hydroxylation, chlorination and glycosylation are closely related to antibiotic activity. Such modifications are difficult to achieve by chemical synthesis. The work undertaken here attempts to advance the knowledge on the enzymatic modification of amino acid residues in antibiotics, and to use this knowledge for producing novel variants of peptide antibiotics. In parallel, the work undertaken for this thesis is also concerned with exploring riboswitches as biosynthetic tools to meet the high demand for good protein expression systems for the functional characterisation of tailoring enzymes for antibiotic engineering. This thesis is focused on (1) the tailoring enzymes responsible for the biosynthesis of OmAsp in the antibiotic A54145; (2) two modifications, chlorination and mannosylation, of the antibiotic ramoplanin; (3) Re-engineering riboswitches as gene expression tools.

1.11.1. The Biosynthetic Origin of OmAsp

The non-proteinogenic amino acid OmAsp at position 9 in A54145 is important for antibiotic activity [133]. It is postulated to be synthesised by two sequential reactions *i.e.* hydroxylation of the β -carbon of L-Asp, followed by methylation of this hydroxyl group. Three tailoring enzymes, two hydroxylases and an MTase, from the A54145 gene cluster were over-expressed and assayed *in vitro* with the free substrate amino acids L-Asp and L-OH Asp respectively. The antibiotic CDA, produced by *S. coelicolor*, possesses similar primary and secondary structure to A54145 and contains the Asp residue at the same relative position as OmAsp in A54145. Thus a transformant *S. coelicolor* with over-expression of both enzymes was expected to produce variant CDA which contains OmAsp.

1.11.2. The Bio-engineering of the Lipoglycopeptide Antibiotic Ramoplanin by Chlorination and Mannosylation

Chlorination and mannosylation have been proven to play important roles in antibiotic activity. The chlorination and mannosylation modifications on the 17th and 11th Hpg residues of the antibiotic ramoplanin respectively are predicted to be catalysed by the putative halogenase Ram20 and mannosyl transferase Ram29 from the ramoplanin biosynthetic gene cluster. Ramoplanin shares similar primary and secondary structure to enduracidin. Due to the slow growth rate of ramoplanin producing strain, we characterised both enzymes in the enduracidin producing strain. Here we set up transformant strains of *S. fungicidicus* with the genes *ram20* or *ram29*, with the aim of producing chlorinated enduracidins or mannosylated enduracidins when the genes were over-expressed.

1.11.3. Re-engineering PreQ₁ Riboswitches

The preQ₁ riboswitch is a transcriptional “off” riboswitch that is regulated by the metabolite 7-aminomethyl-7-deazaguanine (preQ₁). On commencing this research, the preQ₁ riboswitch was fused with the reporter gene *lacZ* and a mutagenesis strategy was applied to the preQ₁-*lacZ* constructs in order to understand the role of the preQ₁ ligand binding residues and to aid in the design of suitable ligands for artificial preQ₁ riboswitches. The engineered preQ₁ riboswitches were expected to respond in a dose-dependent manner to the non-natural synthetic molecules rather than to the natural metabolite preQ₁, and thus provide a regulatory system for controlling gene expression.

Chapter 2. The biosynthetic origins of 3-O-methyl aspartic acid (OmAsp)

2.1. Introduction

Detailed information about the biosynthesis, structure and mode of action of the three Ca^{2+} -dependent cyclic antibiotics CDA, daptomycin and A54145 is given in section 1.8. All three antibiotics contain the L-Asp residue but only in A54145 has this residue been enzymatically modified to 3-O-methyl aspartic acid (OmAsp) [2-4] (Figure 1.21). Regarding stereochemistry of OmAsp, the published literatures on A54145 does not address amino acid stereochemistry. Miao *et al.* (2006) proposed the stereochemistry of A54145 according to the studies of CDA and daptomycin and only showed OmAsp as L-amino acid [2]. We proposed that these enzymatic modifications involve an initial hydroxylation of the β -carbon of L-Asp, followed by methylation of the hydroxyl group, forming the methylether. This biosynthetic pathway is of significance due to the difficulties in accessing the β -methylene group for hydroxylation by classical chemical methods. Additionally, it would be interesting to investigate whether this biosynthetic pathway can be applied to other antibiotic biosynthetic systems such as CDA and daptomycin.

Amino acid modifications can be carried out by specific *cis*-acting domains embedded within NRPS modules or by *trans*-acting tailoring enzymes. In the analysis of the A54145 NRPS LptC module 2, corresponding to OmAsp, we did not find any candidate domain that could be responsible for the modification of L-Asp. Therefore we assumed that the non-standard amino acid OmAsp is synthesised by *trans*-acting modification enzymes, that is, hydroxylation of the β -carbon of L-Asp by a hydroxylase, followed by methylation of the hydroxyl group by an *O*-methyl transferase (*O*-MTase) (Figure 2.1).

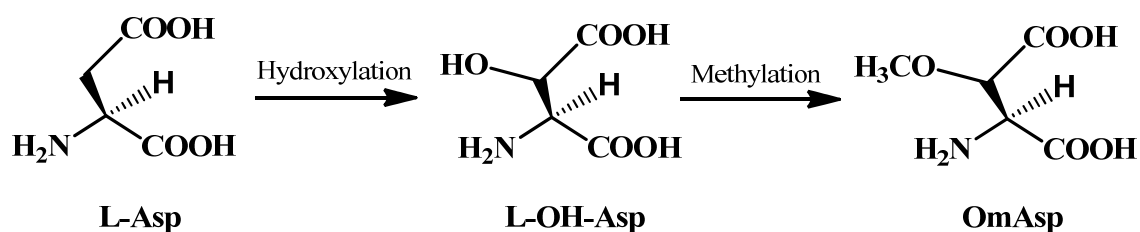


Figure 2.1. Proposed biosynthetic pathway for OmAsp.

Sequence analysis of the A54145 biosynthetic gene cluster showed that there are two putative hydroxylases (*lptJ* and *lptL* genes) and two MTases (*lptK* and *lptI* genes). The MTase LptI has been previously proven to be a SAM-dependent α -KG C-MTase which catalyses the transformation of α -KG to 3-methyl-2-oxoglutarate [118]. Therefore LptJ/LptL and LptK are the best candidate enzymes for hydroxylation and methylation respectively within the biosynthesis of OmAsp.

2.2. LptL: Non-heme Fe(II)/ α -KG-dependent hydroxylase

2.2.1. Bioinformatic Analysis of LptL

The blast analysis of the LptL amino acid sequence shows sequence homology to AsnO (47 % identity), the putative oxygenase TobO (43 %) and the clavamate synthase CAS (38 %) (Table 2.1). AsnO has been shown to be a Fe(II)/ α -KG-dependent oxygenase, which catalyses the β -hydroxylation of the free amino acid L-Asn during CDA biosynthesis in *S. coelicolor* [95, 111]. TobO is a putative oxygenase involved in the biosynthesis of the aminoglycoside tobramycin. CAS is also an oxygenase which catalyses three separate Fe(II)/ α -KG-dependant oxidative steps in the synthesis of clavulanic acid [199, 200].

Table 2.1. Identification of LptL function by a Blast protein sequence search

Protein name	Description	Accession number	Identity (%)
AsnO*	Oxygenase from <i>S. coelicolor</i> A3(2); involved in hydroxylation of L-Asn	NP_627448.1	47
TobO	Putative oxygenase TobO from <i>Streptomyces sp.</i> DSM 40477; involved in tobramycin biosynthesis	CAH18567.1	43
CAS	Clavamate synthase from <i>Synechococcus sp.</i> JA-2-3B'a (2-13)	YP_478072.1	38
CAS	Clavamate synthase from <i>S. clavuligerus</i> ATCC 27064	YP_002195752.1	36

*Crystal structure (2OG6) is available.

Usually the non-heme Fe(II)-dependent hydroxylases contain the conserved '2-His-1-carboxylate facial triad' motif for iron binding. This conserved triad (His-147, Glu-149 and His-279) can also be identified in the amino acid sequence alignment of LptL with that of AsnO and CAS using the online alignment program T-coffee [201] (Figure 2.2). According to the arrangement of the iron binding motif, hydroxylases can be further classified into three subgroups (Table 2.2): subgroup I has an HXDX₅₀₋₇₀HX₁₀RXS motif, *e.g.* isopenicillin *N* synthase (IPNS) and deacetoxycephalosporin C synthase (DAOCS); subgroup II has an HXDX₁₃₈₋₂₀₇HX₁₀₋₁₃R motif, *e.g.* taurine/ α -KG dioxygenase and CAS; subgroup III has an HXDX₇₂₋₁₀₁HX₁₀RXS motif *e.g.* phytanoyl-CoA hydroxylase and proline hydroxylase [202]. The arrangement of the iron-binding motif of LptL matches with that of subgroup II of this superfamily of Fe(II)-dependent hydroxylases.

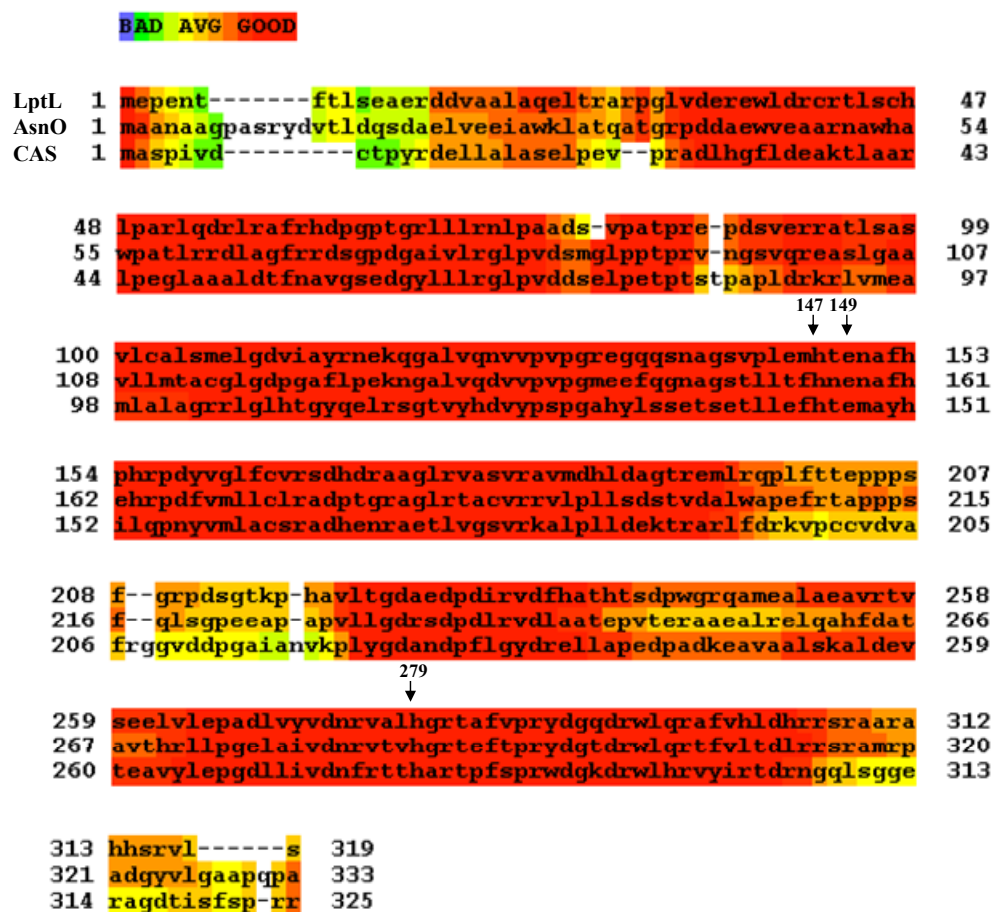


Figure 2.2. The LptL amino acid sequence alignment with AsnO and CAS.

The amino acid sequence of LptL is aligned with that of AsnO and CAS from *S. coelicolor* and *S. clavuligerus* respectively. Arrows (↓) highlight the three conserved residues necessary for iron binding, His-147, Glu-149 and His-279.

Table 2.2. Subgroups of Fe(II)-dependent hydroxylases

Subgroup	Consensus Motif		
Subgroup I	HXD	X_{50-70}	$HX_{10}(R/K)XS$
Subgroup II	HX(D/E)	$X_{138-207}$	$HX_{10-13}R$
Subgroup III	HXD	X_{72-101}	$HX_{10}(R/K)XS$

Note: The subscript number represents the number of amino acids. Adapted from Hogan *et al.* (2000) [202].

A 3D-homology model of LptL was created by the Swiss-Modelling Server, based on the AsnO crystal structure (2OG7) published by Strieker *et al.* in 2007 [203-205] (Figure 2.3A). The superimposition image of AsnO and LptL shows that they share similar secondary structures (Figure 2.3B). In the active site of AsnO, the residues Asp-241, His-155 and His-287 are responsible for iron binding, and the

substrate/product binding-pocket residues can be identified as Glu-125, Gln-144, Asn-146, Asn-158, Asp-241 and Arg-305 (Figure 2.3C-D). Additionally the conserved residue Arg-301 can form a salt bridge to succinate which is derived from the co-substrate α -KG (Figure 2.3C-D). Surprisingly, these residues in the AsnO active site for binding iron, substrate/product and α -KG/succinate are identical to those of LptL (Figure 2.3C-D).

The above bioinformatic analyses including the amino acid blast search, the iron-binding motif search and the arrangement of the iron-binding motif analysis, all suggest that LptL is an Fe(II)/ α -KG-dependent hydroxylase. From *in silico* modelling, it was predicted that LptL is a good candidate for the hydroxylation of L-Asn. We propose here that LptL is also capable of hydroxylating L-Asp, given that OmAsp is found within the A54145 antibiotic.

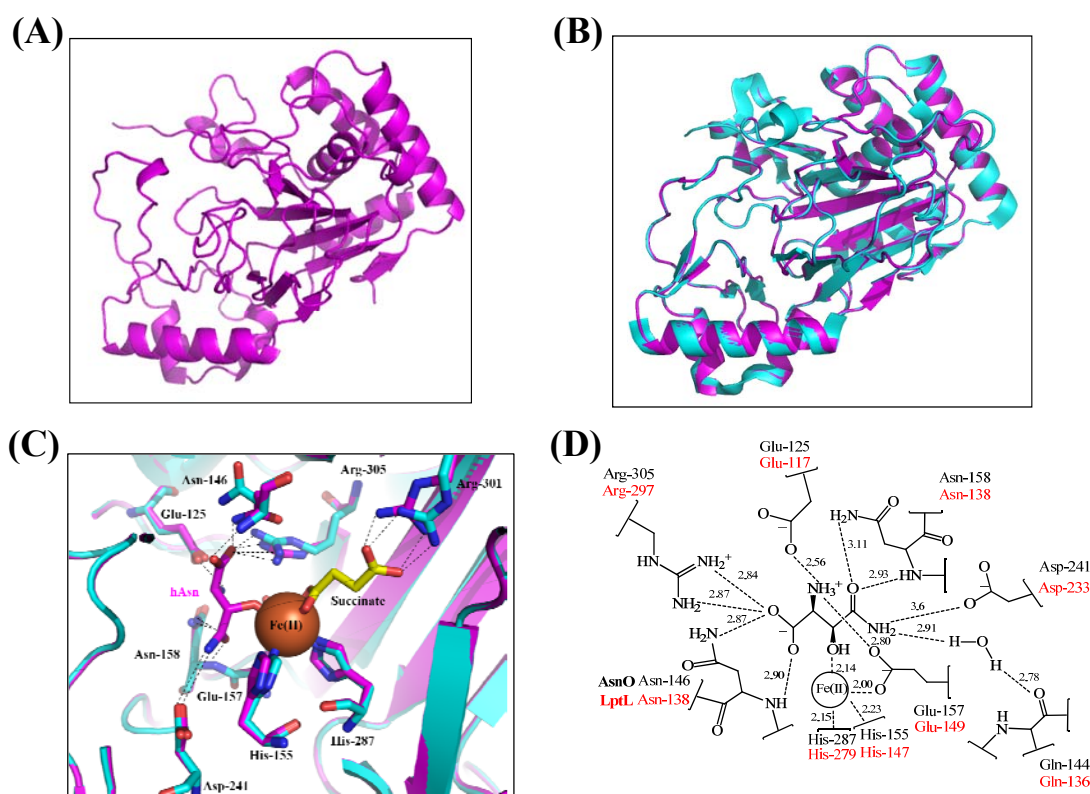


Figure 2.3. Comparison of AsnO and LptL.

(A) 3D-Homology model of LptL based on the AsnO-product complex (2OG7). (B) Superimposition of the structures of AsnO (blue) and LptL (pink). (C) Schematic representation of product-enzyme interactions within the active sites of AsnO and LptL (the numbering of amino acids is taken from the AsnO structure). (D) The interactions of the β -OH-Asn (hAsn) product with the AsnO enzyme are shown to highlight the substrate binding residues of AsnO (black) and LptL (red). (Adapted from Strieker *et al.* (2007)) [111].

2.2.2. Heterologous Expression of *lptL* in *E. coli*

The *lptL* gene was successfully amplified from the genomic DNA of *S. fradiae* with primers *lptL*-F1 and *lptL*-R1 by PCR (Figure 2.4A; PCR conditions and primer information are given in section 6.10). The *lptL* gene fragment of approximately 1000 bp was digested with the restriction enzymes *Nde*I and *Xho*I, then cloned into the pET30b expression vector and expressed with a C-terminal His₆-tag in *E. coli* BL21 (DE3) (the procedures for DNA electrophoresis, DNA extraction from agarose, and DNA digestion, ligation and transformation are given in section 6.11, 6.12 and 6.13). After the sequence was verified by sequencing, the recombinant protein LptL was over-expressed by induction with IPTG and purified by Ni²⁺ affinity chromatography (The detailed procedures are given in section 6.14). As shown in Figure 2.4B, an intense band of approximately 35 kDa was observed by SDS-PAGE, which is consistent with the theoretical molecular weight calculated from the LptL protein sequence analysis. The protein yield was 0.125 mg from a 500 ml culture.

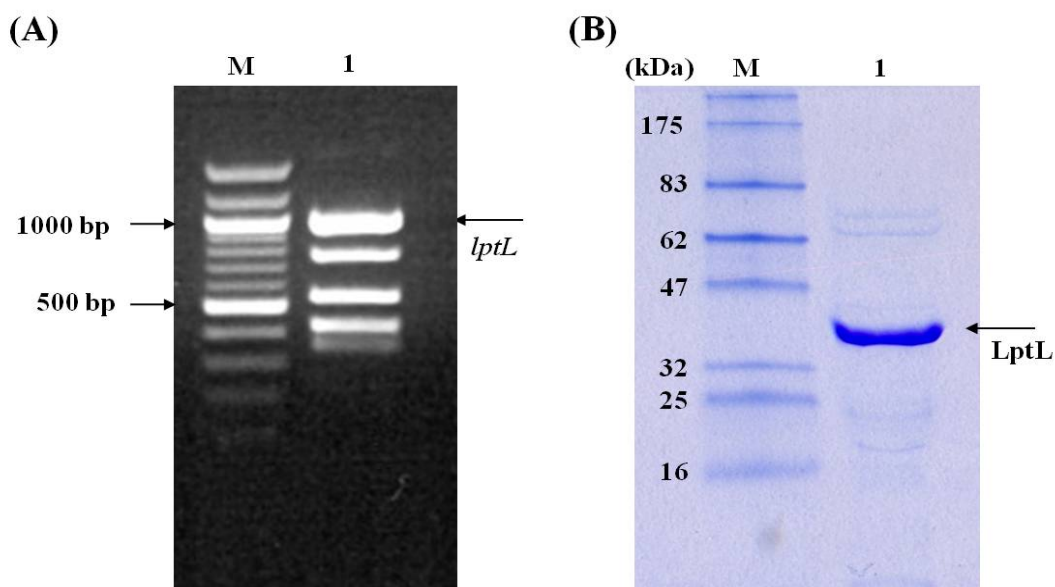


Figure 2.4. Agarose and polyacrylamide gel electrophoretic analyses of the gene *lptL* and protein LptL.

(A) DNA electrophoretic analysis of the gene *lptL* amplified by PCR. Lane M: DNA size marker. Lane 1, PCR products. The arrow indicates the *lptL* DNA fragment band (approximately 1000 bp). (B) SDS-PAGE analysis of the purified LptL. Lane M: Protein molecular weight marker. Lane 1: The sample eluted from the Ni²⁺ affinity column in 250 mM imidazole. The arrow indicates the LptL band (approximately 35 kDa).

2.2.3. Hydroxylase Activity Assay

The free amino acid L-Asn was first used to assay for LptL activity with the cosubstrate α -KG and cofactor iron in HEPES buffer pH 7.5. Reaction was initiated by adding enzyme and incubating at 28 °C for 18 h. To facilitate identification and detection of the modified amino acids, we derivatised the amino acids after the enzyme reaction with phenylisothiocyanate (PITC), followed by reverse phase-high pressure liquid chromatography (RP-HPLC) analysis. The details of the enzyme assay and amino acid derivatisation procedures are given in section 6.25 and 6.26.1 respectively. As shown in Figure 2.5, the substrate L-Asn and product from the incubation of L-Asn with LptL were derivatised with PITC and followed by RP-HPLC analysis. The chromatogram showed that the retention time of the derivatised product (approximately 19 min; Figure 2.5A) was shorter than that of the derivatised substrate (PTC-L-Asn; approximately 24 min; Figure 2.5D), indicating that the product is more polar than the starting material L-Asn (Figure 2.5). The fractions from RP-HPLC which contained products were analysed by mass spectrometry. The mass spectrometric analysis showed that the mass of the product is larger than that of the L-Asn starting material by 16 Da (Figure 2.6), consistent with the product being hydroxylated. Additionally, we showed that L-Asn cannot be hydroxylated by LptL without the cosubstrate α -KG and cofactor iron, demonstrating that LptL is α -KG and iron dependent hydroxylase (Figure 2.5 B and C). As for the other potential amino acid substrates D-Asn and L-Asp, we did not find any new products in the RP-HPLC chromatogram after incubation with LptL (Figure 2.7). These experiments showed that LptL is able to stereospecifically hydroxylate L-Asn and provides the non-proteinogenic amino acid OH-Asn for the biosynthesis of antibiotic A54145. Strieker *et al.* (2007) published the X-ray structure of AsnO and showed that the stereochemistry of L-OH-Asn is (2S, 3R)-*trans* configured [111]. Due to the substrate binding site of LptL being absolutely identical to that of AsnO (Figure 2.3), we predicted that the product of LptL would also be L-(2S, 3R)-*trans*-OH-Asn.

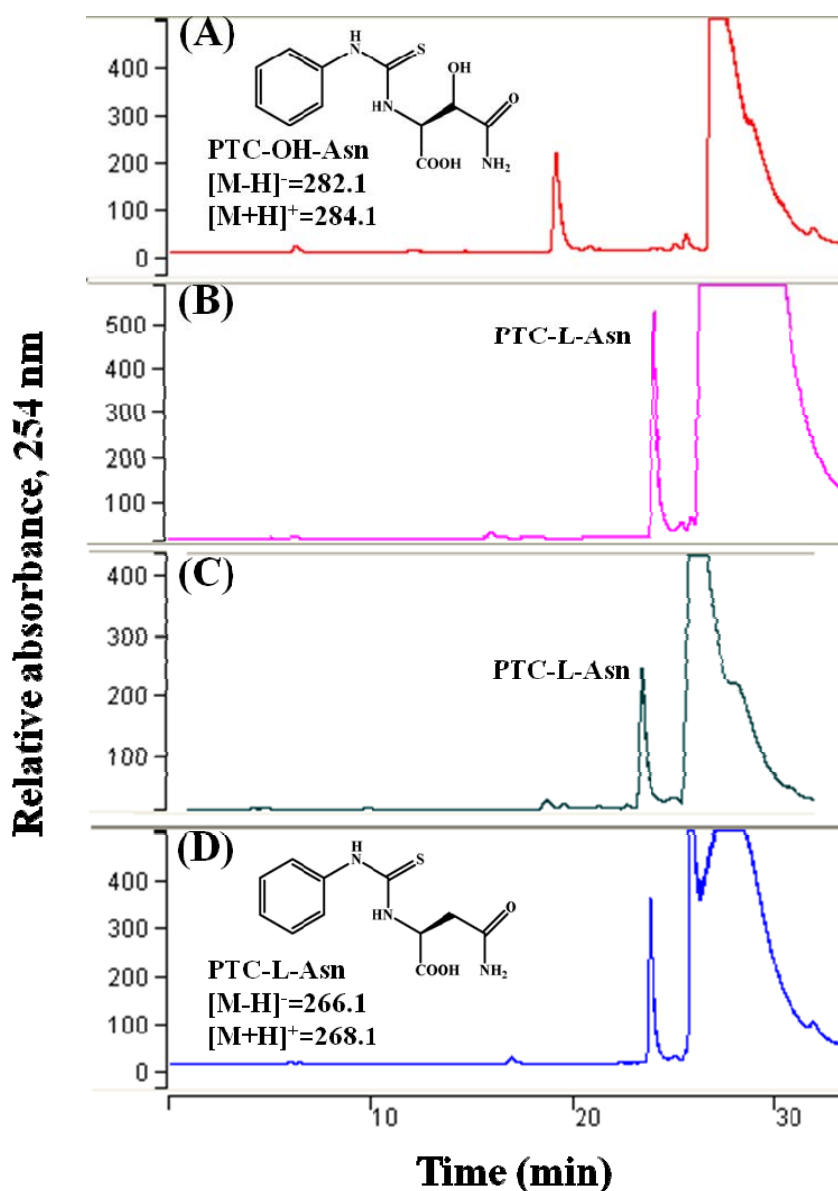


Figure 2.5. RP-HPLC chromatogram of LptL reaction with L-Asn.

(A) The product from LptL reaction with substrate L-Asn, cosubstrate α -KG and cofactor iron, (B) The product from the LptL reaction containing L-Asn and iron without α -KG, (C) The product from LptL reaction containing L-Asn and α -KG without iron, (D) The substrate L-Asn with the denatured LptL. Amino acid after LptL reaction was derivatised with PITC.

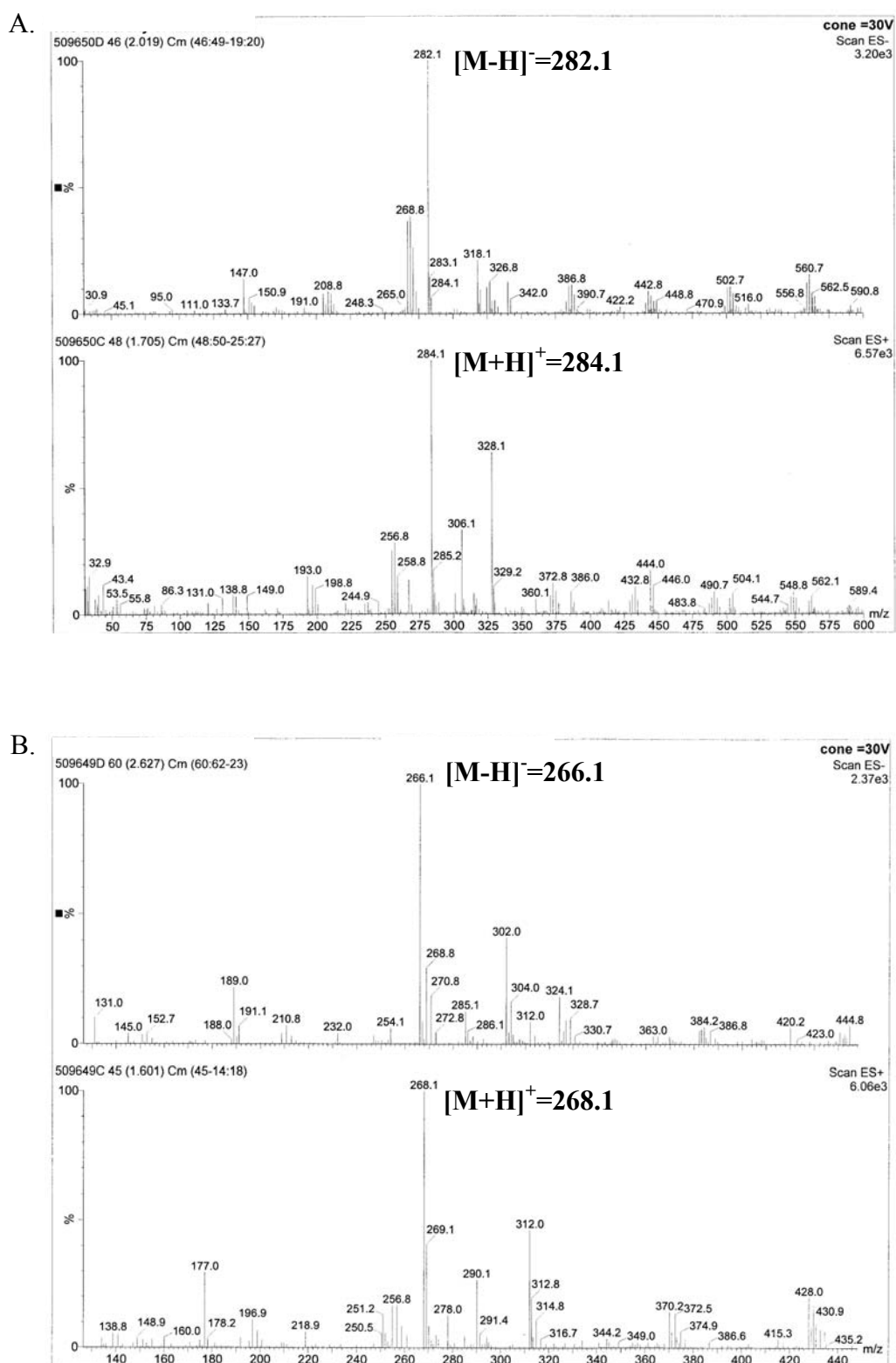


Figure 2.6. Mass spectra of the products from the LptL enzyme assays.

(A) The negative and positive mass spectra for the peak with the retention time of approximately 19 min (see Figure 2.5A). (B) The negative and positive mass spectra for the peak with the retention time of approximately 24 min (see Figure 2.5D).

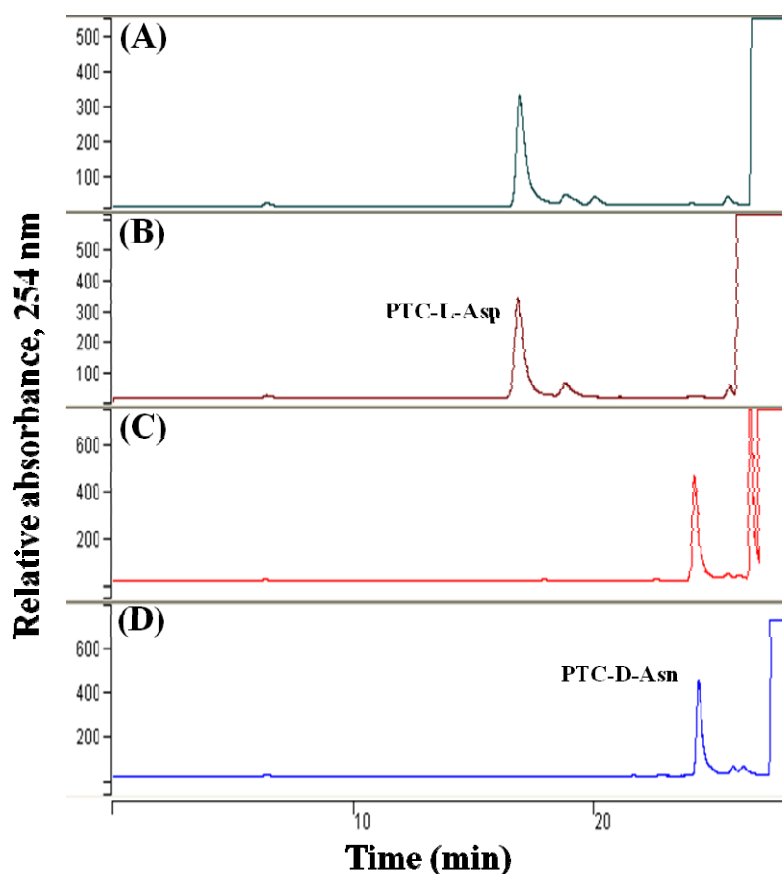


Figure 2.7. RP-HPLC chromatogram of LptL incubation with L-Asp and D-Asn. (A) The product from LptL reaction with L-Asp, (B) The substrate L-Asp with the denatured LptL, (C) The product from LptL reaction with D-Asn, (D) The substrate D-Asn with the denatured LptL. Amino acids after incubation with LptL were derivatised with PITC. Retention times for reactions (A) and (C) were identical to that of the negative controls (B) and (D), indicating that these substrates are not modified by LptL.

2.2.4. Modification timing for formation of OmAsp

The formation of the non-proteinogenic amino acid OmAsp may occur either before or during peptide assembly. For the former, the OmAsp-activating A domain would need to select substrates from a pool of L-Asp and L-modified-Asp, as A54145 antibiotic contains both L-Asp and L-OmAsp amino acids in its backbone. A comparison of substrate recognition sequences of A domains corresponding to L-Asp and L-OmAsp should reveal some variance in the substrate binding pocket residues if modification does indeed occur on free amino acids, that is, the A domain that activates L-OmAsp must have a selective substrate binding motif, distinct to that of Asp-activating A domain. Conversely, if modification occurs during peptide assembly,

then we should expect the substrate specificity motifs to be invariant for all A domains selective for L-Asp, as modification then occurs while L-Asp is bound to the PCP domain.

In the analysis of the substrate binding pocket residues of the corresponding OmAsp- and Asp-activating A domains from the A54145, CDA and daptomycin synthetases (Table 2.3), we found that the active site of the OmAsp-activating A domain of module LptC2 is identical to that of the Asp-activating A domains of modules Cda1-4, Cda1-5 and Cda2-1. Additionally the active site for OmAsp is similar to that of the Asp-activating A domains of modules LptB2, DptA3, DptB2 and DptC1, differing by only one residue. This information indicates the A domain of the OmAsp module recognises and binds L-Asp rather than L-OmAsp, and suggests that the timing of the hydroxylation modification occurs when L-Asp is tethered to the PCP domain. We therefore hypothesised that PCP-Asp is first hydroxylated by LptJ, then methylated by LptK to form OmAsp (Figure 2.8).

Table 2.3. The substrate-binding pocket residues for the Asp- and OmAsp-activating A domains

Activated amino acid	Module	A domain substrate-binding pocket residues
OmAsp	LptC2	DLTK I GAVN
	LptB2	DLTK V GAVN
	DptA3	DLTK L GAVN
	DptB2	DLTK L GAVN
Asp	DptC1	DLTK L GAVN
	Cda1-4	DLTK I GAVN
	Cda1-5	DLTK I GAVN
	Cda2-1	DLTK I GAVN

Note: Lpt, Dpt and Cda represent the A54145, daptomycin and CDA synthetases respectively. The substrate binding pocket residues of each A domain are obtained by the alignment of the A4-5 motifs with that of GrsA. Variant residue is highlighted in red.

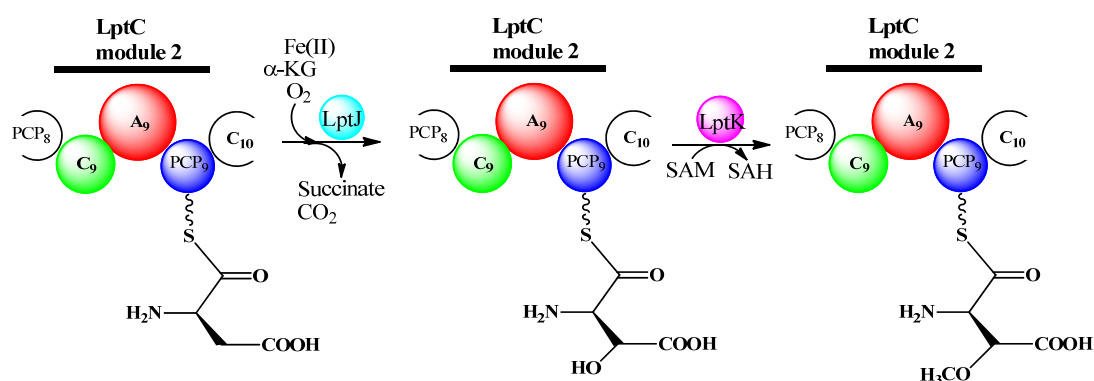


Figure 2.8. Proposed biosynthetic pathway for OmAsp.

2.2.5. Summary

The protein LptL was successfully produced in an active soluble form and was shown to hydroxylate L-Asn but not D-Asn and L-Asp, demonstrating its substrate stereoselectivity. The 3D-homology structure of LptL shows that its substrate binding site is identical to that of AsnO, which has been characterised by X-ray crystallography and produces L-(2S, 3R)-*trans*-OH-Asn [95, 111]. Moreover given that high degree similarity of protein sequences between AsnO and LptL, we predicted that the product of LptL is also L-(2S, 3R)-*trans*-OH-Asn. Regarding the biosynthesis of OmAsp, we proposed that OmAsp formation occurs during peptide assembly line (Figure 2.8) according to the comparison of the substrate binding pocket residues of the corresponding OmAsp and Asp-activating A domains (Table 2.3).

2.3. LptJ: SyrP-like protein/Hydroxylase

2.3.1. Bioinformatic Analysis of LptJ

The gene downstream of *lptL*, *lptJ*, shares homology not only to SyrP protein but also to a known non-heme Fe(II)-dependent hydroxylase, TauD (Table 2.4). The SyrP protein is usually a regulatory protein for antibiotic production, but SyrP from the syringomycin biosynthetic gene cluster in *P. syringae* has been shown to be a hydroxylase, hydroxylating a PCP-tethered Asp on the 8th module of the syringomycin synthetase [105]. From sequence alignment, the conserved “2-His-1-carboxylate facial triad” motif for iron binding, which is a characteristic for hydroxylases, is present in the LptJ amino acid sequence (Figure 2.9). Based on the case study of syringomycin and the LptJ alignments, we predicted that LptJ might be the hydroxylase which hydroxylates L-Asp during peptide assembly. Here we examined whether LptJ can hydroxylate L-Asp before or after peptide assembly, by measuring its activity with both the free amino acid L-Asp and with the two antibiotics CDA and daptomycin (the antibiotics A54145, CDA and daptomycin all contain the conserved L-Asp residue; see section 1.8).

Table 2.4. Identification of LptJ function by a Blast protein sequence search

Protein name	Description	Accession number	Identity (%)
SyrP-like protein	<i>Streptomyces avermitilis</i> MA-4680	NP_824815	62
SyrP-like protein	<i>Streptomyces verticillus</i>	AAG02342	42
TauD dioxygenase	<i>Cyanothece sp.</i> PCC 8801	NC_011726.1	34

	BAD	AVG	GOOD			
LptJ	1	mep	ta	-----	wrpa---vispdsha-lpatadalag	28
SaSyrP	1	mt	dtqt	-----	wtpl---eieteceggaeelvervag	29
SvSyrP	1	mt	gsvt	-----	ltplggiiiprprgeg-lttgaeydlg	31
TauD	1	mi	skplktikrravniaaasqlvtvscfeqkpi	-----	iiiiqpnm-ldlia-----	47
	29	llg	-----	dsartdellaahkvflsgfgvgp	---lelekimp111pdr1	70
	30	m	-----	geefgkllvaekglvfrfgvtp	---edldraldallpnrl	68
	32	plgdagpdwvrahgprlrerlatdglillhglptdgdgvdgfhdvvgsvggdpl				85
	48	-----	watyhqe	vinnylqqggailfrqfsink	---laqfeelmtalfgsll	91
				▼ ▼		
	71	pyvfgnsprtkvghnvytsteypaeftismhsemsyaarwparllfyceraadt				124
	69	ayvhgnsprtkvgsnvytsteypreftismhnemsyahawpsrlafycqvpggg				122
	86	pyterstprsvvkgniytsteypadgppimhnemsyahawpstlyffchtapdt				139
	92	dysyqstprhkvkgsiytsteyppeqfip1hnemsyasnwpekiqffclkaatq				145
	125	ggatpvdnaawyralkdvrdaya-	gglrytqnlhggrglgkswgdtfetdr			177
	123	ggatpvvdaavwygsldaevreafa-	ggvryvqnlhdgyglgkswgdtfettsr			175
	140	ggatpiadgravldlipaevrarfs-	ggvvytrtfrad--mg1swgeafqtedr			190
	146	ggetpiansrriqgridpkirekfqekgilyvrnyseq--ldlpwqkvfgttnk				197
				▼		
	178	seveeysrtgatwgwnarnglrsvshvrpatiehpatgerlwfngsdgwhpatl				231
	176	eeveaflgptgatwewkadggirvssvrpattrhpvtgaevwfngsdgwhpagl				229
	191	gdverhcr ahggefswdg- dvlrtrhhrpatavdpgtgaevwfngahl fhps1				243
	198	lqvenycr qsgiewewnd- nhlktrqicqavanhpqt nemvwfngahl fhvss1				250
	232	gge-aaalmellppeelpqsvafadgspipaeyarqvrdrglehavdnndwrpgd				284
	230	gddtaaalagilpedelpqsvtfadgspipaeyvaqirdrglanavdvdragd				283
	244	dpdlrqvllletygenlprdal fadgtpipdadlatvraaytraalalpwrsgd				297
	251	ns sfrds1levlkeedlprnayygdgtplevsvleeirtiyqeemvifswqsgd				304
	285	lmlvdnvqaahgrrpftgdrilvamsdhgrphrpgqpprppqeghr				331
	284	l11idnvl1ahgrrpfvgdrvlvams-----d				311
	298	imlvdn1rmahgrepftgerrvlvants-----ads				328
	305	l111dnmltahgrmpftgerrv1vamaqphd-----lvvktwt1i				345

Figure 2.9. The LptJ amino acid sequence alignment with SaSyrP, SvSyrP and TauD.

The amino acid sequence of LptJ is aligned with that of SaSyrP (from *S. avermitilis* MA-4680), SvSyrP (from *S. verticillus*) and TauD (from *Cyanothece sp.* PCC 8801) using the online alignment program T-coffee [201]. (▼) shows the three conserved residues for iron binding.

2.3.2. Heterologous expression of *lptJ* in *E. coli*

The cloning and heterologous expression of *lptJ* was carried out in a similar manner to that of the *lptL* (section 2.2.2). The *lptJ* gene of approximately 900 bp was amplified from the genomic DNA of *S. fradiae* by the primers lptJ-F1 and lptJ-R1 using PCR (Figure 2.10A) and was cloned into the pET30b expression vector for protein expression with a C-terminal His₆-tag in *E. coli* BL21 (DE3). The recombinant protein LptJ was then purified by Ni²⁺ affinity chromatography and analysed by SDS-PAGE (Figure 2.10B). The result showed a protein band of approximately 36.7 kDa, which is consistent with the theoretical molecular weight calculated from the LptJ protein sequence analysis. The protein yield was 0.25 mg from a 500 ml culture.

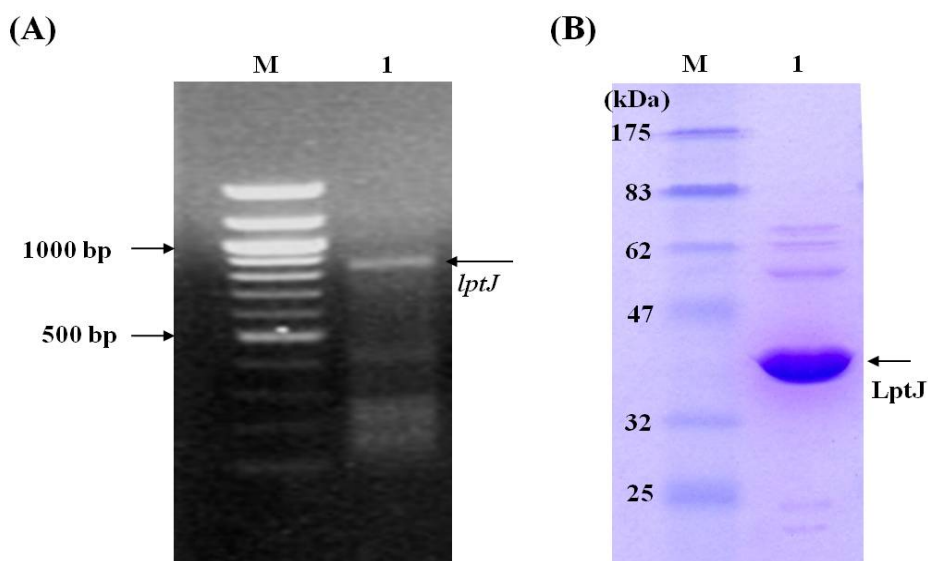


Figure 2.10. The agarose and polyarylamide gel electrophoretic analyses of the gene *lptJ* and protein LptJ.

(A) The DNA electrophoretic analysis of the gene *lptJ* amplified by PCR. Lane M: DNA size marker. Lane 1: PCR products. The arrow indicates the *lptJ* DNA fragment band (approximately 900 bp). (B) The SDS-PAGE analysis of the purified LptJ. Lane M: Protein molecular weight marker. Lane 1: LptJ eluted from Ni²⁺ affinity column in 250 mM imidazole. The arrow indicates LptJ band (approximately 36 kDa).

2.3.3. Hydroxylase Activity Assay

The free amino acid L-Asp and the antibiotics daptomycin and CDA (a mixture of CDA3a and 4a) were used to assay for LptJ activity, although we proposed that the complex of L-Asp tethered to PCP domain of OmAsp module is the substrate for LptJ. The assay conditions were as for LptL and are given in section 6.26.1. As shown in Figure 2.11, the RP-HPLC retention times of the derivatised substrate L-Asp and the L-Asp product post-incubation with LptJ were the same, indicating that no modification had occurred on L-Asp. Additionally, daptomycin and CDA were purified after from the LptJ reaction by RP-HPLC and applied to liquid chromatography-mass spectrometric (LC-MS) for analysis. The daptomycin mass spectra (Figure 2.12) showed no difference in mass between the daptomycin from the LptJ reaction products and the control. Although a product with a mass of 1636.4 Da can be observed with an increase of 16 Da over wild-type, this product was also observed in the control sample at a similar ratio to wild-type. Therefore we presumed that this was not due to LptJ activity. In the CDA mass spectra (Figure 2.13), CDA3a and CDA4a can be identified. However we did not find any new CDA products after LptJ reaction.

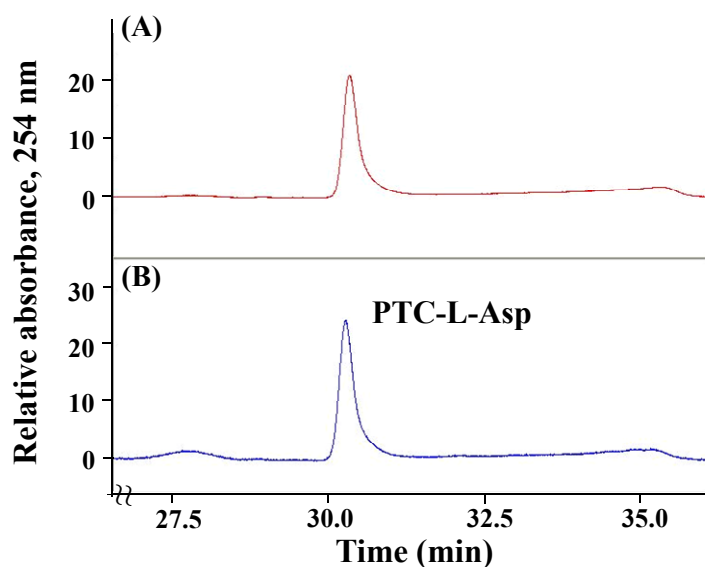


Figure 2.11. RP-HPLC spectra of LptJ reaction products.

(A) The product from incubation of L-Asp with LptJ, followed by PITC derivatisation, (B) The substrate L-Asp after PITC derivatisation.

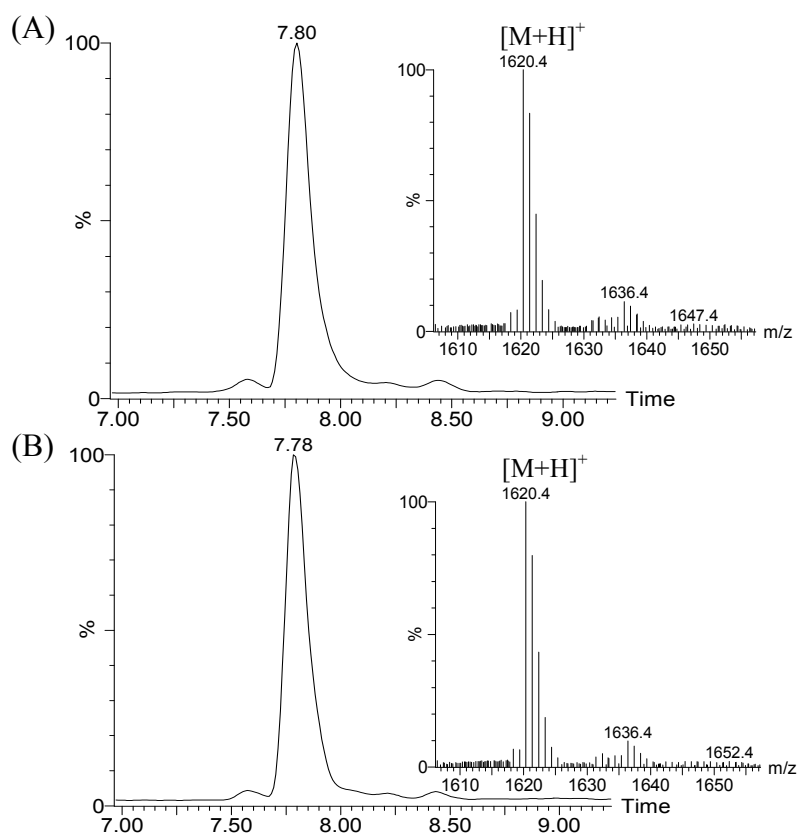


Figure 2.12. LC-MS analysis of daptomycin from LptJ reaction.

(A) The product from the incubation of daptomycin with LptJ. Mass spectra of the RP-HPLC peak gave an m/z value of 1620.4, which relates to the $[M+H]^+$ ion of the daptomycin standard. (B) Control: daptomycin standard.

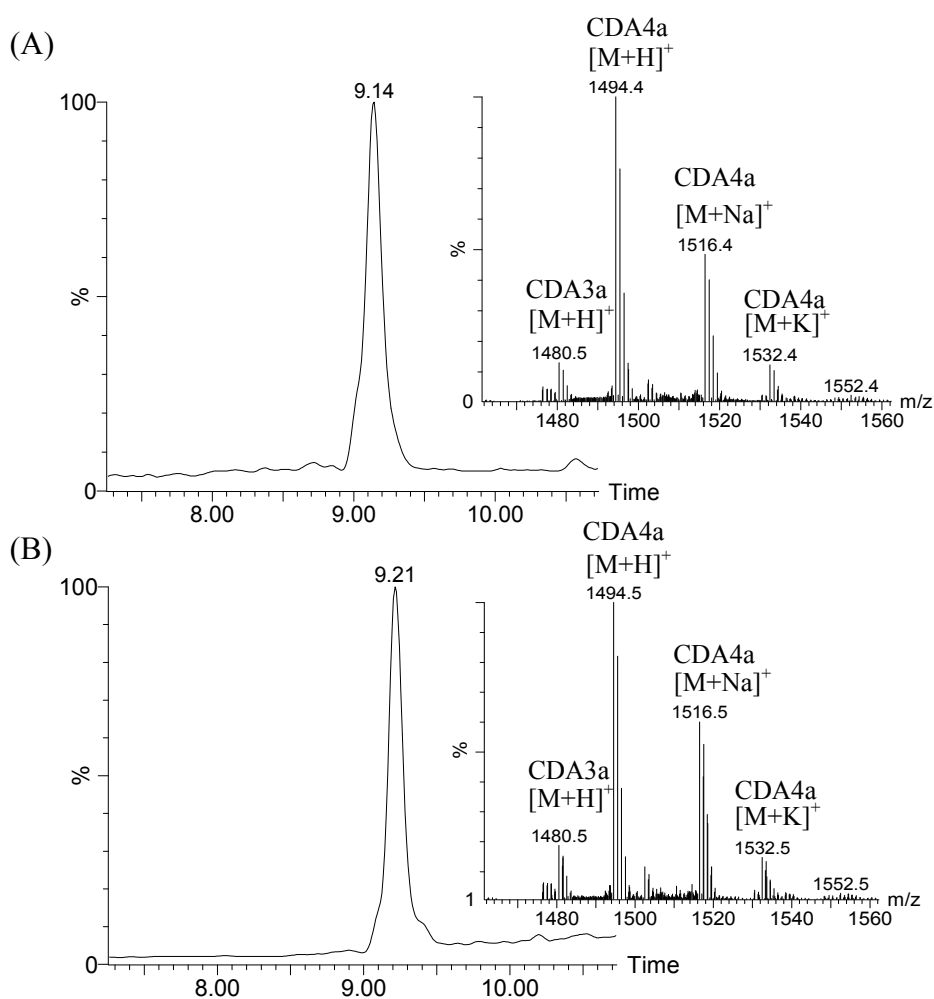


Figure 2.13. LC-MS analysis of CDA mixture from LptJ reaction.

(A) The products from the incubation of CDA mixture with LptJ. Mass spectra of the chromatogram peaks at 8.9-9.4 mins gave an m/z value of 1480.5 and 1494.4, which corresponds to the $[M+H]^+$ ion of the CDA3a and CDA4a standards respectively. (B) Control: CDA mixture standards.

2.3.4. Summary

LptJ, which we predicted to be a hydroxylase, was successfully cloned and over-expressed in soluble form. LptJ activity was then assayed with free amino acid L-Asp and the two antibiotics, CDA and daptomycin, however no hydroxylation occurred. The results of these assays support the hypothesis that LptJ works during peptide assembly (Figure 2.8). In section 2.5, we will examine the function of LptJ using an *in vivo* strategy.

2.4. LptK: O-methyl transferase (O-MTase)

2.4.1. Bioinformatic Analysis of LptK

LptK was assigned to the O-MTase family in accordance with its amino acid sequence homology with the known O-MTases RapI and FkbM, involved in biosynthesis of the antibiotic rapamycin and the immunosuppressant FK506 respectively [206, 207] (Table 2.5). Therefore we predicted that LptK could methylate OH-Asp to produce the non-standard amino acid OmAsp which can be found within the lipopeptide A54145 [2].

Table 2.5. Identification of LptK function by a Blast protein sequence search

Protein name	Description	Accession number	Identity (%)
Putative methyl transferase	<i>Streptomyces griseus</i> NBRC 13350	YP_001822082.1	45
RapI (O-methyl transferase)	<i>S. hygroscopicus</i> NRRL5491; involved in rapamycin biosynthesis	CAA60470.1	40
FkbM (O-methyl transferase)	<i>Streptomyces sp.</i> MA6858; involved in immunosuppressant FK506 and FK520 biosynthesis	AAC44360.1	40

Additional evidence which supports the role of LptK as an MTase is the presence of a conserved motif found within known MTases. In 1994 Kagan and Clarke aligned 84 MTase (non-DNA methylase) sequences, includes protein carboxyl MTases, O-MTases, N-MTases, S-MTases, porphyrin precursor MTases, lipid MTases and RNA MTases, from diverse organisms ranging from bacteria to humans [208]. Three conserved motifs (I, II, III) were identified and proposed as SAM-binding elements (Table 2.6). However the distribution of the conserved motifs throughout the range of identified MTases is diverse. For example, in the study of the 84 MTase sequences, the individual motifs I, II or III can be found in 69, 46 and 61 of the 84 sequences respectively, while all three motifs are present in 45 of the enzymes. In the LptK amino acid alignment with RapI and FkbM (Figure 2.14), motif I can be found in all

three *O*-MTase sequences.

Table 2.6. The conserved motifs in MTases

Motif I	(V/I/L) (L/V) (D/E) (V/I) G (G/C) G (T/P) G
Motif II	(P/G) (Q/T) (F/Y/A) D A (I/V/Y) (F/I) (C/V/L)
Motif III	LL (R/K) PGG (R/I/L) (L/I) (L/F/I/V) (I/L)

	BAD AVG GOOD	Motif I
LptK	1 MTIALAD-----VEGLNQHETEFlyDE IFTRRAYLPEALHLPEAPVVFVDVG	46
FkbM	1 MSDVVE TLRLPNGATVAHVNAGEAQFLYRE IFTRCYLRHGVELRPGDVVFVDVG	54
RapI	1 MSASVQTIKLPYGRPSAHVNPGEAQFLYQE IFAERCYLRRGLELRAGDVVFVDVG	54
	Motif I	
	47 ANIGMFTLFVRSERPGATVHSFEPVPPVRDI LCRNRERHAVAGLVHPYGLAEAE	100
	55 ANIGMFMLFAHLEHPGVTVHAFEPAPVPFAALRANAVRHRVAGRVDQCAVSDEA	108
	55 ANIGMFSLFAHLECPDVTVHAFEPAPVPAALRANAERYAIAGRFEQCAVSDVA	108
	101 QEVFTHYPGYSTMSTRSTLADTEAERAFVRGQVRTADLPEAERMLDELLAFRF	154
	109 GVRRMTFFYPDATLMSGFHPDAAARKE LLRTLGLNGGYTAEDVDMMLAQLPDT--	160
	109 GRGKMTFFYTDTTMSGFHPDPATRAELLRRLAINGGYSAAADRMLAELPDT--	160
	155 REEKVTCRLRPLSAVLDEHPVDRIDLLKIDVQGEREVLRGLEDHRHWPLVVRQIA	208
	161 -GEEIETSVVRLSDVIAERGI AAI GLLKIDVEKSERRVLAGVEDADWPRI RQVV	213
	161 -SQVIETSVVRLSDVIAERGIT S IGLLKIDVEKNERHVMAGIDAGDWPRI RQVV	213
	209 MEVHDSPPGGSTAGRLRAVADELERRGFDVLTEQEDRYAGTDRHSVFAVAEPRRG	262
	214 AEVHDVD-----GALGEVVALLRGMGFTVVAEQDPLFAGTEIHQVA--ARRTAG	260
	214 TEVHDID-----GRLDEVLTLLRGQGFVLSQEPLFAGTDIYQVV--ARRGDA	260

Figure 2.14. The LptK amino acid sequence alignment with FkbM and RapI.

The amino acid sequence of LptK is aligned with that of the FkbM and RapI from *S. hygroscopicus* NRRL 5491 and *Streptomyces sp.* MA6858 respectively, using the online alignment program T-coffee [201].

2.4.2. Heterologous Expression of *lptK* in *E. coli*

The cloning and heterologous expression of *lptK* was carried out in a similar manner to that of *lptL* (section 2.2.2). The *lptK* gene of approximately 800 bp was amplified with the primers lptK-F1 and lptK-R1 from the genomic DNA of *S. fradiae* by PCR (Figure 2.15A). The gene fragment was digested by *Nde*I and *Xho*I, and then cloned into the pET30b expression vector for protein expression with a C-terminal His₆-tag in *E. coli* BL21 (DE3). After induction with IPTG, the recombinant LptK protein was expressed and then purified by Ni²⁺ affinity chromatography. As shown in Figure 2.15B, one protein band of approximately 30 kDa can be observed by SDS-PAGE, consistent with the theoretical molecular weight calculated from the LptK protein sequence analysis. The protein yield was 0.10 mg from a 500 ml culture.

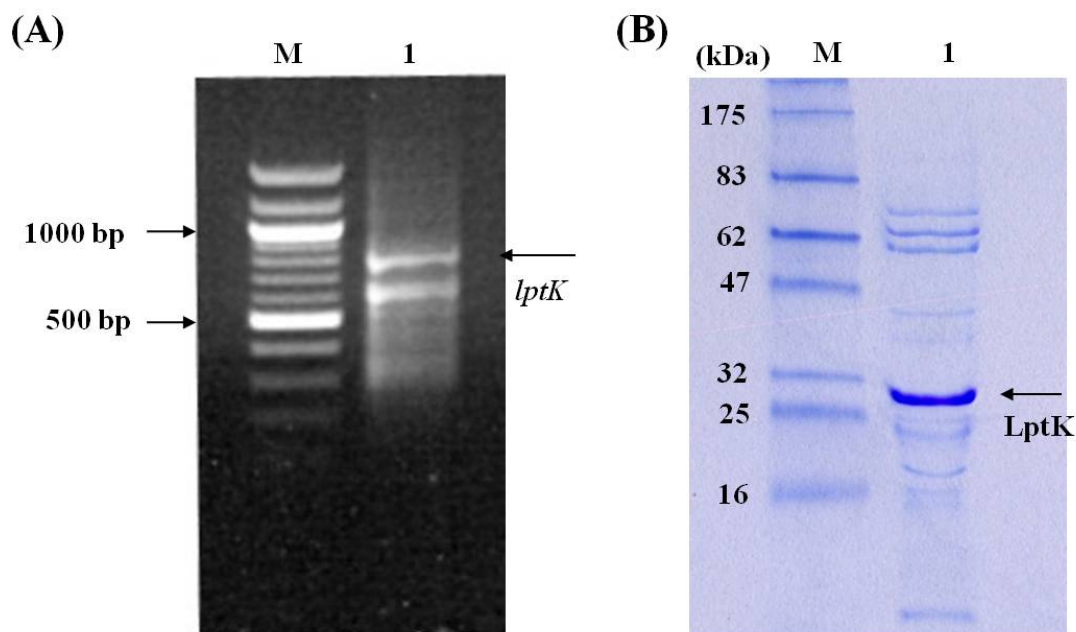


Figure 2.15. The agarose and polyarylamide gel electrophoretic analyses of the gene *lptK* and protein LptK.

(A) The DNA electrophoretic analysis of the gene *lptK* amplified by PCR. Lane M: DNA size marker. Lane 1: PCR products. The arrow indicates the *lptK* DNA fragment band (approximately 800 bp). (B) The SDS-PAGE analysis of the purified LptK. Lane M: Protein molecular weight marker. Lane 1: LptK eluted from Ni²⁺ affinity column in 250 mM imidazole. The arrow indicates the LptK band (approximately 30 kDa).

2.4.3. O-MTase Activity Assay

The commercially bought D, L-threo-OH-Asp was used to assay for LptK activity although we proposed that the modifications of Asp may occur during peptide assembly. The enzyme assay was conducted in HEPES buffer pH 7.5 with amino acid substrate, SAM, MgSO₄ and enzyme at 28 °C for 18 h. The assay conditions, derivatisation method and detection methods are given in section 6.25 and 6.26.2. After substrate incubation with LptK, products were derivatised with PITC and analysed by RP-HPLC. From analysis of HPLC chromatogram, starting materials and products showed no difference in retention times (Figure 2.16). The mass spectrometric analysis also showed that the starting materials and products have the same molecular weights (data not shown). This indicates that LptK was unable to methylate the free OH-Asp.

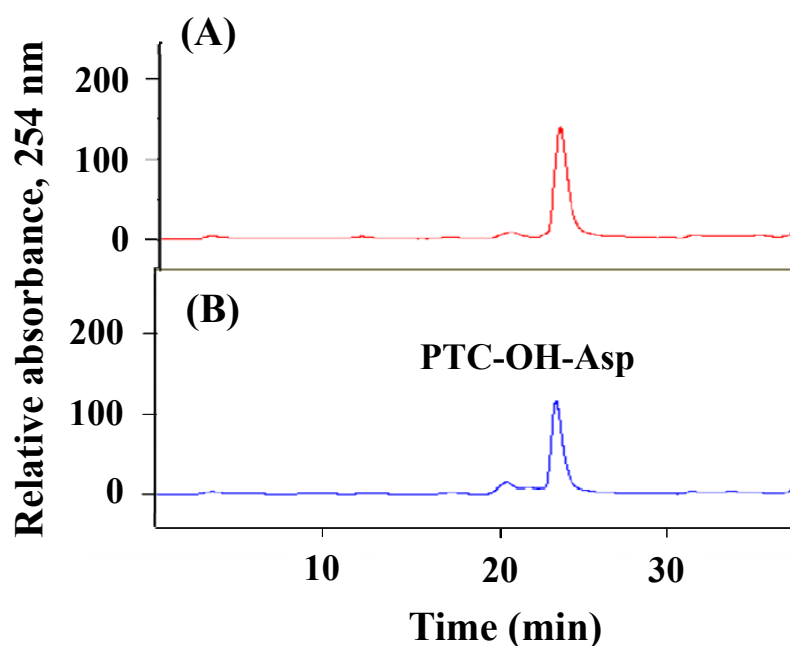


Figure 2.16. RP-HPLC spectra of LptK reaction products.

(A) The product from incubation of OH-Asp with LptK, followed by PITC derivatisation, (B) The substrate OH-Asp after PITC derivatisation.

2.4.4. Summary

The LptK protein was successfully over-expressed in soluble form. However the MTase assay showed that LptK did not methylate OH-Asp. This result is consistent with the assumption that methylation of substrate amino acids occurs during peptide assembly. Therefore we adopted an *in vivo* experimental strategy to examine LptK function and activity (the details of which are given in the next section).

2.5. In Vivo Studies of the genes *lptJ* and *lptK*

The antibiotic CDA shares a similar structure with the antibiotic A54145, containing an L-Asp residue at the same relative position as the OmAsp in A54145. We expected that the biosynthetic modification system for OmAsp biosynthesis from the A54145 biosynthetic gene cluster could be used to modify the L-Asp residue of CDA. Therefore we chose the *S. coelicolor* MT1110 strain, which produces CDA, as a target organism for this *in vivo* experiment. We employed two *Streptomyces* protein expression vectors, the integrative pMT3226 vector containing a glycerol inducible promoter (*glyP*) (Section 2.5.1) and the non-integrative pIJ86 vector containing a constitutive promoter *ermE** (Section 2.5.2), to set up a single gene *lptJ* construct and a two-gene *lptJ-lptK* construct. The pMT3226 vector integrates into the chromosome as a single copy, with the gene insert(s) expressed by the addition of 1 % glycerol. The high copy number non-integrative pIJ86 vector replicates autonomously in the cytoplasm, with the cloned insert(s) constitutively expressed due to the *ermE** promoter. Four constructs were cloned: pMT3226-*lptJ*, pMT3226-*lptJK*, pIJ86-*lptJ* and pIJ86-*lptJK*. In the A54145 biosynthetic gene cluster, the genes *lptJ* and *lptK* are potentially translationally-coupled due to their bi-cistronic nature; the stop codon of *lptJ* (TGA) overlaps with the start codon of *lptK* (ATG) to form the sequence ATGA, so we cloned *lptJ* and *lptK* together as a continuous sequence. We expected that the expression in *S. coelicolor* of the single enzyme LptJ, or both enzymes LptJ and LptK, would result in the modification of Asp that is tethered to the PCP domain of CDA synthetase, leading to the production of CDA with OH-Asp and OmAsp residues respectively.

2.5.1. Glycerol Inducible Vector pMT3226

2.5.1.1. Verification of pMT3226 Functionality in *S. coelicolor*

The pMT3226 vector is transferable from *E. coli* to *Streptomyces spp.* (the map of pMT3226 is given in section 6.2.1). It can be replicated in *E. coli* due to the presence of a pUC18 origin and can be transferred from *E. coli* (ET12567/pUZ8002) to *Streptomyces spp.* by conjugation via its *oriT* element (the details of the conjugation procedure are given in section 6.15). Upon entry into *Streptomyces spp.*, the

vector-encoded PhiC31 integrase is expressed, which catalyses the integration of the vector into the PhiC31 *attB* site of the *Streptomyces spp.* chromosome. The gene cloned into the *glyP* cassette of pMT3226 may then be expressed by induction with 1 % glycerol [209]. Before investigating the expression of our target genes, we verified the vector's functionality by expression of the reporter gene *xylE*.

The reporter gene *xylE*, encodes the enzyme catechol dioxygenase which converts the colourless substrate catechol to the yellow product muconic acid [210] (Figure 2.17). This reporter acts as an indicator of the expression of upstream target genes due to its location downstream of the multiple cloning site and its transcriptional control by the same *glyP* promoter.

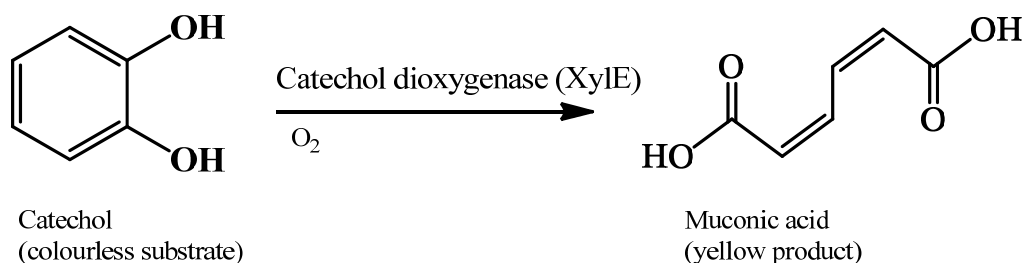


Figure 2.17. Catechol dioxygenase reaction.

Catechol dioxygenase oxidizes the colourless substrate catechol, forming the yellow product muconic acid.

S. coelicolor transformants containing plasmid pMT3226, which we named ScpMT3226, were selected by resistance to apramycin. As shown in Figure 2.18A, the wild-type *S. coelicolor* cannot grow on the antibiotic-containing Mannitol soya flour medium plate (MS plate, medium composition is given in section 6.3.2); only bacteria with pMT3226 successfully integrated can grow on selective medium. Additionally the wild-type *S. coelicolor* did not show any catechol dioxygenase activity when plates were overlaid with an aqueous solution of catechol (Figure 2.18B; assay conditions are given in section 6.16).

Eight transformant strains of ScpMT3226 grew on the MS plates containing 1 % glycerol and so these were analysed by the catechol dioxygenase assay to verify promoter activity. The result showed that four of the eight transformants displayed the yellow colour associated with the product muconic acid (Figure 2.18C). Following this, four transformant colonies (three yellow and one colourless) were used to verify

the integration of the vector sequence into the chromosome. The extracted genomic DNA from these four transformant strains was subjected to PCR analysis with the primers (glyP and xylE), which amplify the region between the glycerol promoter and the 5'-end of the reporter gene *xylE*. PCR products of approximately 500 bp were generated in all four strains, matching the expected size of this region, confirming the presence of the vector sequence in their genomes (Figure 2.18D). Therefore, since the bacteria can grow on the antibiotic-containing selective medium, their genome must be integrated with the vector pMT3226. In terms of the catechol dioxygenase assays, not all of the transformants with promoter activity show yellow colouring. The lack of colour may be issues with permeability of the substrate catechol. Therefore we still used this vector for our gene functionality studies.

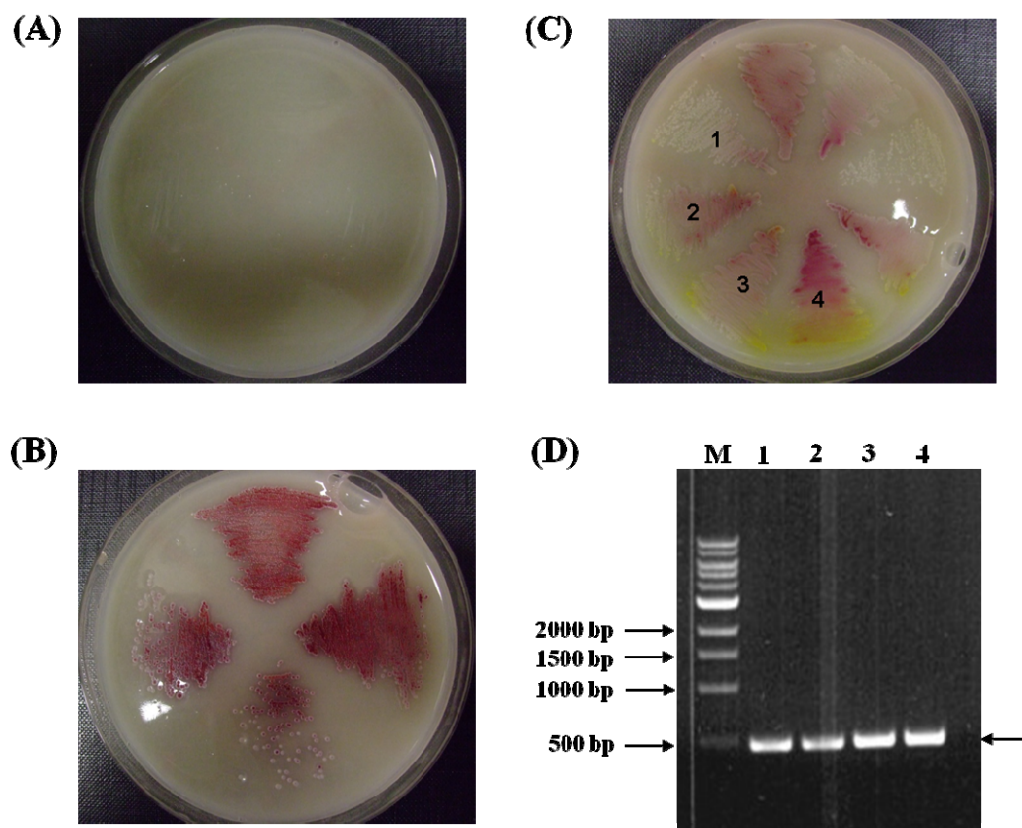


Figure 2.18. Verification of the vector pMT3226 functionality by expression of XylE.

(A) Wild-type *S. coelicolor* on an apramycin-containing plate. (B) Wild-type *S. coelicolor* colonies were assayed for catechol dioxygenase activity. (C) The transformant strains which were integrated with pMT3226 were assayed for catechol dioxygenase activity which turns colourless substrate catechol to the yellow product muconic acid (D) The PCR results from the genomic DNA of four transformant strains. Lane M: DNA size marker. Lanes 1-4: PCR products from the four transformant strains (numbered 1-4 in (C)).

2.5.1.2. Cloning of the *lptJ* and *lptJK* genes into pMT3226

The *lptJ* and *lptJK* genes were amplified by the forward primer sclptJK-F1 and the reverse primers sclptJ-R1 and sclptJK-R1 respectively, using genomic DNA from *S. fradiae* as a template. Both gene fragments contained an additional thirteen nucleotides from the sequence upstream of the *lptJ* start codon that included the Shine-Dalgarno (SD) sequence “GAAAGAG”. Fragments were cloned into the *Bam*HI site of the pMT3226 and replicated in *E. coli*, yielding pMT3226-*lptJ* and pMT3226-*lptJK* (Figure 2.19). After verification of insert by sequencing, both constructs were transformed into the *E. coli* ET12567 strain (with pUZ8002) and transferred into *S. coelicolor* via conjugation.

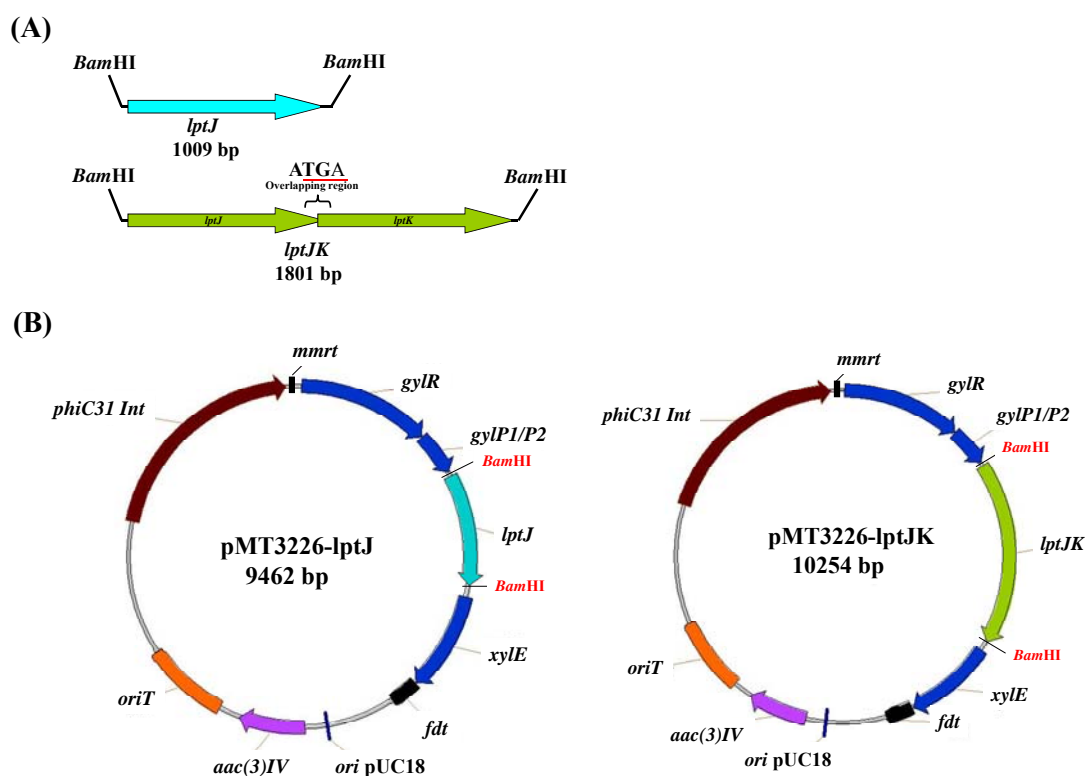


Figure 2.19. Schematic representation of the pMT3226-*lptJ* and pMT3226-*lptJK* constructs.

(A) The genes *lptJ* and *lptJK* were amplified and cloned into pMT3226 using the *Bam*HI restriction site. (B) The resulting constructs were designated pMT3226-*lptJ* and pMT3226-*lptJK*.

2.5.1.3. Characterisation of the Transformant Strains ScpMT3226-lptJ and ScpMT3226-lptJK

The pMT3226-lptJ and pMT3226-lptJK constructs were integrated into the *attB* site of the *S. coelicolor* chromosome, forming transformant strains ScpMT3226-lptJ and ScpMT3226-lptJK respectively. Catechol dioxygenase assays were carried out to verify protein expression. We did not observe any yellow colouration from the transformant strains ScpMT3226-lptJ and ScpMT3226-lptJK (Figure 2.20 A and B). The lack of yellow colour may be due to the inability of the glycerol promoter to drive tandem gene expression. Four colonies each, from both ScpMT3226-lptJ and ScpMT3226-lptJK, were taken for genomic DNA extraction and subjected to PCR analysis with the primer pair (glyP and xyle). PCR products approximately 1500 bp and 2200 bp were generated, matching the expected sizes of the *lptJ* and *lptJK* inserts respectively (Figure 2.20 C and D). Therefore the transformant strains must contain the pMT3226-lptJ and pMT3226-lptJK gene constructs in their respective genomes. We chose ScpMT3226-lptJ1 and ScpMT3226-lptJK1 for further characterisation by bioassay and LC-MS.

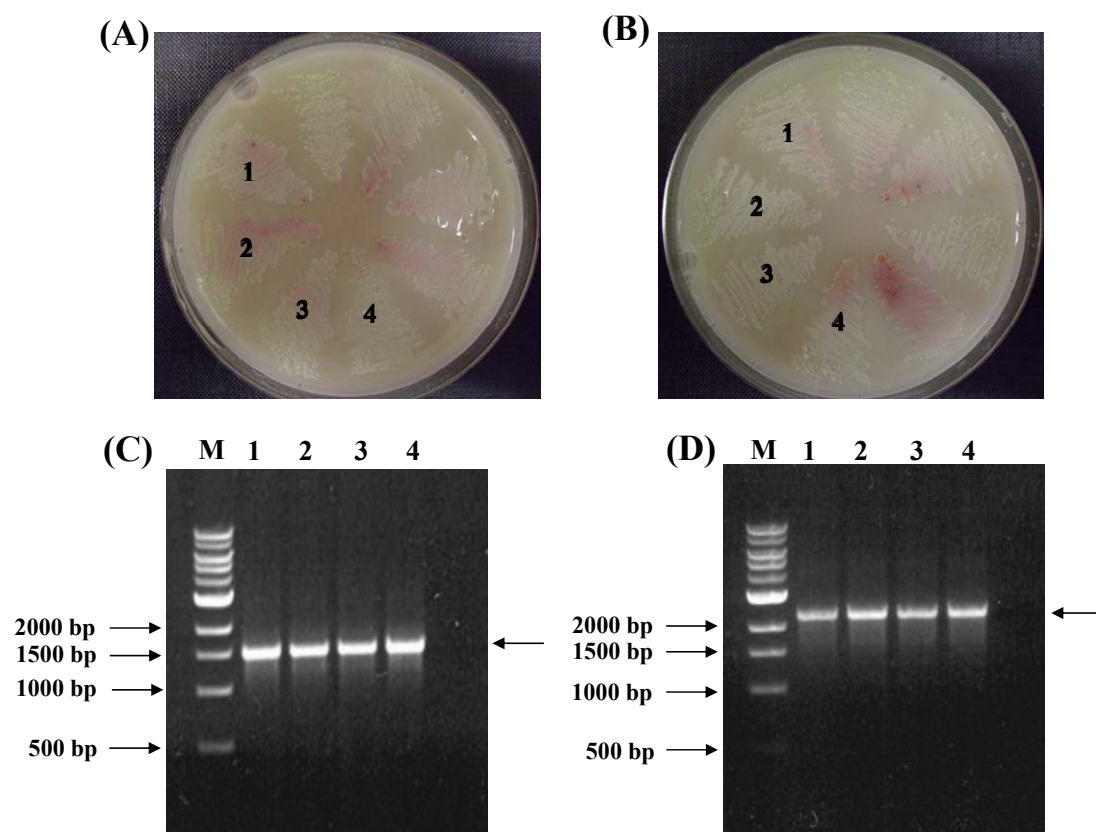


Figure 2.20. Colony screening, catechol dioxygenase assay and PCR analysis of the ScpMT3226-lptJ and ScpMT3226-lptJK strains.

The transformant strains ScpMT3226-lptJ (A) and ScpMT3226-lptJK (B) were selected on apramycin-containing plates and assayed for catechol dioxygenase activity. The PCR results from the genomic DNA of each of the four transformant strains of ScpMT3226-lptJ1-4 and ScpMT3226-lptJK1-4 are shown in (C) and (D) respectively. Lane M: DNA size marker. Lanes 1-4: PCR products from the four transformants (numbered 1-4 in A and B).

2.5.1.4. CDA Bioassay of the ScpMT3226-lptJ and ScpMT3226-lptJK strains

Micrococcus luteus, which produces a yellow pigment, was used as an indicator strain for CDA activity in this assay (the detail of the assay conditions are given in section 6.17). The presence of CDA in *M. luteus* bacterial lawns results in circular inhibition zones where cells are unable to survive due to the bactericidal effect. A negative control was set up in which no calcium was added, to ensure that the antibiotic activity was due to CDA. Both the ScpMT3226-lptJ and ScpMT3226-lptJK strains in the bioassays showed calcium-dependent antibiotic activity and the diameters of the inhibition zones showed no significant differences to that of the wild-type strain (Figure 2.21). It was concluded from this result that CDA production was unaffected in the transformant strains, and that the CDA produced was either wild-type or a variant that retaining biological activity.

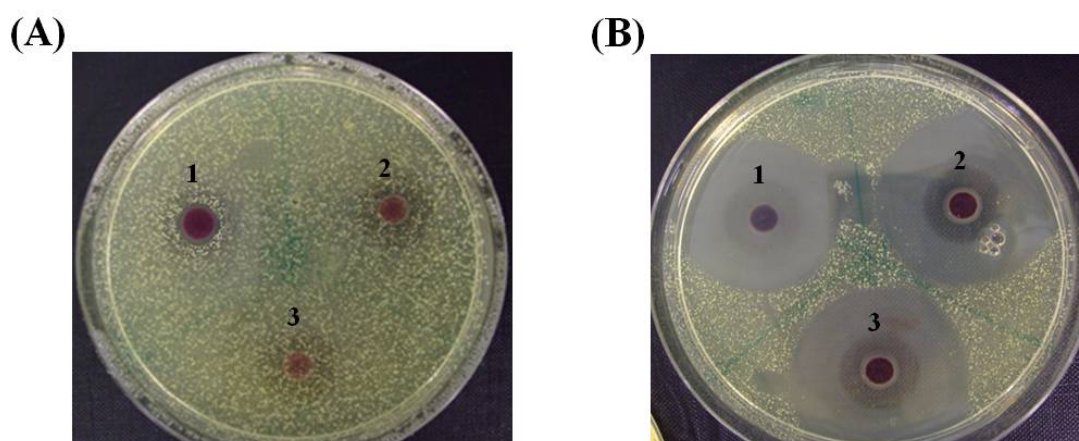


Figure 2.21. Bioassays of the ScpMT3226-lptJ and ScpMT3226-lptJK strains. Wild-type, pMT3226-lptJ1 and pMT3226-lptJK1 strains are labelled 1, 2 and 3, respectively. (A) negative control plate (without calcium). (B) bioassay plate (containing calcium).

2.5.1.5. LC-MS Analysis of the Fermentation Broth from the ScpMT3226-lptJ and ScpMT3226-lptJK strains

The spores (10^5 - 10^7) of the transformant strains ScpMT3226, ScpMT3226-lptJ and ScpMT3226-lptJK were used to inoculate SV2 media and grown at 30 °C for 7 days. The culture broths containing CDAs were collected and purified by using C₁₈ bond elut cartridges with methanol. The purification procedure was sampled at each stage and analysed by LC-MS by Laura Nunns (Micklefield group), with no new products observed. The fractions containing major amount of CDAs from the 100 % methanol eluate were presented here. The detailed procedures for purification and LC-MS characterisation are given in section 6.18 and 6.20. The chromatogram of the control strain ScpMT3226 extracts showed a major two peak (with retention time of 7.93 and 8.05) with a minor peak (with retention time of 7.43). The mass spectra presented the peak of 7.93 and 8.05 contained CDA3a ($[M+H]^+=1480.5$) and CDA4a ($[M+H]^+=1494.5$) respectively (Figure 2.22). The minor peak also contain trace amount of CDA3a and CDA4a (data not shown). The LC-MS analyses of the extracts from the transformant strains ScpMT3226-lptJ and ScpMT3226-lptJK were the same as that of the control, showing production of CDA3a and CDA4a without modification (Figure 2.23 and Figure 2.24). Interestingly the mass spectra of the ScpMT3226 and ScpMT3226-lptJ strain extracts showed a product with a 17 Da more than that of CDA3a and CDA4a respectively. It is not related to LptJ activity since the control ScpMT3226 strain also produces this product, so we did not explore this product further.

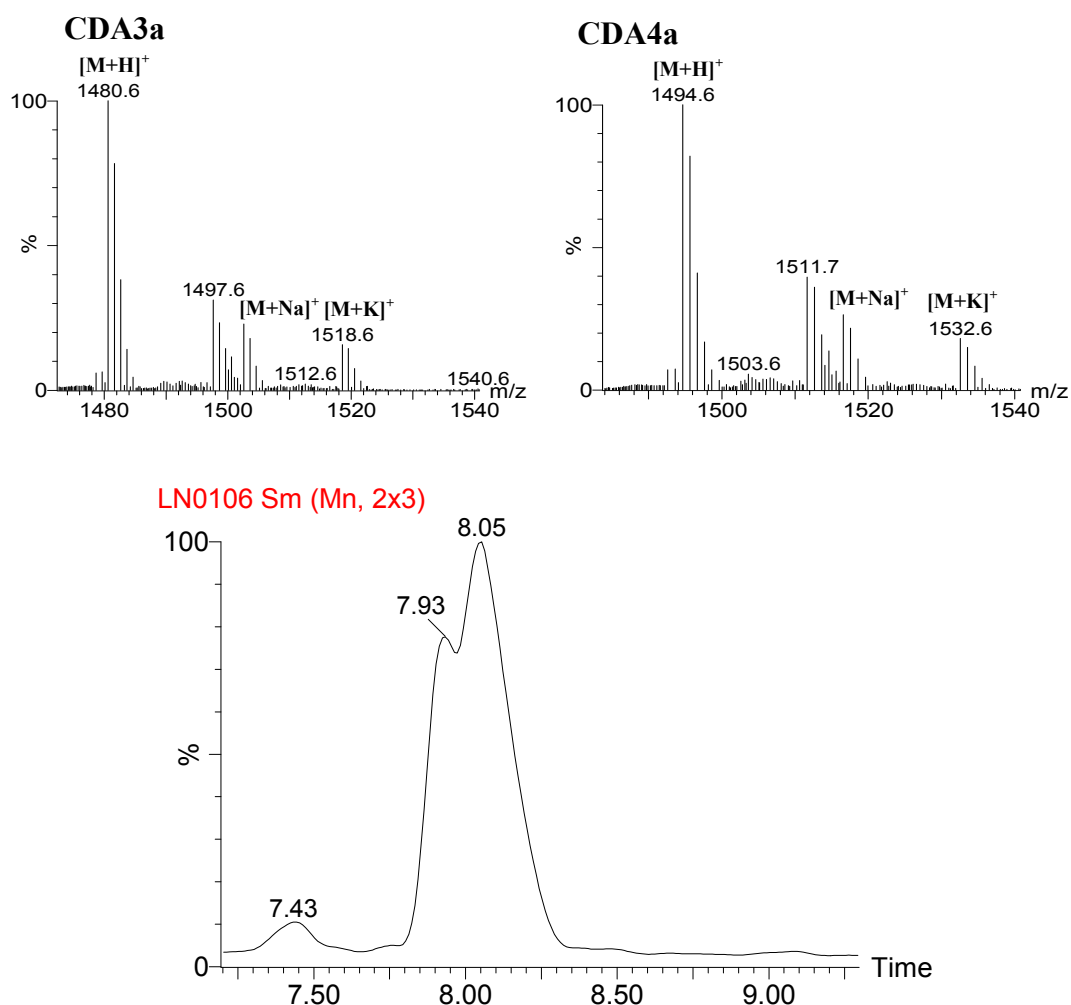


Figure 2.22. LC-MS analysis of the ScpMT3226 fermentation products.

The ScpMT3226 strain produces CDA3a ($[M+H]^+$ observed 1480.6, expected 1480.5; $[M+Na]^+$ observed 1502.6, expected 1502.5; $[M+K]^+$ observed 1518.6, expected 1518.5) and CDA4a ($[M+H]^+$ observed 1494.6, expected 1494.5; $[M+Na]^+$ observed 1516.6, expected 1516.5; $[M+K]^+$ observed 1532.6, expected 1532.5). The mass spectra of CDA3a and CDA4a were extracted from the LC peaks with retention times of 7.8-7.9 and 8.0-8.1 min respectively.

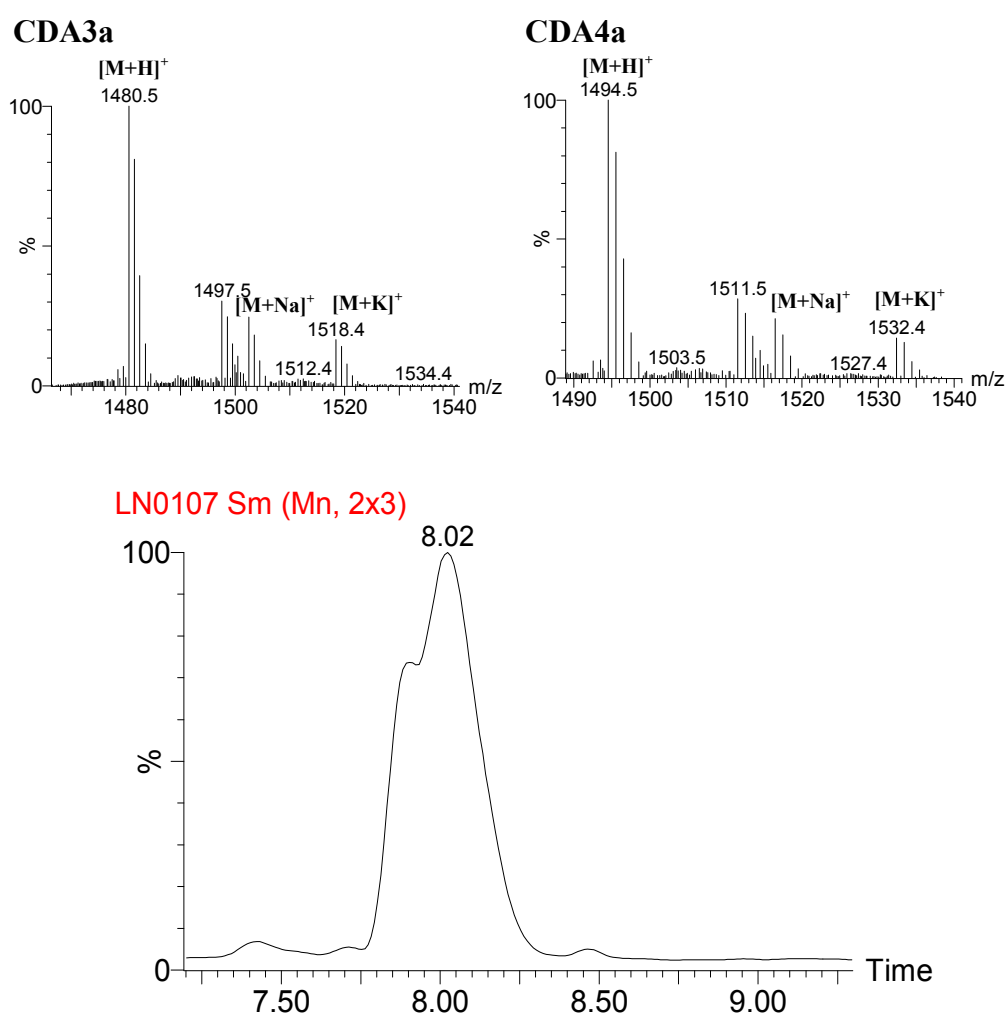


Figure 2.23. LC-MS analysis of ScpMT3226-lptJ fermentation products.

The ScpMT3226-lptJ strain produces CDA3a ($[M+H]^+$ observed 1480.5, expected 1480.5; $[M+Na]^+$ observed 1502.5, expected 1502.5; $[M+K]^+$ observed 1518.5, expected 1518.5) and CDA4a ($[M+H]^+$ observed 1494.5, expected 1494.5; $[M+Na]^+$ observed 1516.5, expected 1516.5; $[M+K]^+$ observed 1532.4, expected 1532.5). The mass spectra of CDA3a and CDA4a were extracted from the LC peaks with retention times of 7.7-7.8 and 8.0-8.1 min respectively.

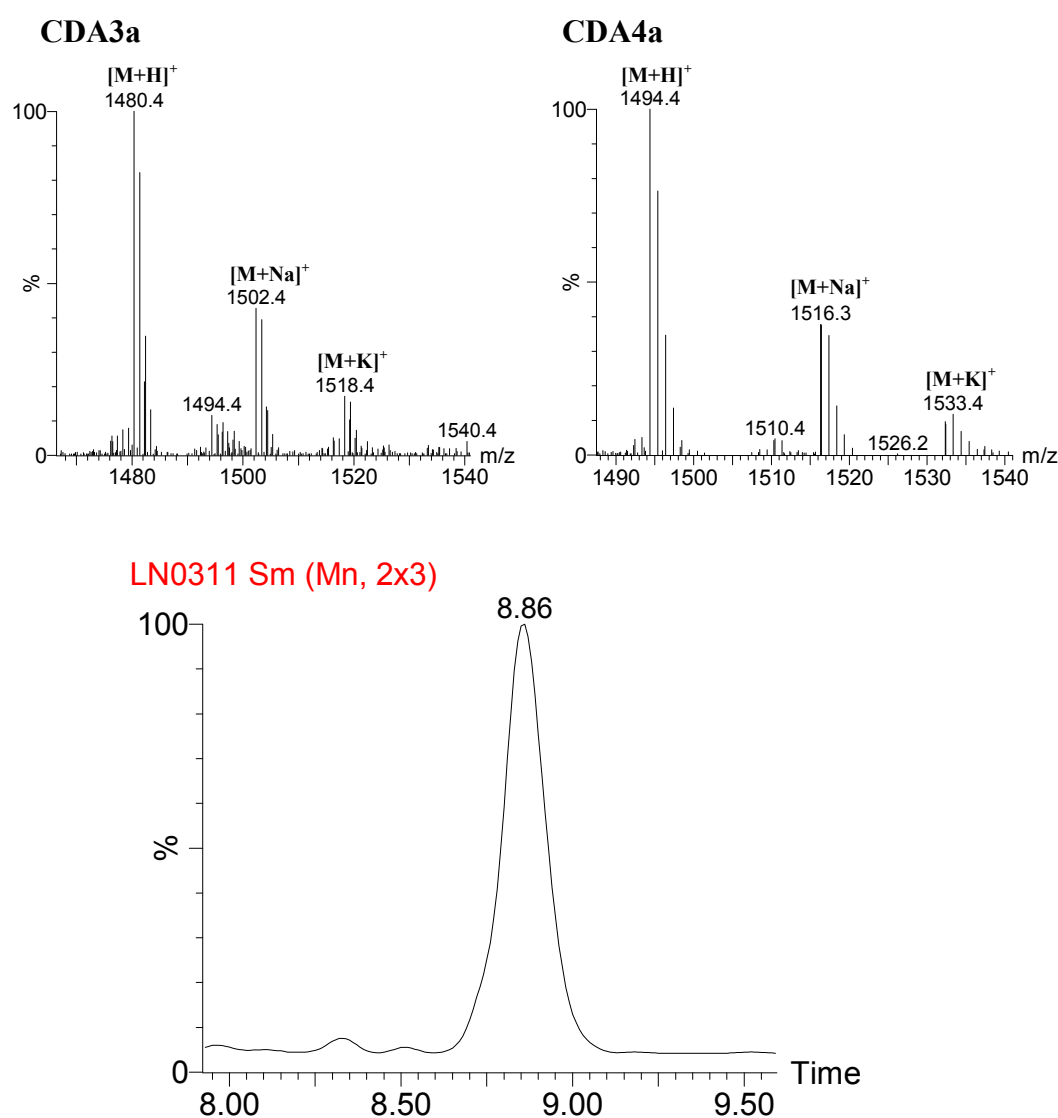


Figure 2.24. LC-MS analysis of ScpMT3226-lptJK fermentation products.

The ScpMT3226-lptJK strain produces CDA3a ($[M+H]^+$ observed 1480.4, expected 1480.5; $[M+Na]^+$ observed 1502.4, expected 1502.5; $[M+K]^+$ observed 1518.4, expected 1518.5) and CDA4a ($[M+H]^+$ observed 1494.4, expected 1494.5; $[M+Na]^+$ observed 1516.3, expected 1516.5; $[M+K]^+$ observed 1533.4, expected 1532.5). The mass spectra of CDA3a and CDA4a were extracted from the LC peaks with retention times of 8.6-8.7 and 8.8-8.9 min respectively.

2.5.2. Non-integrative pIJ86 Vector

The pIJ86 vector is transferable from *E. coli* to *Streptomyces spp.* (the map of pIJ86 is given in section 6.2.2). It is self-replicating in *E. coli* due to the pUC18 origin and can be transferred from *E. coli* (ET12567/pUZ8002) to *Streptomyces spp.* by conjugation via its *oriT* element. Due to the lack of a vector-encoded integrase, pIJ86 does not integrate into the chromosome. Additionally, this vector contains the constitutive *ermE** promoter which has been widely used for protein expression in *Streptomyces spp.*[211, 212].

2.5.2.1. Cloning of *lptJ* and *lptJK* into the pIJ86 Vector and Construction of the Transformant Strains ScpIJ86-*lptJ* and ScpIJ86-*lptJK*

The genes *lptJ* and *lptJK* were amplified by the same procedures as described in section 2.5.1.2. The two gene fragments were digested with the restriction enzyme *Bam*HI and cloned into the compatible restriction enzyme site *Bg*III of pIJ86 (Figure 2.25). After verification by sequencing, both constructs were then used to transform *S. coelicolor* by conjugation with *E. coli*, yielding the transformants ScpIJ86-*lptJ* and ScpIJ86-*lptJK*. Four colonies each from the ScpIJ86-*lptJ* and ScpIJ86-*lptJK* strains were used to inoculate liquid media for plasmid DNA extraction. The plasmid DNAs were then subjected to PCR analysis with primers *ermP-F1* and *ermP-R1*, which amplify the region containing the multiple cloning site of the vector pIJ86. Polymerase chain reaction products of approximately 1300 bp and 2000 bp were generated, corresponding to the expected sizes of *lptJ* and *lptJK* respectively (Figure 2.26), thereby confirming that all transformants contained the pIJ86 constructs.

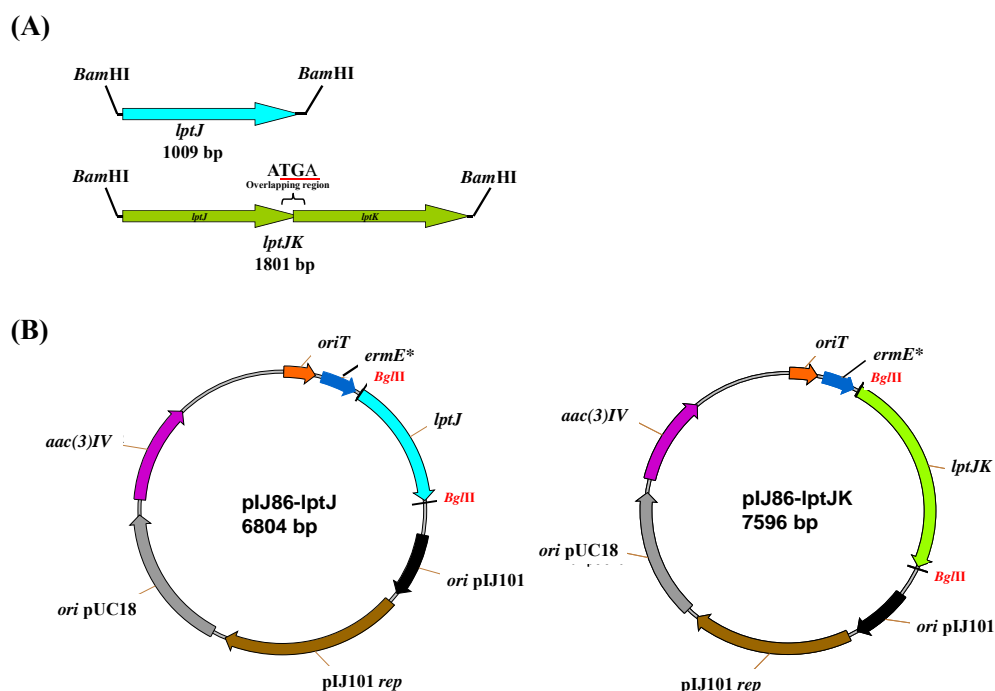


Figure 2.25. Schematic representation of the pIJ86-*lptJ* and pIJ86-*lptJK* constructs.

(A) The genes *lptJ* and *lptJK* were amplified by primers (which contained the *Bam*HI restriction enzyme site) and cloned into pIJ86. (B) The resulting plasmids were designated pIJ86-*lptJ* and pIJ86-*lptJK*.

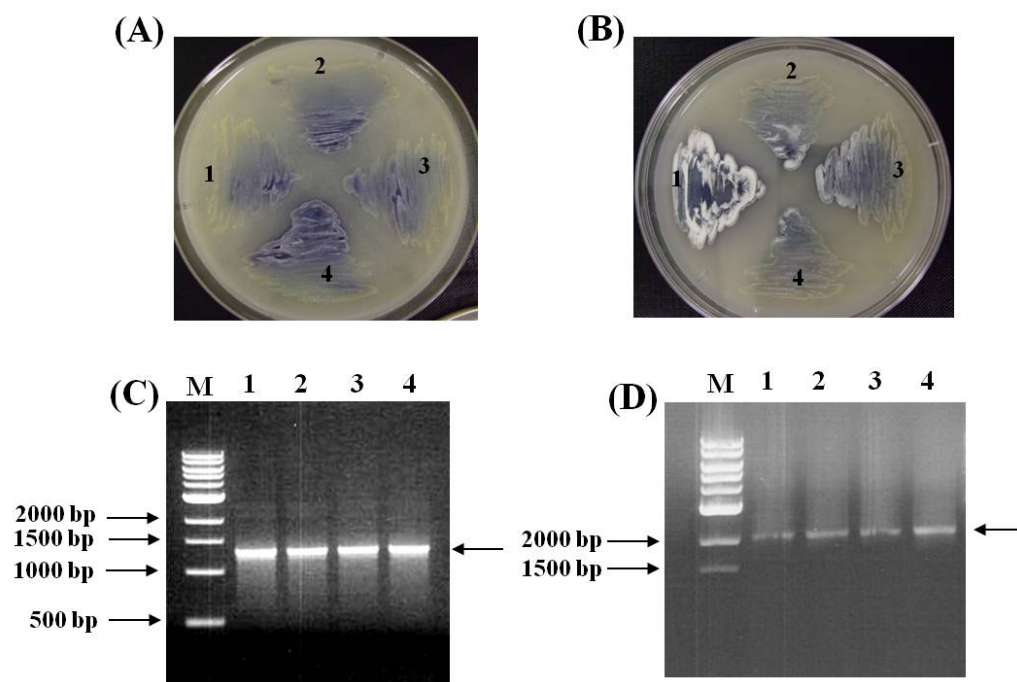


Figure 2.26. Colony screening and PCR analyses of the ScpIJ86-*lptJ* and ScpIJ86-*lptJK* strains.

Four transformant strains of ScpIJ86-*lptJ* (A) and ScpIJ86-*lptJK* (B) were selected on apramycin-containing plates. The PCR results from the plasmid DNA of ScpIJ86-*lptJ*1-4 and ScpIJ86-*lptJK*1-4 are shown in (C) and (D) respectively; Lane M: DNA size marker. Lanes 1-4: PCR products from the four transformants.

2.5.2.2. CDA Bioassay of the ScpIJ86-lptJ and ScpIJ86-lptJK strains

Details of the CDA bioassay procedures are given in section 6.17. No significant differences in the inhibition zone diameters were observed for the ScpIJ86-lptJ and ScpIJ86-lptJK strains compared with that of the wild-type; therefore CDA activity against the indicator strain *M. luteus* appeared unchanged in the transformant strains (Figure 2.27). It was concluded from this result that the transformant strains were either producing wild-type CDA (*i.e.* the pIJ86 constructs had no effect on CDA production), or a variant of CDA with similar biological activity to the wild-type.

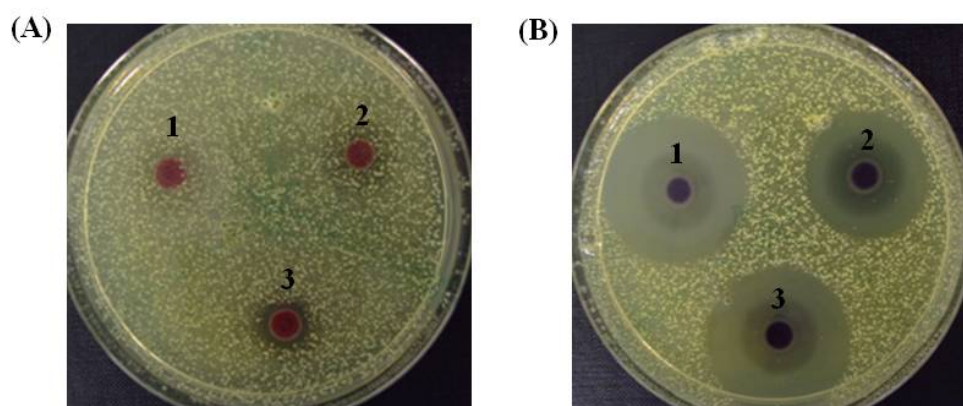


Figure 2.27. Bioassays of the ScpIJ86-lptJ and ScpIJ86-lptJK strains. Wild-type, ScpIJ86-lptJ1 and ScpIJ86-lptJK1 strains are labelled 1, 2 and 3 respectively. **(A)** negative control plate (without calcium). **(B)** bioassay plate (containing calcium).

2.5.2.3. LC-MS Analysis of the Fermentation Broth from the ScpIJ86-lptJ and ScpIJ86-lptJK strains

From 7-day cultures of the ScpIJ86, ScpIJ86-lptJ and ScpIJ86-lptJK strains, crude CDA mixtures were purified using a C₁₈ bond elut cartridge and analysed by LC-MS. The mass spectra from the control strain extract produced CDA3a ($[M+H]^+=1480.5$) and CDA4a ($[M+H]^+=1494.5$) (Figure 2.28). The transformant strains ScpIJ86-lptJ and ScpIJ86-lptJK also produced CDA3a and CDA4a with the same masses to those observed for the control strain, *i.e.* with no modification (Figure 2.29 and Figure 2.30).

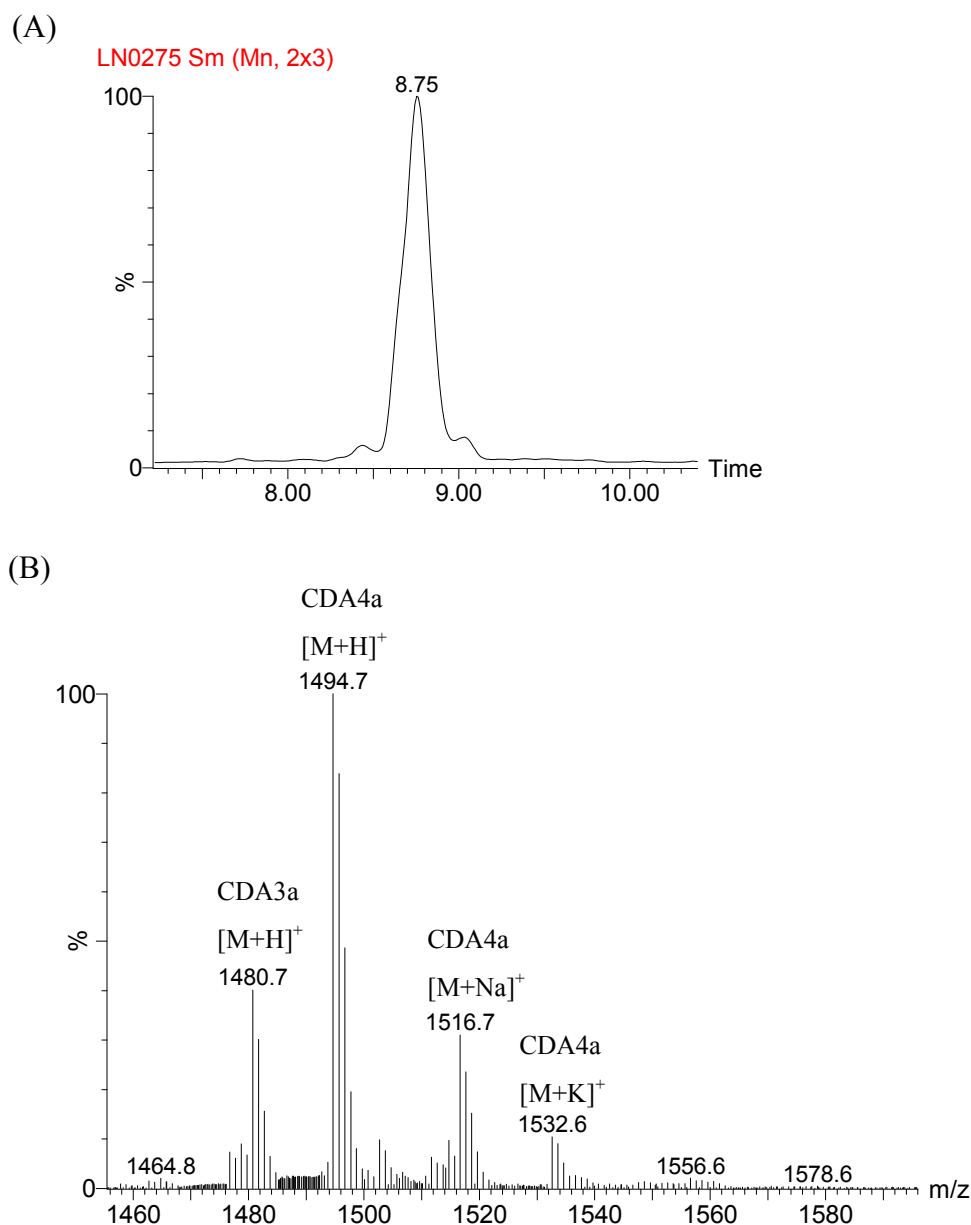


Figure 2.28. LC-MS analysis of ScpIJ86 fermentation products.

(A) Chromatogram of fermentation products from the strain ScpIJ86. (B) The mass spectra of CDA3a ([M+H]⁺ observed 1480.7, expected 1480.5) and CDA4a ([M+H]⁺ observed 1494.7, expected 1494.5; [M+Na]⁺ observed 1516.7, expected 1516.5; [M+K]⁺ observed 1532.6, expected 1532.5). The mass spectra were extracted from the LC peaks with retention times of 8.3-9.3 min respectively.

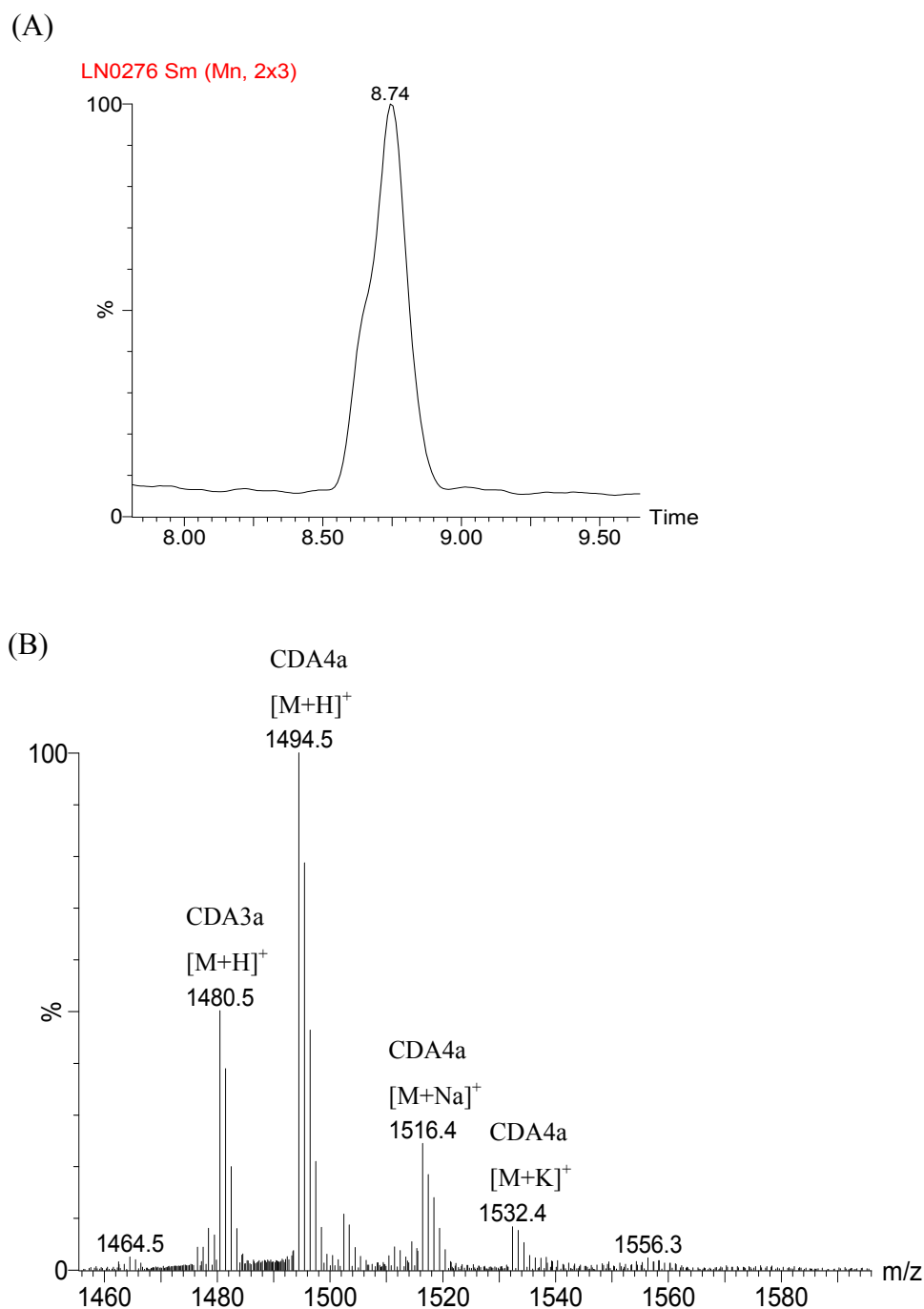


Figure 2.29. LC-MS analysis of ScpIJ86-lptJ fermentation products.

(A) Chromatogram of fermentation products from the strain ScpIJ86-lptJ. (B) The mass spectra of CDA3a ($[M+H]^+$ observed 1480.5, expected 1480.5) and CDA4a ($[M+H]^+$ observed 1494.5, expected 1494.5; $[M+Na]^+$ observed 1516.4, expected 1516.5; $[M+K]^+$ observed 1532.4, expected 1532.5). The mass spectra were extracted from the LC peaks with retention times of 8.5-9.0 min.

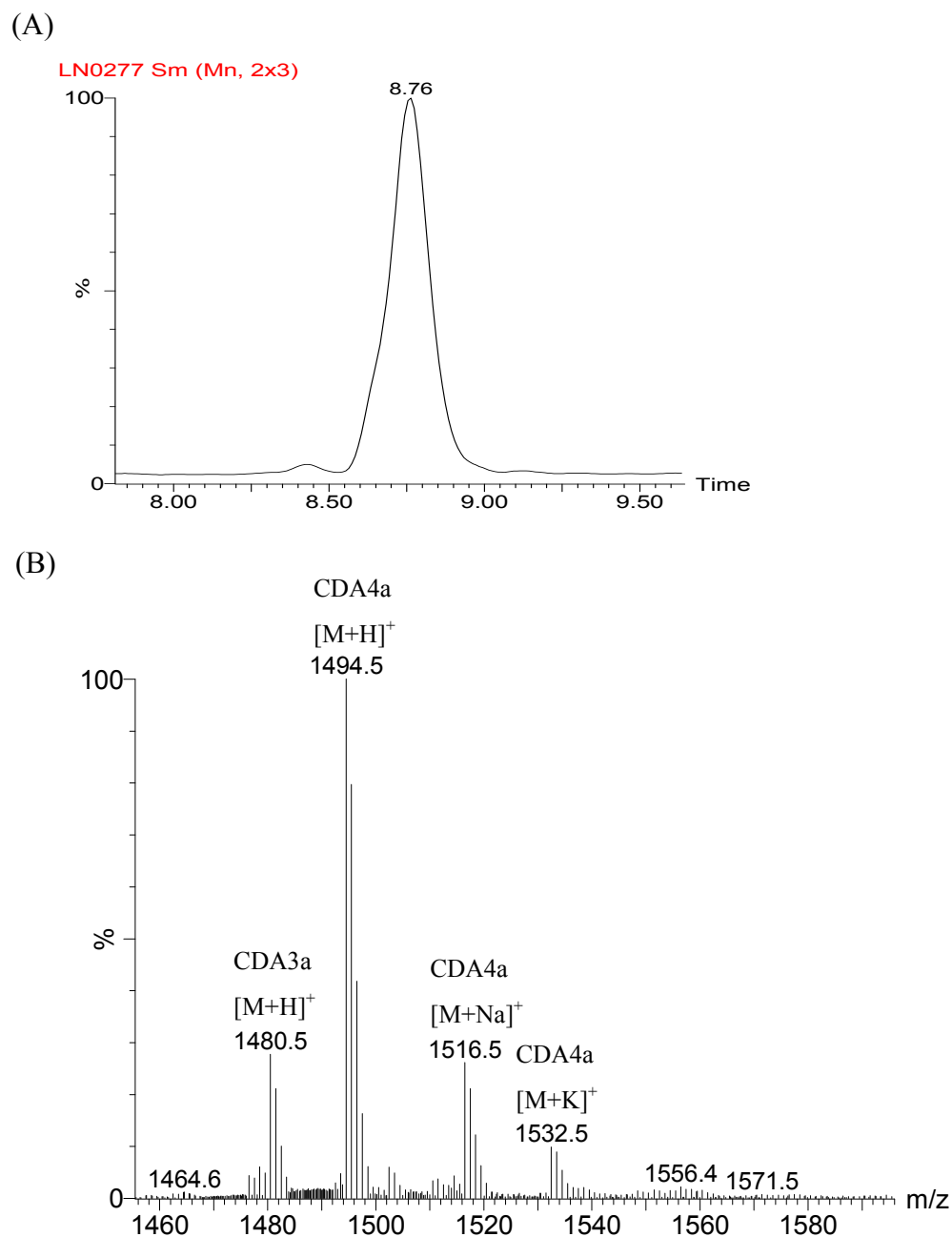


Figure 2.30. LC-MS analysis of ScpIJ86-lptJK fermentation products.

(A) Chromatogram of fermentation products from the strain ScpIJ86-lptJK. (B) The mass spectra of CDA3a ($[M+H]^+$ observed 1480.5, expected 1480.5) and CDA4a ($[M+H]^+$ observed 1494.5, expected 1494.5; $[M+Na]^+$ observed 1516.5, expected 1516.5; $[M+K]^+$ observed 1532.5, expected 1532.5). The mass spectra were extracted from the LC peaks with retention times of 8.4-9.1 min.

2.6. Discussion

From the *in vitro* enzymatic characterisations, we found that LptL is responsible for the biosynthesis of L-OH-Asn. LptL shows amino acid sequence homology to that of AsnO of which Strieker *et al.* (2007) published the X-ray crystal structure and showed the stereochemistry of L-OH-Asn being (2S, 3R)-*trans* configured [111]. Subsequently they engineered AsnO D241N to stereospecifically hydroxylate L-Asp to yield L-*trans*-OH-Asp in 2008 [213]. They also showed that they found LptL was able to stereospecifically hydroxylate L-Asn [214]. Therefore we predicted that the product of LptL is also L-(2S, 3R)-*trans*-OH-Asn and did not do any more characterisation on the LptL reaction product OH-Asn or engineering works on LptL.

We postulated that OmAsp is biosynthesised by two sequential steps of hydroxylation and then methylation. Comparison of the Asp- and OmAsp-activating A domains shows near identical substrate recognition sequences for Asp, indicating that OmAsp is not recognised as a separate substrate and is not incorporated as a free amino acid. From this we speculated that OmAsp biosynthesis may occur during peptide assembly using PCP-tethered Asp as a substrate.

The *in vitro* characterisations of the LptJ and LptK enzymes showed that they do not act to modify free amino acids although this may be due to incorrect protein folding leading to reduced enzymatic activity. Additionally we did not have the ideal substrate *i.e.* PCP-tethered Asp for the assays. Even with correct folded active enzymes, it would be unlikely that we would observe modified products due to the lack of correct substrate. Therefore we did not deal with the issue of protein folding in an *E.coli* host and directly adopted an *in vivo* strategy using a *Streptomyces* host which contains a PCP-tethered Asp substrate to study OmAsp biosynthetic pathway.

In our *in vivo* experiments, two different expression systems were employed to over-express the LptJ and LptK proteins in *S. coelicolor*. However we did not observe the hydroxylation and methylation modifications on the Asp residue of CDA from the transformant strains pMT3226-lptJ, pMT3226-lptJK, pIJ86-lptJ, and pIJ86-lptJK.

We proposed several factors which might be responsible for these results. **Firstly, protein structure:** the enzymes LptJ and LptK may fold improperly in *S. coelicolor*. Additionally, there is a significant diversity between the organisation of the A54145 and CDA synthetases (Figure 1.24 and Figure 1.28). The CDA cluster contains three synthetases possessing six, three and two modules, whereas the A54145 cluster has four synthetases with five, two, four and two modules. Therefore the Asp-recognition synthetases of CDA and A54145 are expected to have different conformations which may affect LptJ and LptK substrate recognition and NRPS association. Even if LptJ and LptK were expressed and folded correctly, they may possess high specificity for their substrate binding sites, *i.e.* their active sites only accept Asp tethered to the PCP domain of the antibiotic A54145 synthetase. **Secondly, the C domains of NRPSs** have been demonstrated to act as gatekeepers for incorporation of specific substrates into the peptide chain [78]. The C domain of the Asp activating-module of CDA synthetase may be unable to recognise OmAsp, resulting in the non-production of CDA containing OmAsp, even with the expression of functional enzymes of LptJ and LptK in *S. coelicolor*. Recently Alexander *et al.* (2011) assigned the functions of hydroxylation and methylation during the biosynthesis of the OmAsp to LptJ and LptK by using mutagenesis strategy [215]. Their results support our earlier assumption as to the functions of LptJ and LptK.

Chapter 3. Bio-engineering of the Lipopeptide Antibiotic Enduracidin by Chlorination and Mannosylation

3.1. Introduction

Details of the biosynthesis, structure and mode of action of the antibiotics enduracidin and ramoplanin are given in section 1.9. Due to their potent activity and distinct mode of action, they are potential new therapeutic agents for combating β -lactam- and vancomycin-resistant pathogens. It is therefore worth producing new variants of enduracidin and ramoplanin to optimise the properties of these antibiotics. For example, enduracidin is poorly soluble in water which may possibly be resolved by the addition of a sugar moiety.

Here we would like to characterise two tailoring enzymes; a halogenase (Ram20; section 3.4) and a mannosyl transferase (Ram29; section 3.5) from the ramoplanin biosynthetic gene cluster [6]. The enzymes Ram20 and Ram29 are predicted to chlorinate hydroxyl phenylglycine-17 (Hpg-17) and mannosylate Hpg-11 of ramoplanin respectively. Due to the slow growth rate of the ramoplanin producer strain *Actinoplanes* ATCC33076, and the close structural similarity between ramoplanin and enduracidin, we decided to over-express both enzymes in the enduracidin producer *S. fungicidicus* to characterise their functions. We anticipated that chlorination and mannosylation by Ram20 and Ram29 may also be possible at amino acids Hpg-17 and Hpg-11 respectively of enduracidin, resulting in the production of new enduracidin analogues (Figure 3.1).

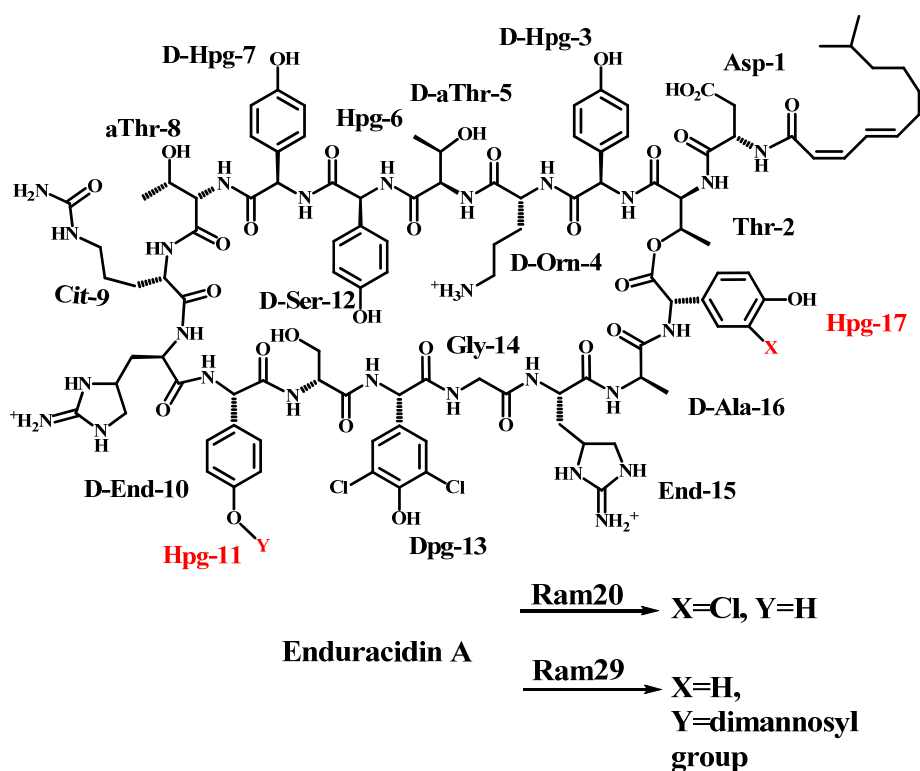


Figure 3.1. Predicted enduracidin A analogues after modification by the tailoring enzymes Ram20 and Ram29.

3.2. Isolation and Characterisation of Enduracidin from *S. fungicidicus* ATCC31731 (Strain No. Emt 2-140)

Enduracidin was first isolated from *S. fungicidicus* ATCC21013 in 1968 by Takeda Chemical Industries [145]. This parent strain cannot efficiently excrete enduracidin, limiting the utility of this strain for screening assays during antibiotic engineering [216]. Therefore a mutant strain, *S. fungicidicus* Emt 36-3, with the ability to excrete enduracidin was created using the mutagen m-fluoro-DL-tyrosine [216]. Subsequently a double mutant strain, *S. fungicidicus* Emt 2-140, which has both the properties of excretion and high production of enduracidin, was obtained from the selection of *S. fungicidicus* Emt 36-3 using other mutagenesis methods including ultra-violet, X-ray irradiation and treatment with N-methyl-N'-nitro-N-nitrosoguanidine [216]. The Takeda Chemical company confirmed that the mutant strain *S. fungicidicus* Emt 2-140 produced enduracidin at a minimal 4 grams per litre of medium after a 9-day fermentation [216]. Therefore we chose this strain (*S. fungicidicus* ATCC31731) for our study.

The bioassay system was set up following the method developed by Higashide *et al.* in 1968 [145]. A bouillon agar plate supplemented with 5 % glycerol allows for enduracidin diffusion and therefore can be used as a visible assay of antibiotic activity against a *B. subtilis* PY79 indicator strain [145]. By using this assay system we can observe the ability of the strain *S. fungicidicus* ATCC31731 to excrete enduracidin (Figure 3.2A). An enduracidin standard was used to correlate the diameter of the inhibition zone with the concentration of enduracidin, thus enabling estimation of the yield of enduracidin from samples of fermentation broth (Figure 3.2B and C).

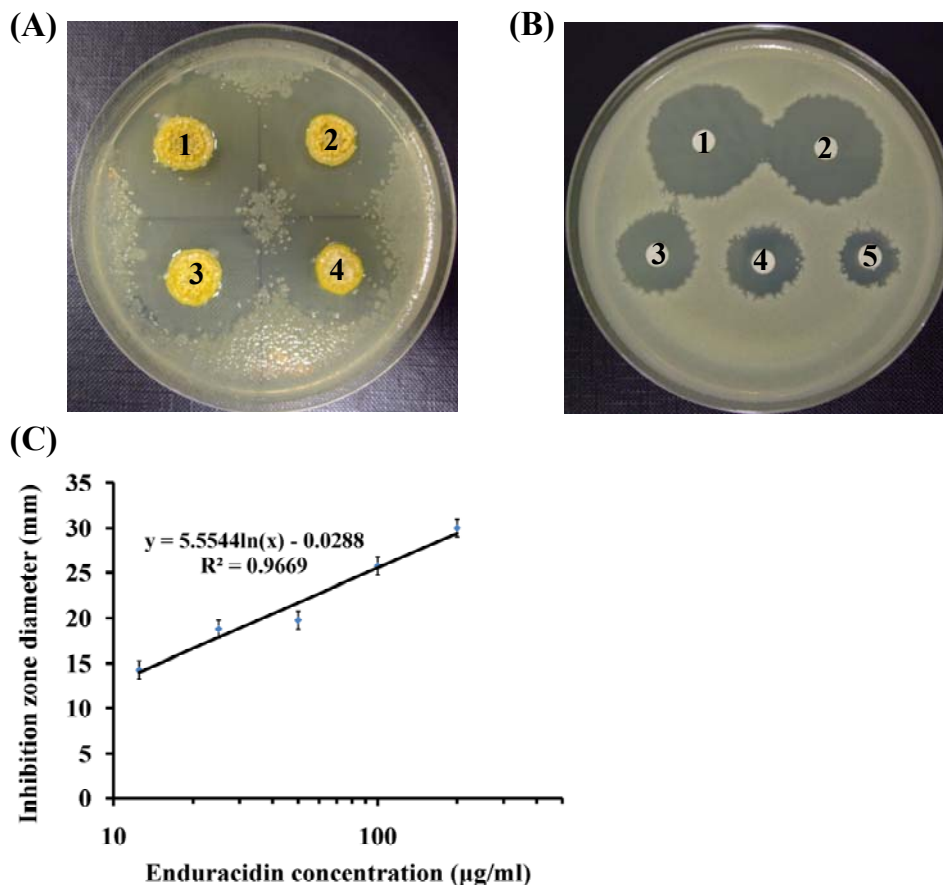


Figure 3.2. Enduracidin activity test.

(A) Spot inoculation test of *S. fungicidicus* (ATCC31731) enduracidin excretion from 1×10^5 (spot 1 and 2) and 2×10^5 (spot 3 and 4) spores after 5 days incubation. (B) Activity test of enduracidin standard, concentrations are 200, 100, 50, 25 and 12.5 µg/ml of enduracidin (spots 1-5). (C) Standard curve for enduracidin. Enduracidin activities are correlated to the diameter of the inhibition zone (mm; left axis) and the error bars represent the mean±SD of four replicates. A linear regression line can be obtained for the 12.5-200 µg/ml range of enduracidin.

The purification procedure of enduracidin from *S. fungicidicus* ATCC31731 was modified from the protocol developed by Asai *et al.* in 1968 [146]. Enduracidin excreted from *S. fungicidicus* ATCC31731 is poorly soluble in water at neutral pH (but soluble in acidic water pH 2.0) and therefore can be collected together with the mycelia from the culture broth after centrifugation. The purification procedure is summarised in Figure 3.3 (A) and the details are given in section 6.19. The presence of enduracidin at each step was analysed by bioassay (Figure 3.3B and C). The results showed that most of the enduracidin was recovered in the mycelial fraction, rather than in the supernatant fraction of the culture broth (spot 1 and 2; Figure 3.3B). Additionally we found that a large amount of enduracidin was present within the mycelia when we lysed the cells with 70 % acidic methanol (spot 3; Figure 3.3B).

Therefore in this study we employed 70 % acidic methanol for lysis, followed by a 1-Butanol/Water purification procedure. We can obtain at least 15 mg of crude enduracidin from a one litre culture based on the bioassay result. Finally the crude powder product was purified further by RP-HPLC.

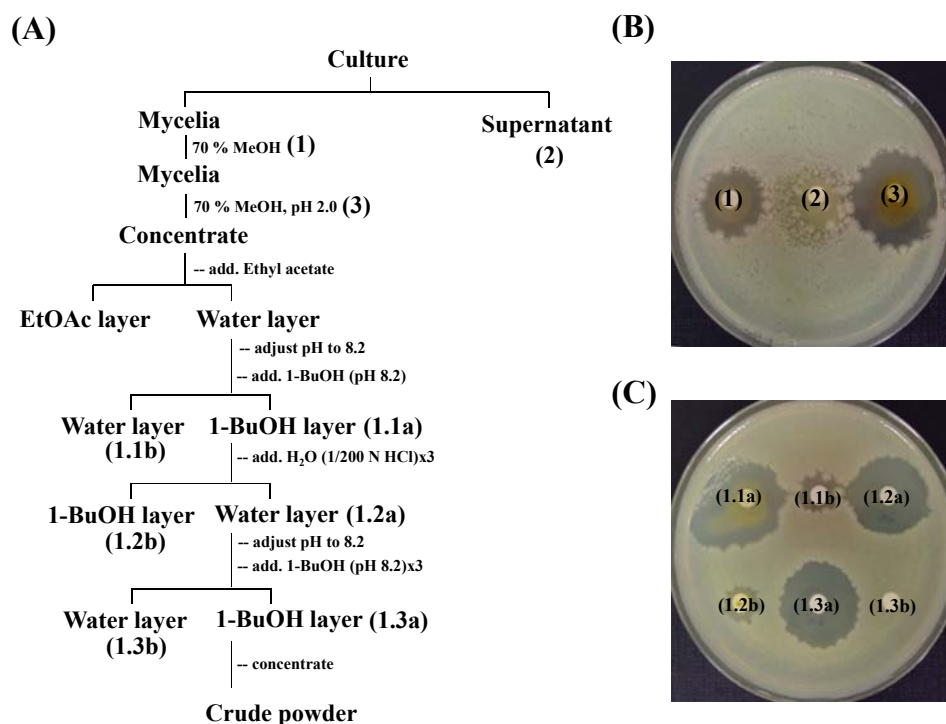


Figure 3.3. Enduracidin purification procedure.

(A) The isolation procedure. (B) Preliminary fractions tested for activity: mycelia (1), supernatant (2), mycelial lysate (3). (C) Enduracidin activity test of fractions from the 1-Butanol/Water extraction steps.

In the RP-HPLC chromatogram enduracidin A and B from *S. fungicidicus* have the same retention time as that of the enduracidin A and B standards (Figure 3.4). The mass spectrometric analysis of the enduracidin A and B revealed that the parent ions are observed mainly in the triply and doubly charged state ($[M+3H]^{3+} = 785.3$ and 790.0 ; $[M+2H]^{2+} = 1177.5$ and 1184.4 respectively), with few singly charged ions ($[M+H]^+ = 2352.3$ and 2366.3 respectively) (Figure 3.5). Enduracidin contains three readily protonated basic amino acids (two Enduracididine (End) and one Ornithine (Orn) residues); explaining the observed prevalence of multiply charged states.

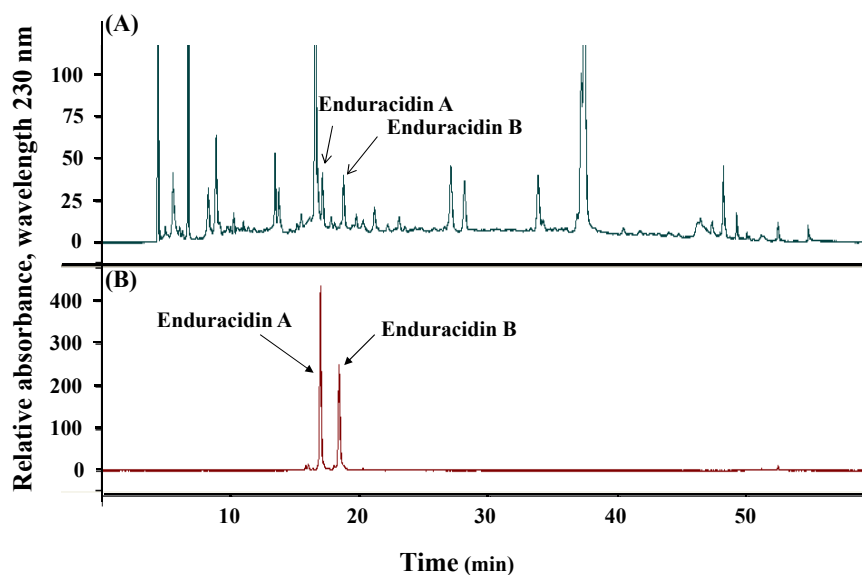


Figure 3.4. RP-HPLC analysis of the *S. fungicidicus* crude extract.

(A) Crude extract from the fermentation broth of the wild-type *S. fungicidicus*. (B) Enduracidin A and B standards.

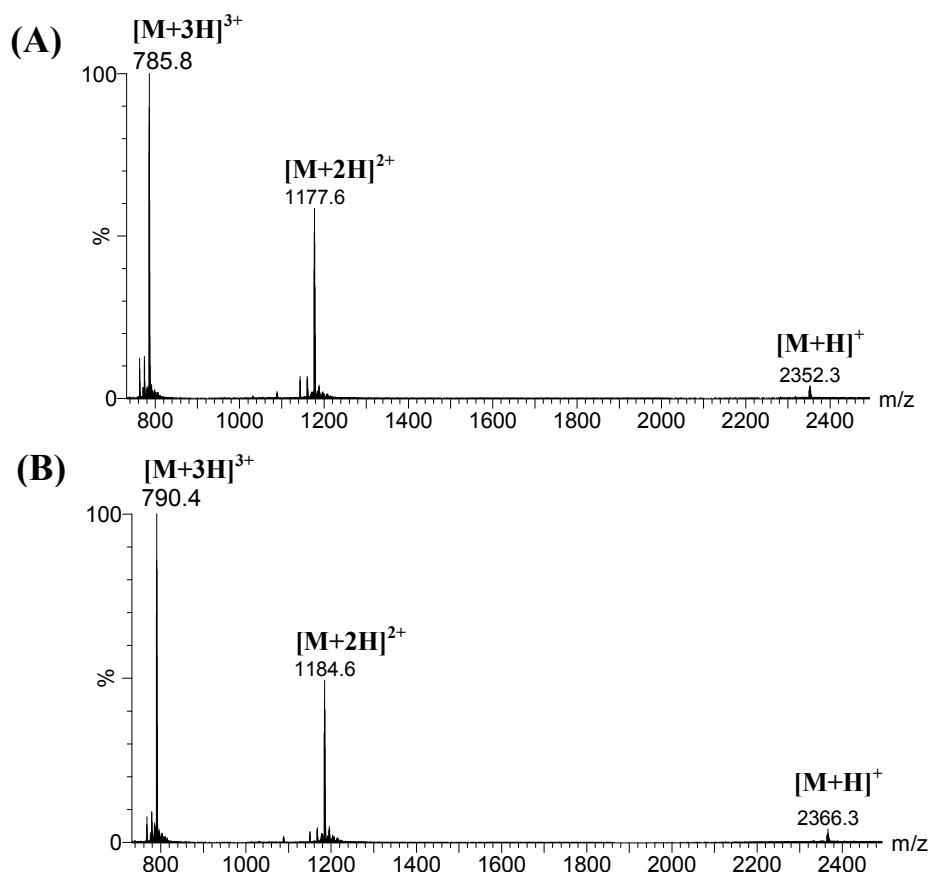


Figure 3.5. LC-MS analysis of enduracidin A and B.

The spectra show the singly, doubly and triply charged ions for (A) Enduracidin A ($[M+H]^+$ observed 2352.3, expected 2353.9; $[M+2H]^{2+}$ observed 1177.6, expected 1177.5; $[M+3H]^{3+}$ observed 785.8, expected 785.3) and (B) enduracidin B ($[M+H]^+$ observed 2366.3, expected 2367.9; $[M+2H]^{2+}$ observed 1184.6, expected 1184.4; $[M+3H]^{3+}$ observed 790.4, expected 790.0).

3.3. Verification of the pMS17 Vector Functionality in *S. fungicidicus*

Earlier we cloned the genes *ram20* and *ram29* along with their own SD sequences into two different *Streptomyces* expression vectors, pIJ86 (a non-integrative and high copy number vector containing a constitutive promoter *ermE**) and pMS17 (an integrative vector containing a tetracycline-inducible promoter). Four *S. fungicidicus* transformant strains were created: pIJ86-*ram20*, pIJ86-*ram29*, pMS17-*ram20* and pMS17-*ram29*. However, we did not find any modification of enduracidin from these four transformants, possibly due to problems with protein expression. Therefore we verified protein expression from the pMS17 vector using enhanced green fluorescence protein (eGFP) as a reporter in *S. fungicidicus*.

The *eGFP* sequence was obtained from the Bibb group as a construct in pIJ8668 and has previously been confirmed to act as a reporter for spatial and temporal gene expression studies in *Streptomyces spp.* [217-219]. Its sequence differs from wild-type GFP by two amino acid substitutions (Ser-65 to Thr and Phe-64 to Leu) and by optimisation of codons for expression in *Streptomyces spp.* (Figure 3.6A). These changes lead to detectable eGFP expression and fluorescence in *Streptomyces spp.* [217]. The gene fragment of approximately 750 bp was amplified using the primers eGFP-F1 and eGFP-R1 (with eGFP-F1 encompassing the SD sequence) and cloned into the *EcoRV* site of pMS17 (Figure 3.6B and C). The pMS17-eGFP construct was then used to transform *S. fungicidicus* by conjugation, creating the transformant strain SfpMS17-eGFP.

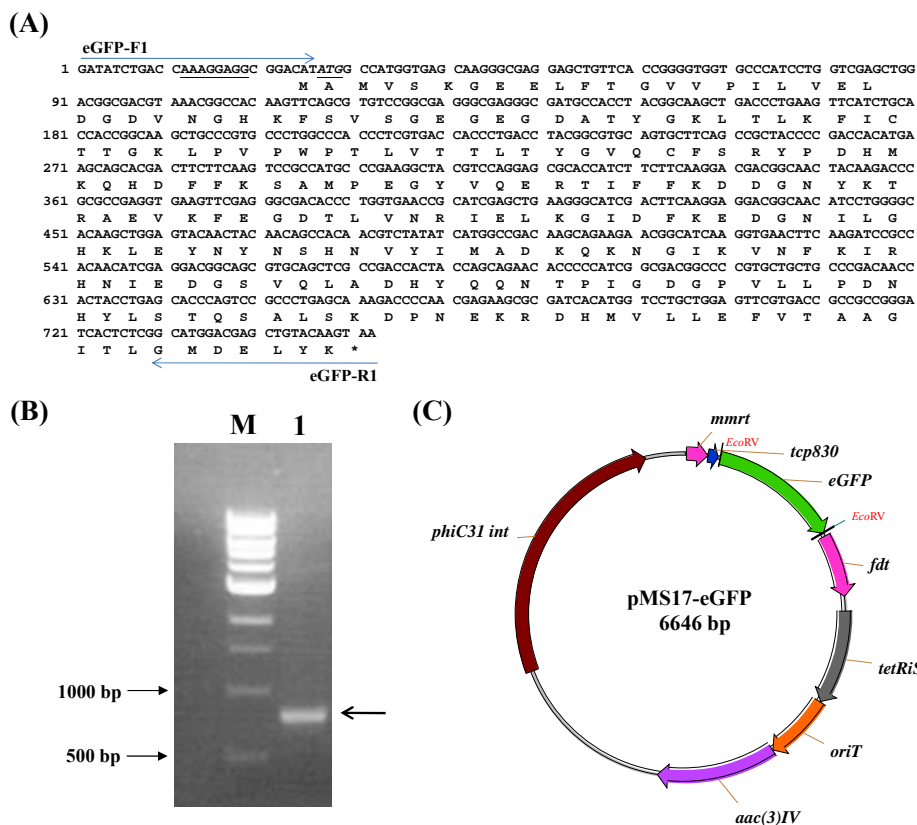


Figure 3.6. The sequence and PCR amplification of eGFP.

(A) The eGFP ORF comprises the sequence between the start codon ATG (underlined in italics) and the stop codon TAA (marked by an asterisk). The SD sequence “AAAGGAGG” is also underlined. Forward and reverse primers (eGFP-F1 and eGFP-R1) are denoted by arrows. (B) Agarose gel showing the *eGFP* fragment of approximately 750 bp (lane 1). Lane M: DNA size marker. (C) Schematic representation of the construct pMS17-eGFP.

The SfpMS17-eGFP strain was cultured in Bouillon medium with the antibiotic apramycin for 2 days and eGFP expression was induced by the addition of 0.1 $\mu\text{g/ml}$ anhydro-tetracycline (aTc). The mycelia were collected after 24 h, washed with water and fixed on glass slides for microscopy. As shown in Figure 3.7, the green fluorescence was observed by fluorescence microscopy (DeltavisionRT QLM Deconvolution Microscope). Although wild-type *S. fungicidicus* showed a faint green background fluorescence, we can observe significantly stronger green fluorescence from the SfpMS17-eGFP strain. Results from inducer concentrations higher than 0.1 $\mu\text{g/ml}$ showed no significant increase in eGFP expression. This result matches that of Smith and coworkers who cloned luciferase into the pMS17 vector and showed that expression reached saturation when using 0.1 $\mu\text{g/ml}$ of aTC [220]. Therefore this concentration was adopted for subsequent experiments.

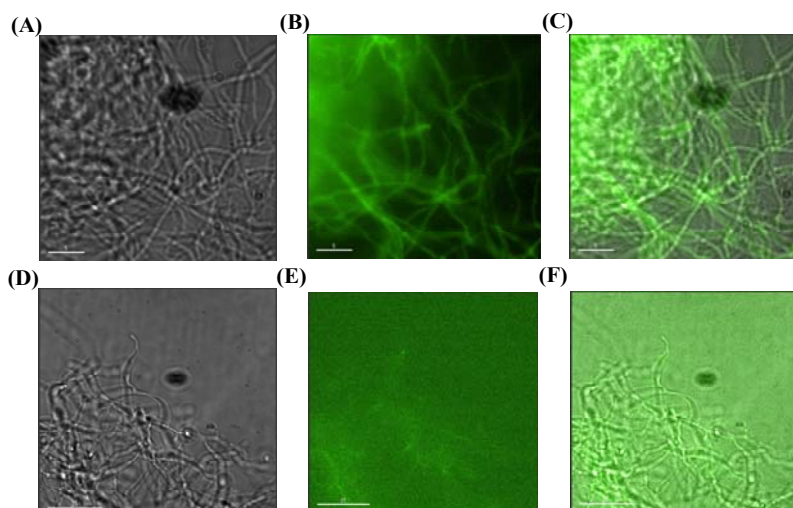


Figure 3.7. eGFP expression in SfpMS17-eGFP strain (A-C) and wild-type *S. fungicidicus* (D-F).

Bright field image (A and D), Fluorescence image (B and E), Merge image (C and F).

The eGFP expression studies in *S. fungicidicus* confirmed the ability of the vector pMS17 to over-express this protein. Comparing the upstream sequence of both *ram20* and *ram29* to that of *eGFP* highlighted a number of key differences (Figure 3.8). The start codon of both genes is GTG, a rare start codon in bacteria, with limited protein expression, which is often found for genes that are involved in the biosynthesis of secondary metabolites. Additionally, the SD sequence for *ram29* is composed of six purine nucleotides rather than the eight purine nucleotides for *eGFP*, which may lower its affinity for the ribosomal 30S subunit. Furthermore, the distance between the SD sequence and start codon is also a crucial parameter for protein expression [221]. Therefore we replaced the upstream sequence of both genes (encompassing the SD sequence and start codon) with the equivalent region from *eGFP*. The details of the *ram20* and *ram29* studies using these modified upstream sequences are given in sections 3.4.4 and 3.5.4.1.

<i>ram20</i>						
5'	AGCTCAGATG	GTCGCTGAGG	AGGATTCGTG	GCTGCTCAAC	CGGAAGAGTT	CGATGTCATC
3'	TCGAGTCTAC	CAGCGACTCC	TCCTAAGCAC	CGACGAGTTG	GCCTTCTCAA	GCTACAGTAG
<i>ram29</i>						
5'	CGAAACGCCG	GAGGATGCCT	TGGCCGCCGT	GATGTCCCG	AGGGTGCGCC	CGCCCGGTGC
3'	GCTTTGCCGC	CTCCTACGGA	ACCGGCCGCA	CCTACAGGGC	TCCCACGCCG	GCGGGCCACG
<i>eGFP</i>						
5'	TGACCAAAGG	AGGCGGACAT	ATGGCCATGG	TGAGCAAGGG	CGAGGAGCTG	TTCACCGGGG
3'	ACTGGTTTCC	TCCGCCTGTA	TACCGGTACC	ACTCGTTCCC	GCTCCTCGAC	AAGTGGCCCC

Figure 3.8. The upstream sequences of *ram20*, *ram29* and *eGFP*.

The SD sequences of *ram20*, *ram29* and *eGFP* are labelled in red; the start codons are labelled in green.

3.4. Bio-engineering of Enduracidin with Chlorination by the Halogenase Ram20

3.4.1. Introduction

Chlorination is found in many natural antibiotic structures such as vancomycin [97], teicoplanin [94], complestatin [222] and syringomycin E [128, 223] (Figure 3.9) as well as the lipopeptide enduracidin [5] and the glycolipopeptide ramoplanin [6]. In the case of vancomycin, teicoplanin and syringomycin E, it has been confirmed that the incorporation of chlorine can enhance antibiotic activity [128, 224, 225]. Chlorination in vancomycin aids peptide dimerisation, thus increasing peptide affinity towards the bacterial cell wall and therefore its efficacy at killing bacteria [224, 226]. However the chemical synthesis methods for integrating chlorine into such compounds are laborious and lack both specificity and regio-selectivity. Therefore the exploitation of chlorinase enzymes is a promising strategy for generating new chlorinated products.

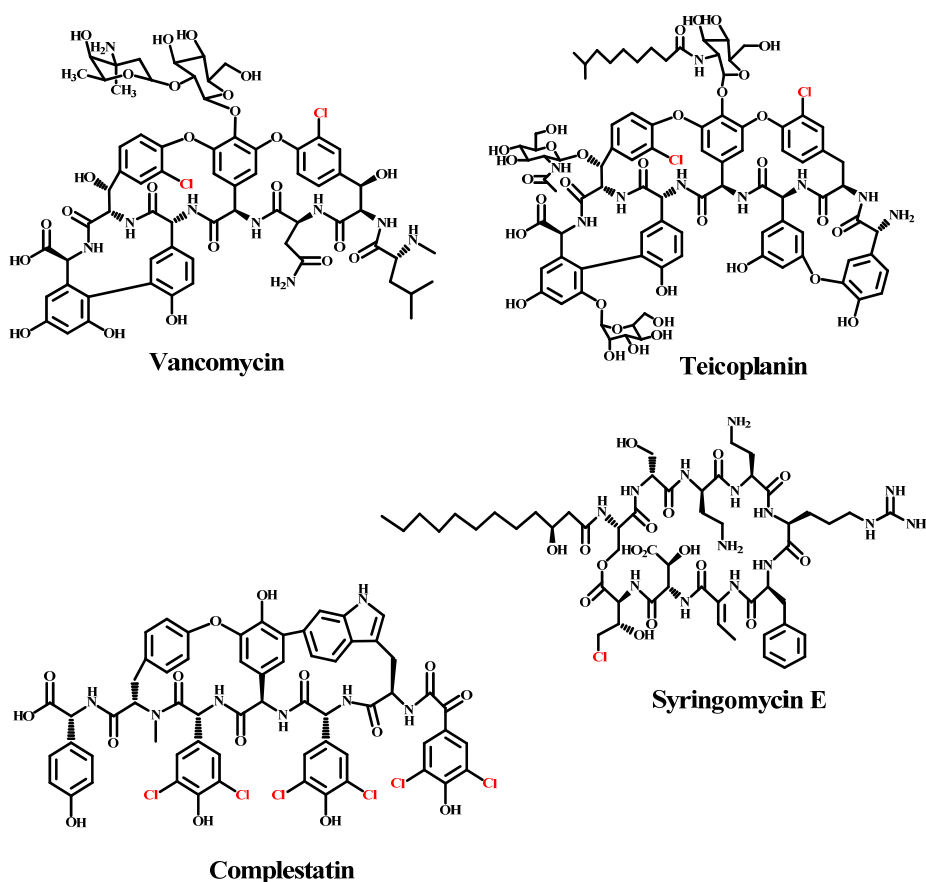


Figure 3.9. Structures of chlorine-containing natural products.

Chlorine atoms in vancomycin, teicoplanin, complestatin and syringomycin E are labelled in red.

Halogenases, such as ComH, Orf18, Tcp21 and SyrB2 have been identified by sequence analysis from the biosynthetic gene clusters of complestatin, vancomycin, teicoplanin and syringomycin E respectively [123]. ComH is predicted to dichlorinate the three Hpg amino acids of complestatin, and Orf18 and Tcp21 are predicted to chlorinate Tyr and β -hydroxyl Tyr residues in vancomycin and teicoplanin. All three halogenases are predicted to be FADH₂-dependent halogenases. Alternatively, SyrB2 has been confirmed as a non-heme Fe(II)/ α -KG-dependent halogenase and can chlorinate Thr when tethered to the PCP domain of syringomycin E synthetase [128].

3.4.2. Bioinformatic Analysis of Ram20

We are interested in the monochlorinated-Hpg (Chp) residue of ramoplanin. In the analysis of the ramoplanin biosynthetic gene cluster, a putative chlorinase encoded by *ram20* showed homology (above 60 % identity) to putative chlorinases from the biosynthetic gene clusters of balhimycin (vancomycin-type antibiotic), complestatin, teicoplanin and enduracidin (Table 3.1). The gene *bhaA* from the balhimycin biosynthetic gene cluster has been confirmed as a chlorinase by gene disruption [227]. These enzymes are predicted to be FADH₂-dependent chlorinases due to the presence of two conserved motifs, the “GGGxxG” and “GWTWxIP” motifs [123, 124] (Figure 3.10). Accordingly, Ram20 is the only chlorinase present in the ramoplanin gene cluster and is predicted to carry out monochlorination of Hpg¹⁷ to form Chp¹⁷.

Table 3.1. Identification of Ram20 function by a Blast protein sequence search

Protein name	Source	Accession number	Identity (%)
ComH	Complestatin biosynthetic gene cluster in <i>Streptomyces lavendulae</i>	AAK81830	65
Tcp21	Teicoplanin biosynthetic gene cluster in <i>Actinoplanes teichomyceticus</i>	CAG15020	65
End30	Enduracidin biosynthetic gene cluster in <i>S. fungicidicus</i>	ABD65950	64
BhaA	Balhimycin biosynthetic gene cluster in <i>Amycolatopsis balhimycina</i>	CAA76550	63

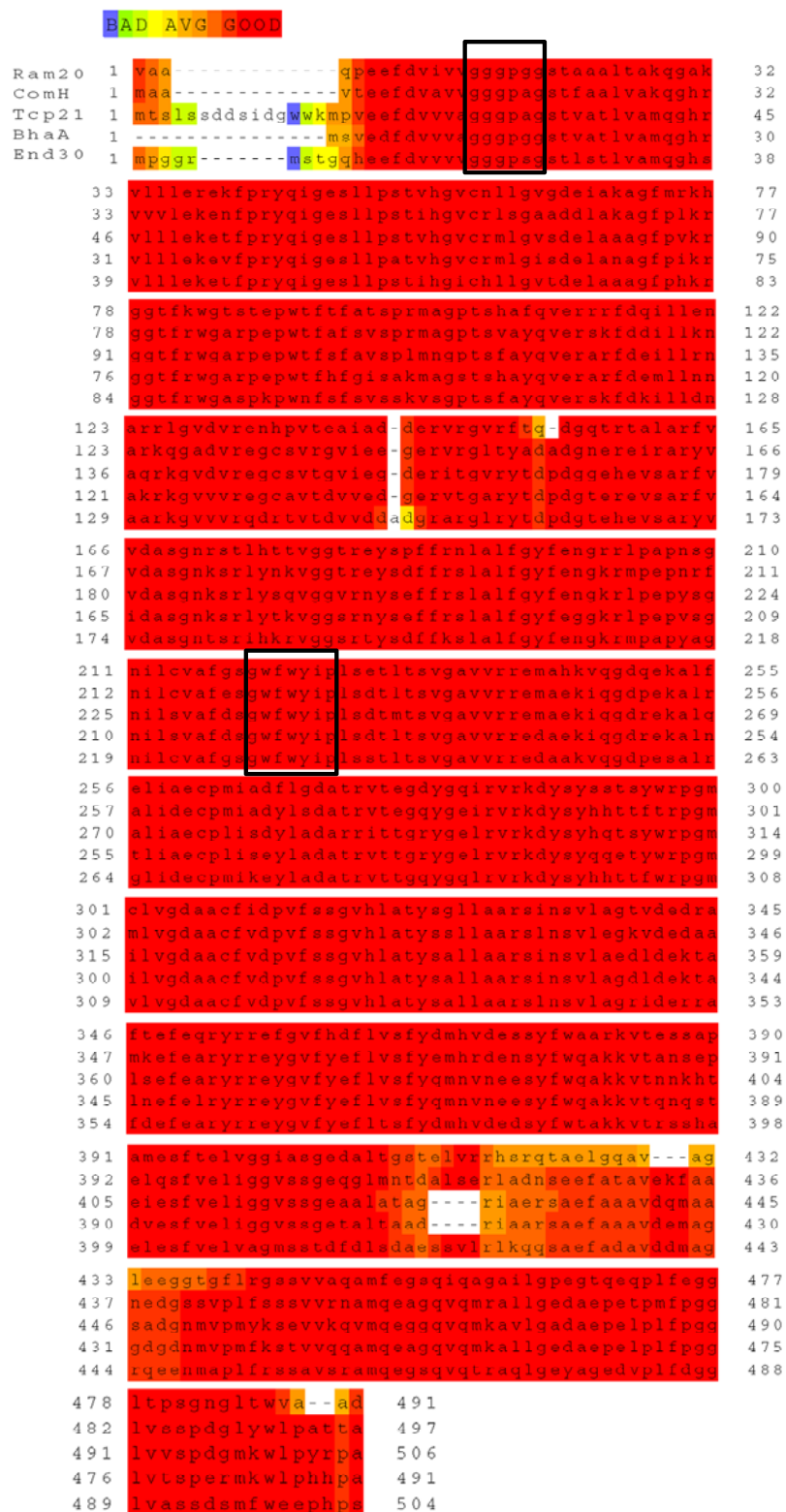


Figure 3.10. Ram20 amino acid sequence alignment with known chlorinases.

The amino acid sequence of Ram20 was aligned with that of ComH, Tcp21, End30 and BhaA from *S. lavendulae*, *A. teichomyceticus*, *S. fungicidicus* and *A. balhimycina* respectively using the online alignment program T-coffee [201]. The two conserved “GGGxxG” and “GWTWxIP” motifs, indicative of FADH₂-dependent chlorinases, are highlighted by open boxes.

3.4.3. Timing of Chlorination

The timing for covalent incorporation of chlorine into antibiotic products is still unclear. Chlorination may occur on free amino acids (before peptide assembly), on amino acids tethered to the PCP domains (during peptide assembly) or peptide products (after peptide assembly). With the former, as ramoplanin and enduracidin contain both chlorinated and non-chlorinated Hpgs, their corresponding A domains might possess a specific substrate-binding pocket to differentiate between chlorinated and non-chlorinated Hpgs.

A comparison of the substrate recognition sequences of A domains from the modules which activate the amino acids Hpg, Chp and Dpg within the enduracidin and ramoplanin synthetases reveals the invariance of these motifs. As shown in Table 3.2, the substrate-binding pocket residues of the A domain from the Ram14-m8 module which corresponds to Chp-17, are identical to those of the other four A domains predicted to activate Hpgs in ramoplanin, with only one amino acid difference occurring in the Ram13-m2 module. Additionally, in enduracidin the substrate-binding pocket residues of the A domain for recognition of Dpg are the same to those of the other four Hpgs, with only one amino acid difference occurring in the EndC-m2 module. Interestingly, the substrate-binding pocket residues of the A domains for Chp and Dpg in enduracidin and ramoplanin are identical to those of the other eight Hpgs, with only one amino acid difference each in the Ram13-m2 module and EndC-m2 module. Accordingly we expected that the timing of chlorination on Hpg residues occurs during peptide assembly, after Hpg selection by the A domain.

Table 3.2. The substrate-binding pocket residues of Hpg-, Dpg- and Chp-activating A domains

Module	Activated amino acid	Substrate-binding pocket residues of A domain
Ram13-m1	D-Hpg-3	DAYHLGLLC
Ram13-m4	L-Hpg-6	DAYHLGLLC
Ram13-m5	D-Hpg-7	DAYHLGLLC
Ram13-m2	L-Hpg-11	DA F HLGLLC
Ram14-m4	L-Hpg-13	DAYHLGLLC
Ram14-m8	L-Chp-17	DAYHLGLLC
EndB-m1	D-Hpg-3	DAYHLGLLC
EndB-m4	L-Hpg-6	DAYHLGLLC
EndB-m5	D-Hpg-7	DAYHLGLLC
EndC-m2	L-Hpg-11	DAYHL G MLC
EndC-m4	L-Dpg-13	DAYHLGLLC
EndC-m8	L-Hpg-17	DAYHLGLLC

Note: Ram and End represent the ramoplanin and enduracidin synthetases respectively. The substrate-binding pocket residues of the A domains are obtained by alignment of the A4-5 motifs with that of GrsA.

3.4.4. Construction of the Transformant Strain SfpMS17-ram20

The *ram20* gene (approximately 1500 bp) was amplified from the *Actinoplanes* strain ATCC33076 genome by the primers sfram20-F2 and sfram20-R1. The forward primer sfram20-F2 was designed with an additional 23 bp from the upstream sequence of *eGFP*, encompassing the SD sequence and start codon ATG. After digestion with the restriction enzyme *EcoRV*, *ram20* was cloned into the pMS17 vector (Figure 3.11A and B). The resultant construct pMS17-ram20 was then transformed into *S. fungicidicus* by conjugation and transformants were selected by the antibiotic apramycin (Figure 3.11C). Four transformants were chosen for further PCR analysis, confirming that the gene *ram20* was in each transformant by amplification from the genomic DNA, using the vector primers pMS17-F1 and pMS17-R1 which flanked the multiple cloning site region. The resultant PCR products were approximately 1700 bp, matching the expected sequence size of *ram20* in pMS17 (Figure 3.11D). From this, it was confirmed that the resultant strain SfpMS17-ram20 contains the gene *ram20*.

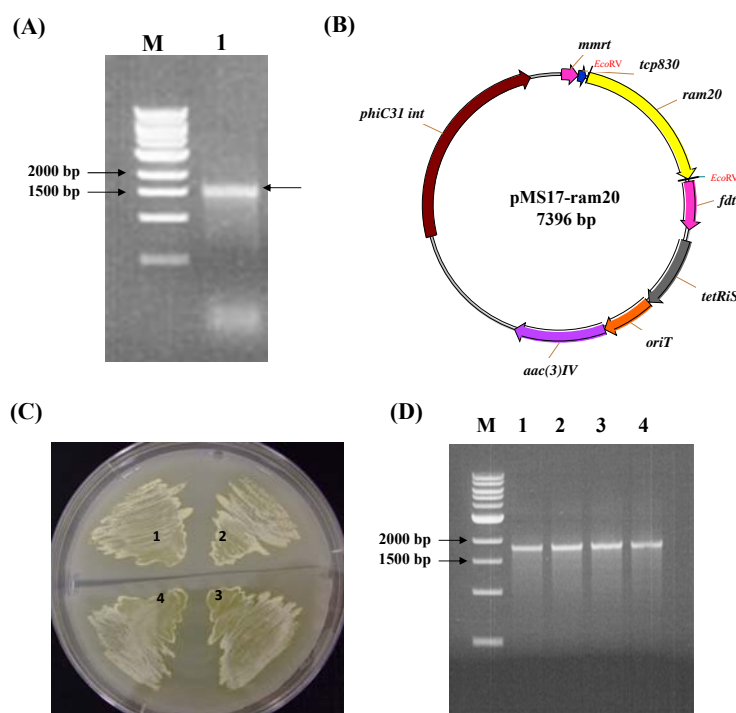


Figure 3.11. PCR amplification and cloning of *ram20* into pMS17, and screening of the SfpMS17-*ram20* transformant strains.

(A) PCR amplification of *ram20*; Lane M: DNA size marker, Lane 1: PCR product (arrow). (B) Schematic representation of the construct pMS17-*ram20*. (C) Selective plate containing apramycin: SfpMS17-*ram20* strains 1-4. (D) PCR amplification of *ram20* from genomic DNA of transformants 1-4. Lane M: DNA size marker, Lanes 1-4: PCR products of SfpMS17-*ram20* strains 1-4.

3.4.5. MS analysis of the Fermentation Broth of the SfpMS17-*ram20* Strain

The culture and induction conditions are described in section 6.19. The crude enduracidin was purified from a 21-day fermentation broth of the SfpMS17-*ram20* strain using a butanol extraction procedure (summarised in Figure 3.3(A) and the details are given in section 6.19). The resultant crude white powder was further purified by RP-HPLC. We found that the addition of a single chlorine to enduracidin does not significantly affect the retention time relative to wild-type enduracidin. This was observed during RP-HPLC analysis where a single enduracidin A peak, when analysed by MS, was found to contain both wild-type and chlorinated enduracidin A. A separate peak for chlorinated enduracidin was never observed. We suspect that the low yield of the modified enduracidin analogues relative to the wild-type, results in it being difficult to separate the two as wild-type absorption saturates this point of the chromatogram. However we were able to determine which fractions contained

wild-type and chlorinated enduracidin after MS analysis. After several rounds of RP-HPLC purification, samples containing enduracidin were concentrated and then directly analysed by MS. We found that SfpMS17-ram20 produces small amounts of chlorinated enduracidin A $[M+2H]^{2+}=1195.1$ which was not found from the extracts of wild-type *S. fungicidicus* strain (Figure 3.12). Unfortunately, a characterisation of the chlorinases involved in ramoplanin and enduracidin biosynthesis was published by Yin *et al.* in 2010 at the time of our own characterisation studies [228]. We therefore did not do further characterisation on our products and changed our direction to produce other chlorinated variants of enduracidin. The strategies for producing different chlorinated enduracidin are given in the discussion (section 3.6) and future work (Chapter 5).

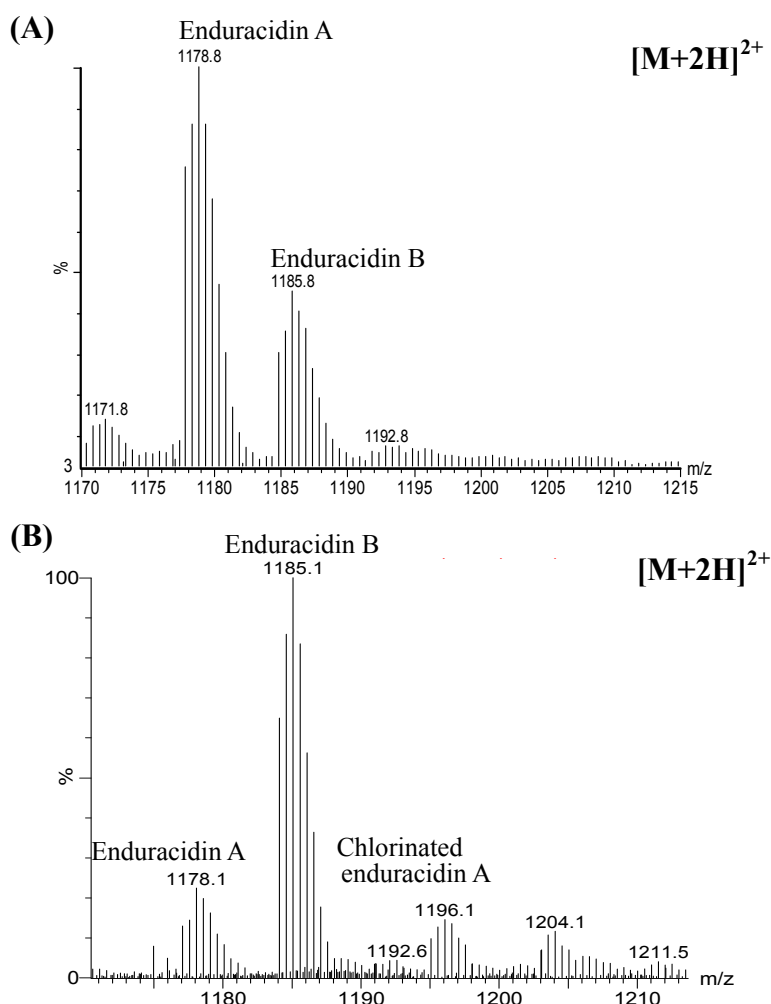


Figure 3.12. MS analysis of SfpMS17-ram20 fermentation products.

(A) Control: the spectrum of wild-type *S. fungicidicus* fermentation products. **(B)** The spectrum of SfpMS17-ram20 extracts shows wild-type enduracidin A ($[M+2H]^{2+}$ observed 1177.1, expected 1177.5) and B ($[M+2H]^{2+}$ observed 1184.1, expected 1184.4) with minor amounts of chlorinated enduracidin A ($[M+2H]^{2+}$ observed 1195.1, expected 1194.5).

Yin *et al.* (2010) showed that the putative chlorinase *end30* from the enduracidin biosynthetic gene cluster was characterised by creating an *end30* knockout strain (SfΔ*end30*) which produced dideschloroenduracidins (2; Figure 3.13) and by restoration of wild-type enduracidin production by complementing SfΔ*end30* with a plasmid-borne *end30* gene (1; Figure 3.13) [228]. From this, they confirmed that *end30* is responsible for the di-chlorination of Hpg-13 in enduracidin. Additionally they over-expressed *ram20* by the constitutive promoter *ermE** in both the mutant SfΔ*end30* and wild-type *S. fungicidicus* and found differently chlorinated enduracidin products. Through characterisation by tandem mass spectrometry, they showed that Ram20 can monochlorinate residues Hpg-13 and Hpg-11 in the mutant SfΔ*end30* and wild-type *S. fungicidicus* respectively, producing monochlorinated and tri-chlorinated products (3 and 4; Figure 3.13).

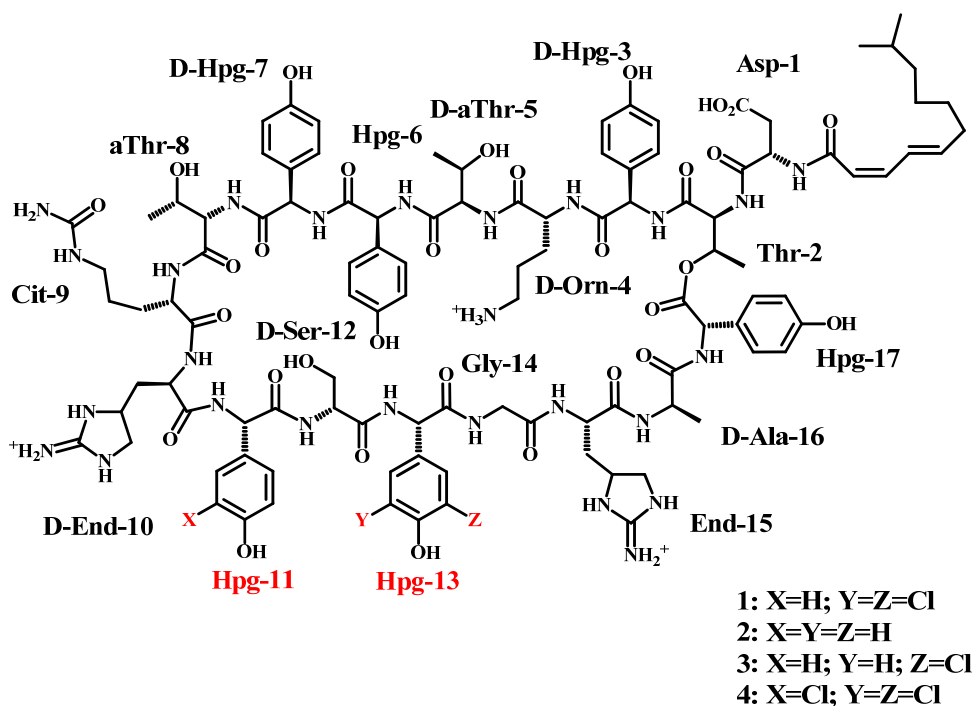


Figure 3.13. Enduracidin A analogues with different halogenation patterns.

3.4.6. Summary

Chlorination is predicted to occur during peptide assembly based on a comparison of the active sites of the Dpg- and Chp-activating A domains with those of the other ten identified Hpg-activating A domains (Table 3.2). Therefore Ram20 was predicted to carry out the chlorination of Hpg-17 when it is tethered to the PCP domain of module 17 of the ramoplanin and enduracidin synthetases.

We successfully produced a transformant *S. fungicidicus* strain SfpMS17-ram20 which can over-express *ram20* under the control of a tetracycline promoter. The MS analysis of SfpMS17-ram20 fermentation extracts shows a new variant of chlorinated enduracidin. However a similar study was published by Yin *et al.*, showing that Ram20 can monochlorinate the Hpg-11 and Hpg-13 residues in enduracidin [228]. Surprisingly, Ram20 was unable to chlorinate Hpg-17 in enduracidin despite this residue being the natural substrate in ramoplanin. A detailed discussion of this is given in section 3.6 and possible approaches for the production of different chlorination patterns in enduracidin are given in Chapter 5.

3.5. Bio-engineering of Enduracidin with Mannosylation by the Mannosyl Transferase (Ram29)

3.5.1. Introduction

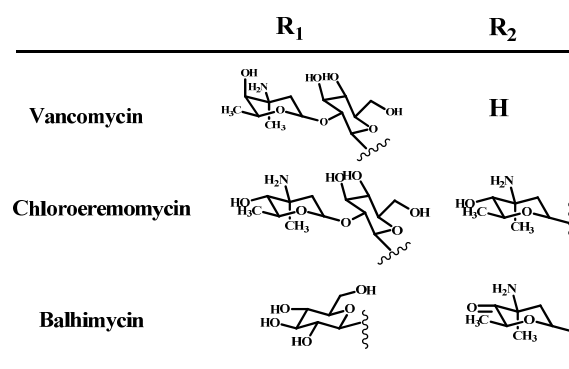
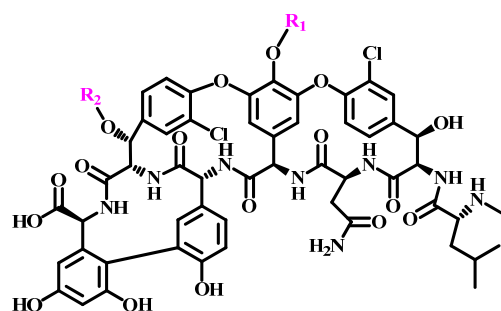
Glycosylation is a post-translational modification carried out by glycosyltransferases (Gtfs). The sugar moieties in glycopeptide antibiotics can increase antibacterial activity by enhancing peptide dimerisation, membrane localisation and structural stability. In addition, the hydroxyl groups on the glycosyl moiety can provide positions for the attachment of other functional groups, such as lipid moieties, leading to further improvements in antibiotic activity. For example in vancomycin (Figure 3.14A) the sugar group enhances peptide dimerisation, resulting in increased affinity towards the bacterial cell wall and thereby increased bactericidal efficacy [226, 229]. For teicoplanin and mannopeptimycin (Figure 3.14B and C), attachment on the sugar moiety of a lipoyl chain and isovaleryl group respectively, enhances peptide anchorage to the membrane, improving antibiotic activity [226, 230]. The dimannose group on ramoplanin contributes to maintain peptide structural stability and protection from acid hydrolysis [231]. However, it has also been found that ramoplanin and its aglycon form share similar potency, indicating that the mannose group is not directly related to antibacterial activity [231].

Different glycosylation patterns can affect peptide activity and mode of action. In vancomycin-type antibiotics, the activity of chloroeremomycin against Vancomycin-resistant *Enterococci* (VRE) is 10-fold higher than that of vancomycin [232]. The only structural difference between the two antibiotics is in the glycosylation pattern (R₂, Figure 3.14A). Vancomycin has the disaccharide D-glucose-L-vancosamine attached to the phenolic group of the fourth residue Hpg. Chloroeremomycin has the disaccharide D-glucose-L-4-*epi*-vancosamine attached at the same residue, and has a second L-4-*epi*-vancosamine attached to the 2-hydroxyl group of the sixth residue, β -OH-Tyr. Moreover, the modification of the disaccharide moiety on vancomycin can alter its mechanism of action. For example, chlorobiphenyl vancomycin has a chlorobiphenyl group on the disaccharide moiety, which changes the original vancomycin target site from the transpeptidation step of peptidoglycan

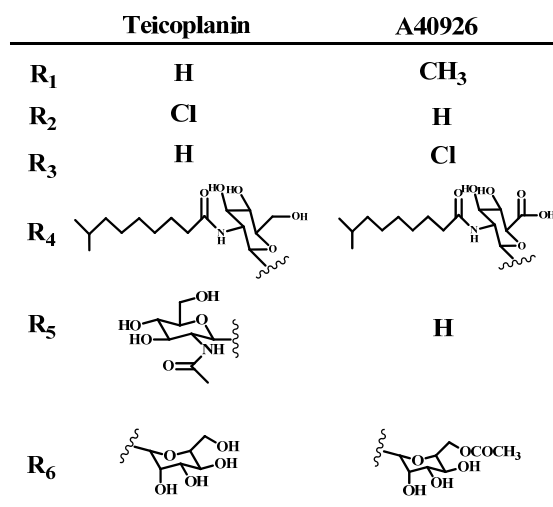
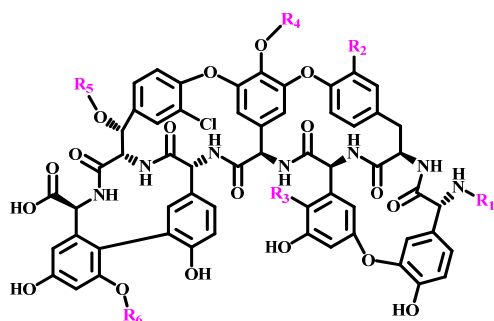
synthesis to the transglycosylation step [31].

Many Gtfs from antibiotic biosynthetic gene clusters such as vancomycin, chloroeremomycin and teicoplanin have been studied extensively. In the case of vancomycin, the enzymes GtfD and GtfE have been proven to act in tandem to produce the *L-epi*-vancosaminyl-1,2-D-glucosyl chain (R_1 of vancomycin in Figure 3.14A) on the Hpg residue [233]. The enzymes GtfB and GtfC, and GtfA from the chloroeremomycin gene cluster were demonstrated to attach the *L-epi*-vancosaminyl-1,2-D-glucosyl chain (R_1 of chloroeremomycin in Figure 3.14A) and the *L-epi*-vancosaminyl group (R_2 of chloroeremomycin in Figure 3.14A) onto Hpg and β -OH-Tyr respectively [234]. In balhimycin biosynthesis, two putative Gtfs, BgtfB and BgtfA, are predicted to transfer glucosyl and dehydrovancosaminyl groups (R_1 and R_2 of balhimycin in Figure 3.14A) onto the Hpg and β -OH-Tyr residues respectively [235]. Additionally, Orfs 1 and 10 from the teicoplanin biosynthetic gene cluster have been confirmed to transfer *N*-acetyl-glucosamine onto the 3-chloro- β -OH-Tyr and Hpg residues (R_5 and R_4 of teicoplanin in Figure 3.14B) respectively [94]. The Dbv9 enzyme from the antibiotic A40926 (precursor to semi-synthetic dalbavancin) biosynthetic gene cluster was predicted to transfer *N*-acetyl-glucosamine onto Hpg (R_4 of A40926 in Figure 3.14B) [236]. However, the transfer of mannosyl groups onto the teicoplanin-type antibiotics (R_6 of Teicoplanin and A40926 in Figure 3.14B), mannopeptimycin and ramoplanin have not yet been characterised. Mannosylation is carried out by mannosyl transferases which are predicted to be membrane proteins and can transfer mannose from polyprenyl phosphomannose. Due to the difficulties in expression of membrane proteins and the poor availability of substrates, this modification is difficult to characterise. Nevertheless we were interested in the mannosylation of ramoplanin, and so attempted to over-express the putative mannosyl transferase Ram29 in *S. fungicidicus* using the optimised *Streptomyces* expression vector pMS17.

(A) Vancomycin-type antibiotics



(B) Teicoplanin-type antibiotics



(C) Mannopectimycin (Epsilon)

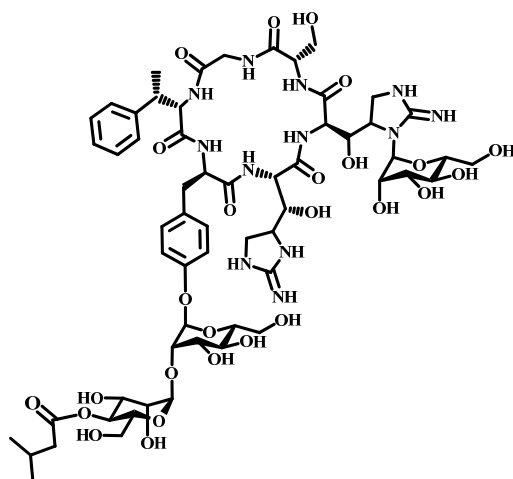


Figure 3.14. Structures of the vancomycin-type, teicoplanin-type, and mannopectimycin (epsilon) antibiotics.

3.5.2. Bioinformatic Analysis of Ram29

The amino acid sequence alignment of Ram29 showed homology to the putative mannosyl transferases Orf3 (62 % identity) and Dbv20 (49 %) from the teicoplanin and A40926 biosynthetic gene clusters (Table 3.3 and Figure 3.15). Therefore we suggest Ram29 is a mannosyl transferase which carries out the mannosylation of ramoplanin.

Table 3.3. Identification of Ram29 function by Blast protein sequence searches

Protein name	Source	Accession number	Identity (%)
Putative mannosyl transferase Orf3	Teicoplanin biosynthetic gene cluster in <i>Actinoplanes teichomyceticus</i>	AAC44360.1	62
Putative mannosyl transferase Dbv20	A40926 biosynthetic gene cluster in <i>Nonomuraea sp.</i> ATCC39727	CAD91215.1	49
Protein mannosyl transferase (VEG22)	uncultured soil bacterium	ACJ60964.1	47

	BAD	AVG	GOOD	
Ram29	1	mdv-prvrppg--a--apapr	rrrrwrfwqspdgppawarpallgiaa	avlyt 49
Orf3	1	mvsetlthp--psqdsapesrphrwa	vwrspagppawarpallgialvaavlya	52
Dbv20	1	mshi-tmtppsacr dpapagr fprwa	vwrspggppwarpallciaataavlya	53
Veg22	1	msa-tlthlplsqdplpegktnrwa	fwrspagppwarpallgialvalllya	53
	50	anlarsgypmyyavavksmsvswpa	fwtgafdpaasitidklagafvpqalsar	103
	53	wmlrlvdyapryadavksmsgswra	flygavdpeatytlldklagsfvpqvvsak	106
	54	wmlplvdyaprysdavksmsenwka	flygtvdvqatytlldklagafvpqaisvk	107
	54	wmlpvvdyaplyseavksmsgswka	flygaidadatytlldklagsfipqvisak	107
	104	vfgfhqwsalalpqaevgviavlvly	ravrwhgpgglaaaglfattpivssmf	157
	107	ifgyhawsalalpqiiegviavlvmyr	vrrwgvvpgllaagiftltpvaasmf	160
	108	ifgfhawalalpqiiegvisvlvmyri	vrrwagvvpgllaavftitpvaasmf	161
	108	ifgysewsalalpqiiegvisvlvmyr	vrrwagvvpgllaagiftltpvaasmf	161
	158	ghsmedgaltlclvlaadafgaavtrg	sparlalagawiglgfqakmmqawlvl	211
	161	ghsmedgalvmclvlavdayqravleg	rlrsligavvwglgfqakmvqawmil	214
	162	ghsmadgalvmclvlavdsyqravleg	rlrslvawgvwglgfqakmlqawmil	215
	162	ghsmhdgfltmclvlavdqyqravseg	rlrslvawgvwglgfqakmlqawmil	215
	212	palvvtylagapvrararvvhvaaava	atlavslwvlatlpgshrpwadgt	265
	215	palavgylltapielrrrrvrlhgaag	vmlavslswmlytltpasdrpyisgt	268
	216	palaigyllsapiglrllrqlhgiagv	vtlvvslwitlyhvtpaadpypisgt	269
	216	palvigyllsapievrrrvkhvavagv	vtlavslswvalyftpaserpvydgs	269
	266	tsgnafamvfyngfdragihvpgaltt	gftdg-----	298
	269	tnnnafamvfynglgrvginlpgae	evngrngvpapggplggpegegrgtvgt	322
	270	tnssaaamvfynglgrlginlpgalp	ppnymgsvigppap---pkrstg	314
	270	tnnsafamvfynglarlginlpgarv	setdvpvldpnpvpp-----	312
	299	-----	-----	298
	323	kaqdqrgtvgtkaqdqrvtvgtkaqd	qrgtvgtkaqdqkgmvppkahgppmgp	376
	315	-----	lprprpvmvipei-----	327
	313	-----	pmidimptirgppwedp	329
	299	-----	gaaaggswtalaadrlatqigwvyp	laltglllglarwt-- 338
	377	galpegyeskmesradpmgwtklfdgh	lgvaigwlfpltlalclglwrrrae	430
	328	-----	giehgggwgklfggrlgvasgwlypl	almallcglwvrrae 369
	330	-----	rkaesrespkgwklfdghlaiavgwly	pltltlvcgllwrrse 375
	339	----aragllfwglwlltaavvlstri	-tiqhnaylavlapplaalaagavqlw	387
	431	rtdpvrgglvmwgvllvtfgfvfsas	-avahtayvaslapaiaavsalgvmfg	483
	370	rtdparggmvmwgvlltfalypsavfvi	phsayvavlapvaalsgigivmfw	423
	376	rtdpvrgglvmwgvlltfglvfsag	-evphtayvaslapavaalsgigivmfw	428
	388	rthrd-gtapwllpavvvvqagwtlwl	atrypsflagltwtapiaavlavvvl	440
	484	rayrrggrqgwilplavavqlawaawl	wshypnflpwarwgtlllgvaalvglv	537
	424	rayrsggrmawifplaiavaelawavl	wsfyptflpwamwgaavalgvavval	477
	429	rayrrggrqawilpavvaaelawgawl	wshypnflpsamwgvialgavaival	482
	441	-----arptarrpavvvvqagllavp	vawgasvlnpryagtsfeagagpsgp	487
	538	larvtrtassalvaaalavgvvmla	apatyavsvldpdysgnsfdanagpasg	591
	478	larlvrrrsslvsagltigvaamla	apatwsavldpryggssfdanagpaar	531
	483	lvrvlknasanlvtaalavgvaa	ilaapatyafsvldpdfagasydanagpesg	536
	488	vgvrldddtdrltpglrrlddylaah	rdgrtylaatsswrtagrliivptghsy	541
	592	t-----	-----	592
	532	t-----	-----	532
	537	ki-----	pgg-----	pesge--mpvgggppgggelgppgsgg 567
	542	lplggfsgaapfplagvqrlvrdgel	ryfvlggpeglggeateayritgwvle	595
	593	-----	-----	592
	533	-----	-----	532
	568	pppgggggmpggg-----	ggmpgggggg-----	592
	596	tcatvppaehgadpdlvtrcdkp	-----	619
	593	-----	-----a-----	593
	533	-----	-----pgg-----	535
	593	---mpggge-----	-----lgpgg-----	603

Figure 3.15. Ram29 amino acid sequence alignment with Orf3, Dbv20 and Veg22. The amino acid sequence of Ram29 was aligned with those of Orf3, Dbv20 and Veg22 from *A. teichomyceticus*, *Nonomurea* sp. ATCC39727 and an uncultured soil bacterium respectively, using the online alignment program T-coffee [201].

3.5.3. Proposed Mannosylation Mechanism

The Ram29 amino acid sequence was analysed by eight different bioinformatics programs which are applied widely to membrane protein analysis, namely: TMPRED [237], HMMTOP [238], TMHMM [239], SCAMPI [240], PRODIV [241], PRO [241], OCTOPUS [242] and TOPCONS [241]. The analyses showed that three-quarters of the Ram29 protein is composed of transmembrane segments (TMS), located away from the C-terminus (Figure 3.16). Although the algorithms did not converge on the precise number of TMS, the results strongly suggest that Ram29 is an integral membrane protein, with a predicted intracytoplasmic N-terminus and extracytoplasmic C-terminus.

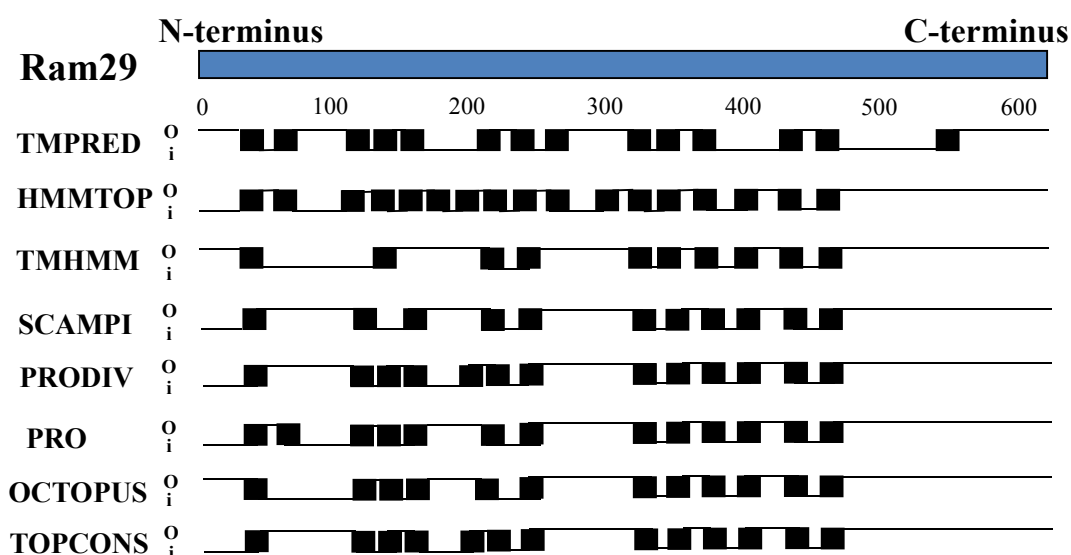


Figure 3.16. Topology prediction of the protein Ram29.

The black boxes correspond to predicted transmembrane α -helices. The predicted locations of the each segment of Ram29 is indicated as intracytoplasmic (i) or extracytoplasmic (o).

Borgh *et al.* (1991) showed that the demannosylated teicoplanin aglycon can be re-mannosylated by incubating with the teicoplanin producing strain, and thus demonstrated that mannosylation occurs on the outside of the cell membrane [243]. Interestingly, this result matches the characteristics of mannosyl transferase, derived from the sequence analysis. For example, the amino acid alignment of mannosyl transferases shows that they have diverse C-termini (Figure 3.15) and the topology

prediction for Ram29 indicates that the C-terminus localises outside of the cell membrane (Figure 3.16). Therefore we proposed that mannosyl transferases would use their diverse C-termini to recognise and bind specific antibiotic aglycons outside of the cell membrane, and then transfer mannose groups from polyprenyl phosphomannoses to form mannosylated antibiotics (Figure 3.17). Polyprenyl phosphomannoses such as the C₃₅- and C₅₀-variants are synthesised by polyprenyl phosphomannose synthases using GDP-mannose [244]. Here we over-expressed *ram29* in *S. fungicidicus* and attempted to characterise its function *in vivo* and *in vitro* using an activity assay towards the enduracidin and ramoplanin aglycons.

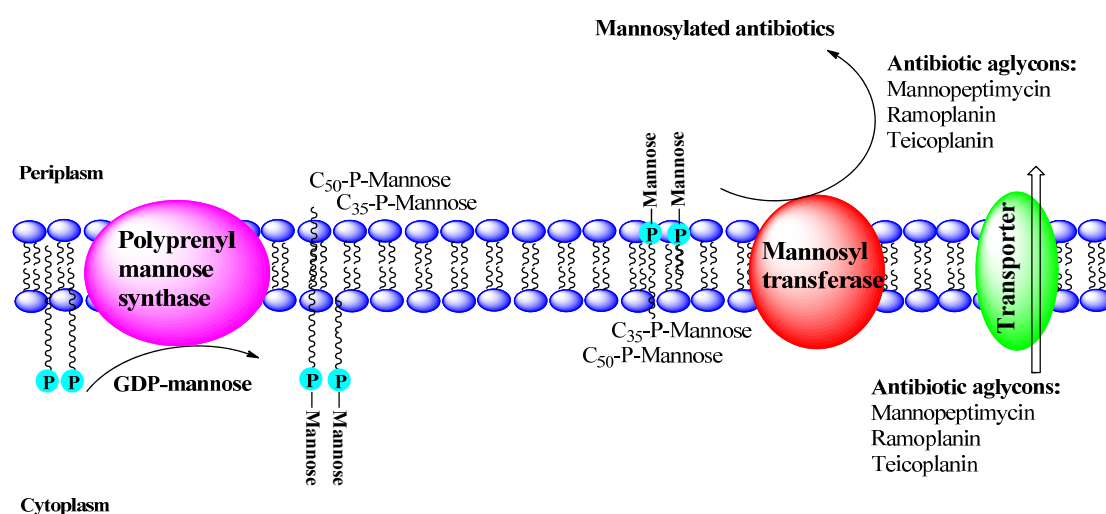


Figure 3.17. A proposed two-component system for mannosylation of the glycopeptides.

Polyprenyl phosphomannose synthase is located in the membrane and catalyses the synthesis of polyprenyl phosphomannoses on the cytoplasmic membrane surface, which then translocate to the periplasmic membrane surface. Subsequently, the membrane protein mannosyl transferase transfers the mannose moiety from polyprenyl phosphomannose to the antibiotic aglycons, which have been transported from the cytoplasm to the periplasm.

3.5.4. In Vivo Studies of the Ram29

3.5.4.1. PCR Amplification and Cloning of *ram29* into pMS17

The *ram29* sequence of approximately 1900 bp was amplified from the *Actinoplanes* strain ATCC33076 genome by the primers sfram29-F2 and sfram29-R1. The forward primer sfram29-F2 incorporated an additional 23 bp from the upstream sequence of

eGFP encompassing the SD sequence and start codon ATG (Figure 3.18A). After digestion with the restriction enzyme *EcoRV*, *ram29* was cloned into the pMS17 vector (Figure 3.18B). The resulting construct pMS17-*ram29* was transformed into *S. fungicidicus* by conjugation and the transformants were selected by the antibiotic apramycin (Figure 3.18C). The presence of the gene *ram29* in the resultant transformant strains was confirmed by PCR. As shown in Figure 3.18D, the PCR products amplified by the vector-specific primers pMS17-F1 and pMS17-R1 have the expected amplification size of approximately 2100 bp, which matches that of *ram29*. From this, it was confirmed that the transformant strains SfpMS17-*ram29* contained the gene *ram29*.

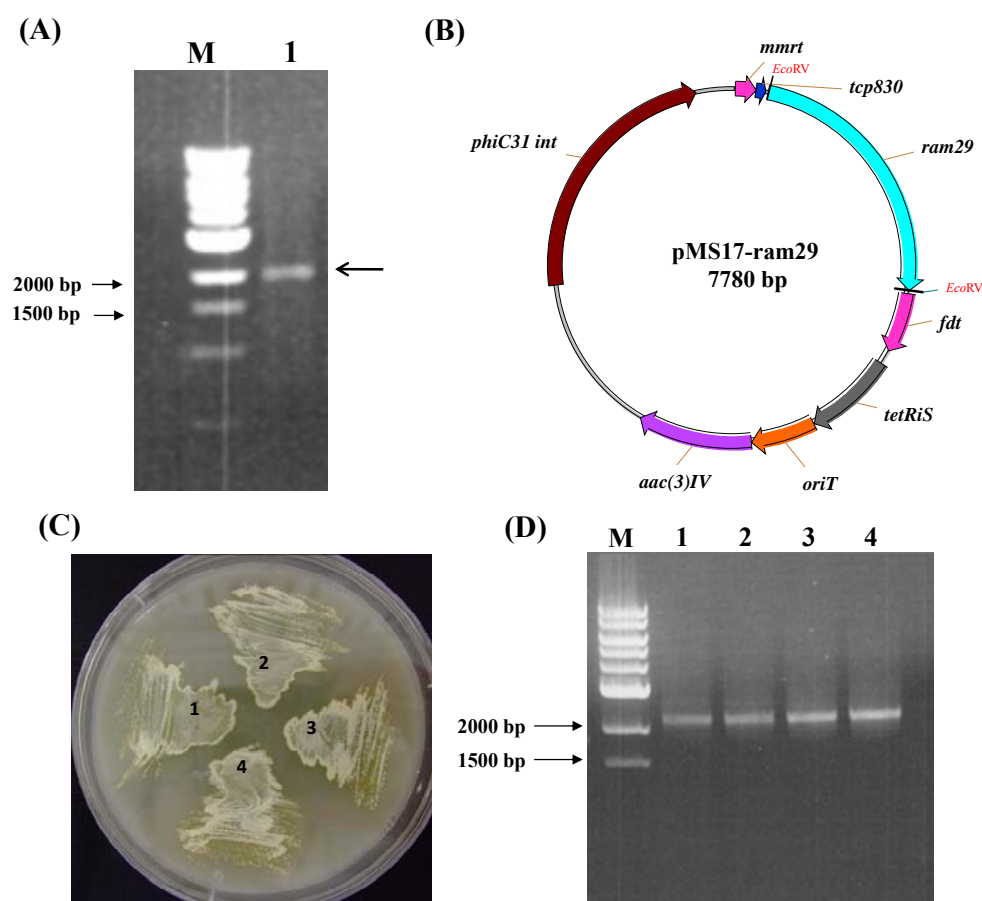


Figure 3.18. PCR amplification and cloning of *ram29* into pMS17, and screening of the SfpMS17-*ram29* transformant strains.

(A) PCR amplification of *ram29*; Lane M: DNA size marker, Lane 1: PCR product (arrow). (B) Schematic representation of the construct pMS17-*ram29*. (C) Selective plate containing apramycin: SfpMS17-*ram29* strains 1-4. (D) PCR amplification of *ram29* from genomic DNA of transformant strains SfpMS17-*ram29* 1-4. Lane M: DNA size marker, Lanes 1-4: PCR products of SfpMS17-*ram29* strains 1-4.

3.5.4.2. MS Analysis of the Fermentation Broth of the SfpMS17-ram29 Strain

The culture and aTC induction conditions are described in section 6.19. Crude enduracidin was purified from a 300 ml 21-day fermentation broth of the SfpMS17-ram29 strain using the butanol extraction procedure (summarised in Figure 3.3 (A) with the details given in section 6.19). The crude white powder was further purified by preparative RP-HPLC. It was difficult to observe any separate product peaks by HPLC chromatogram but we were able to determine which fractions contained the combined wild-type and glycosylated enduracidin after MS analysis. Due to being unable to separate enduracidin and glycosylated enduracidin, and the low yield of glycosylated products, samples containing enduracidin were concentrated from several rounds of RP-HPLC purification and then directly analysed by MS. As shown in Figure 3.19, the SfpMS17-ram29 strain produced predominantly wild-type enduracidin A and B, but with smaller amounts of new products with a mass difference of 162 Da to the wild-type enduracidin. This mass difference is consistent with monoglycosylation, therefore we presumed that these new products are glycosylated enduracidin analogues not observed in the wide-type strain.

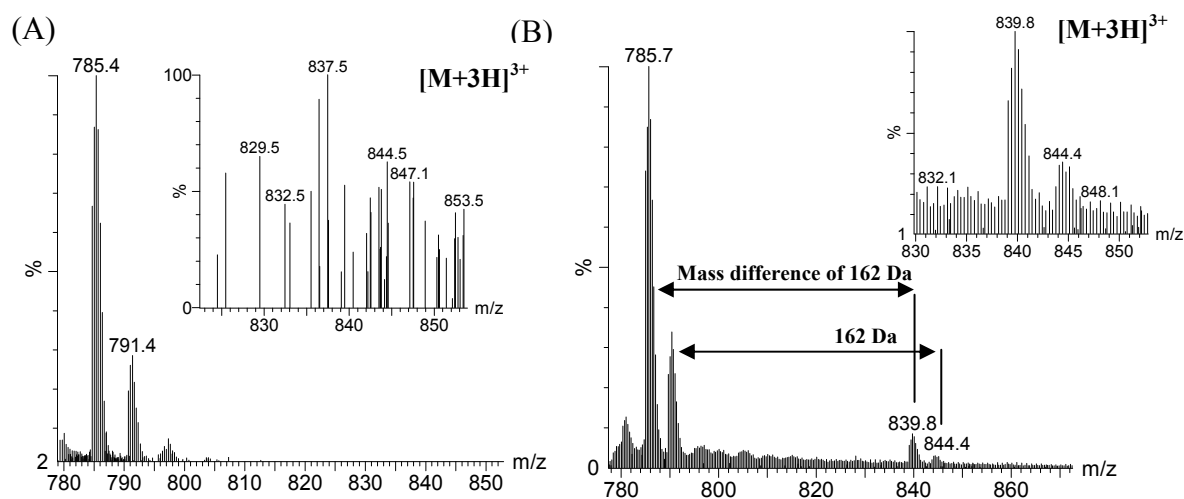


Figure 3.19. MS analysis of the SfpMS17-ram29 fermentation products.

(A) The mass spectrum of wild-type *S. fungicidicus* fermentation products (B) The mass spectrum of SfpMS17-ram29 strain fermentation products shows major amounts of wild-type enduracidin A ($[M+3H]^{3+}$ observed 785.1, expected 785.3) and B ($[M+3H]^{3+}$ observed 790.4, expected 790.4) with smaller amounts of monoglycosylated enduracidin A ($[M+3H]^{3+}$ observed 839.2, expected 839.7) and B ($[M+3H]^{3+}$ observed 844.1, expected 844.4).

Due to the site of mannosylation on Hpg-11 of ramoplanin, we predicted that mannosylation would occur at the corresponding residue in enduracidin. We employed tandem mass spectrometry to perform sequencing of the enduracidin peptide to determine the location of the glycosyl group on the new product. This work was carried out by Matthew Styles (Micklefield group). In the spectra, a prominent y-ion series can be observed for the enduracidin A peptide fragments: End¹⁵-Hpg¹⁷ (y₃, m/z 393.2), Gly¹⁴-Hpg¹⁷ (y₄, m/z 450.2), Hpg¹³-Hpg¹⁷ (y₅, m/z 667.2), Ser¹²-Hpg¹⁷ (y₆, m/z 754.3), Hpg¹¹-Hpg¹⁷ (y₇, m/z 1065.4), Cit⁹-Hpg¹⁷ (y₉, m/z 1376.0) and Thr⁸-Hpg¹⁷ (y₁₀, m/z 1476.7) (Figure 3.20). Analysis of the tandem mass spectrum for the monoglycosylated enduracidin peptide revealed that new fragment ions at m/z 1065.4, 1376.0 and 1476.7 appeared, corresponding to the addition of a single mannose to the y₇, y₉ and y₁₀ fragment ions. As no such additions to the y₃-y₆ fragment ions were observed, we concluded that the site for glycosylation must be Hpg-11 (Figure 3.20 and Figure 3.21).

$$N\text{-Acyl-Asp}^1\text{-Thr}^2\text{-Hpg}^3\text{-Orn}^4\text{-Thr}^5\text{-Hpg}^6\text{-Hpg}^7\text{-Thr}^8\text{-Cit}^9\text{-End}^{10}\text{-Hpg}^{11}\text{-Ser}^{12}\text{-Hpg}^{13}\text{-Gly}^{14}\text{-End}^{15}\text{-Ala}^{16}\text{-Hpg}^{17}$$

Peptide	y fragment	y ₁₀	y ₉	y ₈	y ₇	y ₆	y ₅	y ₄	y ₃	y ₂	y ₁
Monomannosylated Observed		1476.7	1376.0		1065.4	754.3	667.2	450.2	393.2		
Enduracidin A	Expected	1476.5	1375.4	1219.4	1065.3	754.2	667.1	450.2	393.1	239.1	168.0
		1476.7 -	1376.0 -		1065.4 -						
		mannose	mannose		mannose						
		=1315.4	=1214.4		=903.7						
Enduracidin A	Expected	1315.4	1214.4	1057.3	903.7	754.2	667.1	450.2	393.1	239.1	168.0

Figure 3.20. Localisation of the glycosyl group on monoglycosylated- enduracidin A by tandem mass spectrometric analysis.

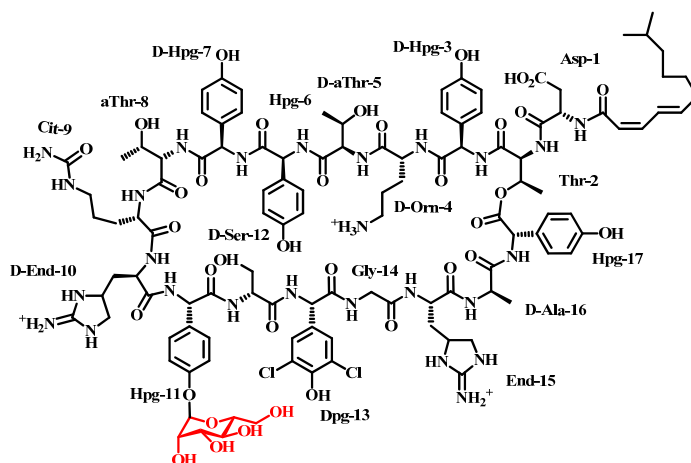


Figure 3.21. Monoglycosylated enduracidin A analogue.

3.5.5. In Vitro Studies of Ram29 Activity

We have confirmed that Ram29 can monoglycosylate enduracidin *in vivo*. Due to the low yield of glycosylated product, we were unable to carry out further characterisation such as antibacterial activity. Glycosylation is a post-translational modification, that is, mannosylation occurs after peptide assembly. Therefore we were able to set up an *in vitro* assay system in order to produce greater quantities of glycosylated products using membrane fractions containing Ram29.

3.5.5.1. Production of the Ramoplanin Aglycon

The demannosylation of glycopeptides can be achieved by chemical or enzymatic methods. Using chemical methods, the Boger group successfully removed glycosyl groups from vancomycin and ramoplanin using anhydrous HF, achieving high yields of peptide aglycons [245]. Using enzymatic methods, Di Palo *et al.* (2007) found that Jack bean α -mannosidase removes single mannose units sequentially from ramoplanin to give first the mannosyl aglycon, then aglycon depending on the length of incubation time [246].

The ramoplanin aglycon was produced by incubating with Jack bean α -mannosidase in phosphate buffer pH 7.0 at 30 °C for 18 h (the detailed procedure is given in section 6.26.3). As shown in Figure 3.22A, the product from the reaction with α -mannosidase has a longer retention time than that of ramoplanin in RP-HPLC. This is likely due to the removal of the hydrophilic mannosyl group, resulting in an enhanced interaction with the stationary phase of the column. The mass of the purified product (peak 2; Figure 3.22A) matched that of the ramoplanin A2 aglycon ($[M+2H]^{2+} = 1115.9$), indicating that the two mannose groups were successfully removed. However a minor amount of residual ramoplanin mannosyl aglycon ($[M+2H]^{2+} = 1197.0$) was still present after the reaction. We also found that the quantity of the aglycon product was less than that of the starting material ramoplanin (250 μ g) by approximately 2-fold after an 18 h reaction. We propose that this phenomenon was caused by the instability of the ramoplanin aglycon together with reduced solubility. Cudic *et al.* (2002) confirmed that the dimannosyl group of ramoplanin aids in maintaining structural

stability and that ramoplanin aglycon is easily decomposed in acidic solution [231]. In the Jack bean α -mannosidase reactions, we found less aglycon product was recovered when using longer reaction times, indicating that the ramoplanin aglycon product was unstable. Additionally, ramoplanin has poor solubility as it easily aggregates and precipitates as a white solid. Therefore when the hydrophilic mannosyl groups were removed, the decreased solubility leads to loss of the ramoplanin aglycon product during purification. After optimisation, we found that an 18 h reaction time resulted in a higher yield of ramoplanin aglycon, with reduced mannosyl aglycons and degradation. The ramoplanin aglycon product was used directly in the Ram29 activity assay.

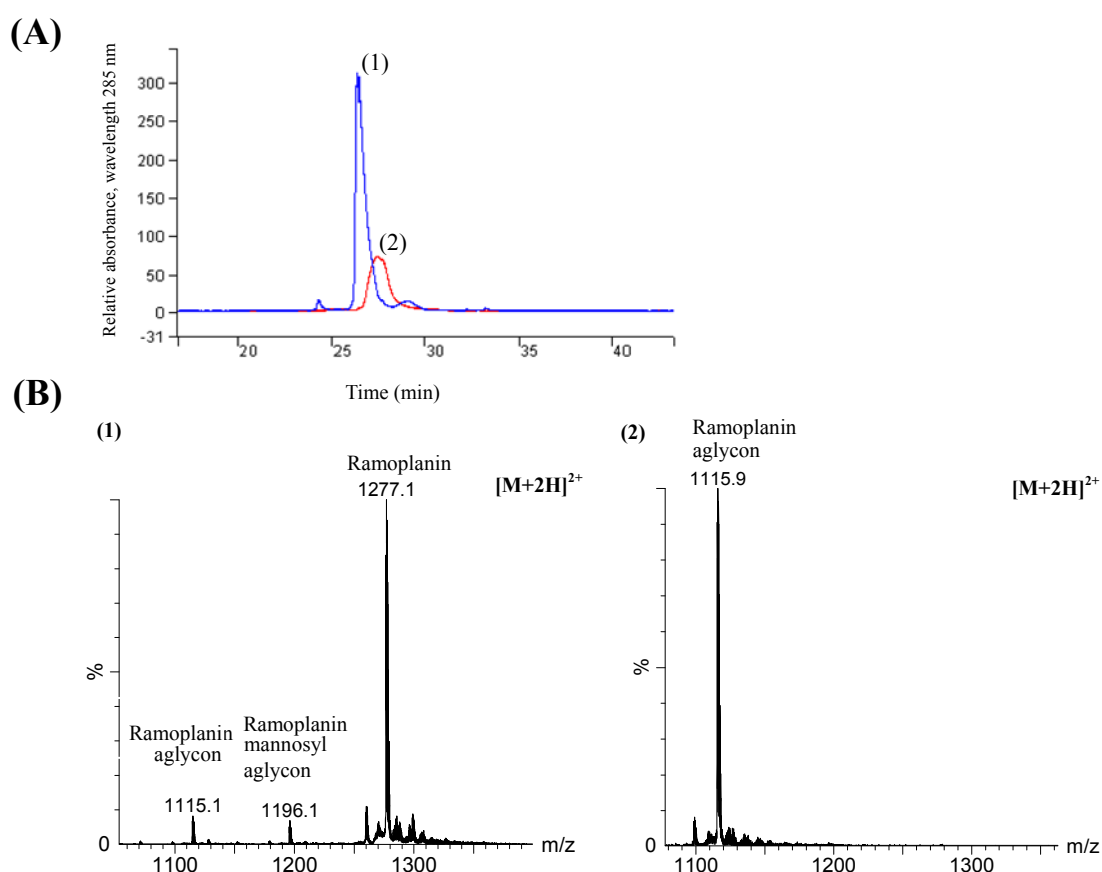


Figure 3.22. Demannosylation of ramoplanin.

(A) HPLC chromatogram of ramoplanin A2 (1) and the product (2) from the reaction with Jack bean α -mannosidase. (B) The mass spectra for the ramoplanin (1) and the product (2).

3.5.5.2. Mannosyl Transferase Assays with Ramoplanin Aglycons

The membrane fractions from the SfpMS17-ram29 strain after aTC induction were prepared for mannosyl transferase assays (the preparation procedures are given in section 6.26.4). Ram29 activity was then determined by the mannosylation of ramoplanin aglycon as determined by mass spectrometry. After a 24 h incubation of ramoplanin aglycon with the membrane protein fraction, mass spectrometric analysis shows production of trace amounts of monomannosylated ramoplanin aglycon ($[M+2H]^{2+} = 1196.8$) and dimannosylated ramoplanin aglycon ($[M+2H]^{2+} = 1277.8$).

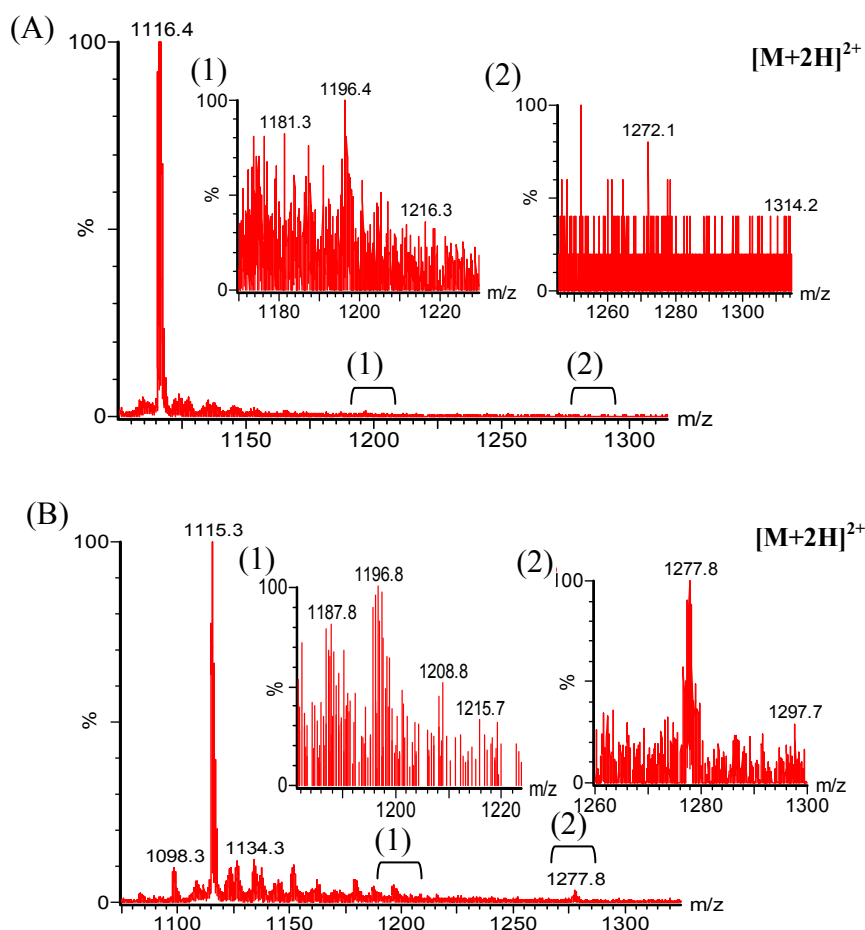


Figure 3.23. The mass spectrum for the product from the reaction of ramoplanin aglycon with the SfpMS17-ram29 membrane fraction.

(A) Control: the product from the incubation of ramoplanin aglycon with boiled SfpMS17-ram29 membrane fraction, (B) Enzyme reaction products: the mass spectrum showed major amounts of ramoplanin aglycon ($[M+2H]^{2+}$ observed 1114.8, expected 1115.0) with smaller amounts of monomannosylated ramoplanin aglycon ($[M+2H]^{2+}$ observed 1196.3, expected 1196.0) and dimannosylated ramoplanin aglycon ($[M+2H]^{2+}$ observed 1277.3, expected 1277.0).

3.5.6. Summary

We successfully obtained an *S. fungicidicus* transformant strain which could over-express Ram29 under the control of a tetracycline promoter. This transformant strain produced monoglycosylated enduracidin during the fermentation process when Ram29 expression was induced. After tandem mass spectrometric analysis, this glycosyl group was found to be attached to the Hpg-11 residue of enduracidin. In an effort to produce more glycosylated products for further characterisation, an *in vitro* assay system was set up.

Ramoplanin aglycon was successfully obtained by incubating with Jack bean α -mannosidase, however there was still a minor amount of ramoplanin mannosyl aglycon present after the reaction. The ramoplanin aglycons were then converted to the monomannosylated and dimannosylated ramoplanin aglycons by using the membrane fraction of SfpMS17-ram29.

3.6. Discussion

The strain *S. fungicidicus* ATCC31731 (from Nogami *et al.* (1984)) supposedly produces 4 g/l of enduracidin [216]. After following the fermentation condition which Nogami *et al.* (1984) published, we only produced 15 mg/l of enduracidin. It is possible that a component of the fermentation medium may have been omitted from the patent. Additionally enduracidin has been found to form amorphous precipitates when they are mixed with peptidoglycans from bacterial cell walls [231]. Therefore we presume enduracidin will form an insoluble membrane complex precipitate when we lyse the cells leading to loss during our purification procedure. This yield is enough for us to identify new variant enduracidin but we are continuing optimisation of our enduracidin fermentation and purification procedures to improve the yield.

After the replacement of the 5'- upstream region of *ram29* with the same region from *eGFP*, their activity could be observed through modification of enduracidin expressed

by *S. fungicidicus*. This demonstrated the importance of the ATG start codon, the sequence of the SD and the distance between the SD sequence and the start codon, as we did not observe any modification of enduracidin when using the original 5'-upstream sequences from *ram29*.

Yin *et al.* (2010) showed that Ram20 chlorinates Hpg-11 or Hpg-13 of enduracidin, rather than Hpg-17 in *S. fungicidicus* at the time we found chlorinated enduracidin from our SfpMS17-ram20 strain. [228]. We did not further characterise our chlorinated enduracidin due to Yin's publication and proposed two possible reasons for the result that they obtained. Firstly, C domains of NRPSs have been demonstrated to act as gatekeepers, regulating the incorporation of specific substrates into the peptide chain [78]. Therefore it is possible that the C domain of module 17 of enduracidin synthetase cannot accept chlorinated Hpg, resulting in only non-chlorinated Hpg-17 being observed in enduracidin. Secondly, the TE domain has also been shown to direct substrate specificity towards the two amino acids at each end of the peptide [82]. It may be that Thr-2 and non-chlorinated Hpg-17 are essential substrates for the formation of the lactone ring of enduracidin; that is, the TE domain of the enduracidin synthetase cannot recognise chlorinated Hpg-17, leading to a failure in the formation of a lactone linkage between Thr-2 and Cl-Hpg-17. This would result in the production of a linear enduracidin peptide which could possibly be targeted for degradation and thus become undetectable.

We have characterised Ram29 as a glycosyl transferase, demonstrating that it can glycosylate Hpg-11 of enduracidin in *S. fungicidicus in vivo*. We found that the yield of glycosylated antibiotics from *in vivo* and *in vitro* experiments was difficult to improve upon, even after trying different protein expression induction times, inducer concentrations and enzyme reaction times. Some *Streptomyces spp.* have been reported to have membrane associated α -mannosidases which remove mannosyl groups from antibiotics. For example, the mycelia of *Streptomyces* GE 91081 can remove one mannose unit from ramoplanin [246], whilst the mycelia of *Nocardia orientalis* NRRL2450 and *Streptomyces candidus* NRRL3218 have been found to remove a mannose unit from teicoplanin [243]. *Streptomyces fungicidicus* contains a putative mannosidase in its enduracidin gene cluster [5], which we presume removes

glycosyl groups from the glycosylated enduracidin and glycosylated ramoplanin aglycons, reaching equilibrium with glycosylation. This may explain why we are able to obtain glycosylated products but cannot improve the yield.

Low substrate solubility may be another factor underlying the low catalytic efficiency observed. It is presumed that the Ram29-catalysed reaction occurs when the substrates (ramoplanin aglycon or enduracidin) associate with the membrane. Ramoplanin and enduracidin have been found to form amorphous precipitates when they are mixed with peptidoglycans from bacterial cell walls [231]. Therefore in the Ram29 assay, the substrates might form insoluble precipitates which would adversely affect the enzymatic reaction. Even when the reaction products (glycosylated ramoplanin or glycosylated enduracidin) were successfully formed, aggregation at the membrane to form an insoluble membrane complex precipitate would lead to failure of product recovery after the reaction.

Using an *in vitro* system, we produced monomannosylated and dimannosylated ramoplanin aglycon. However the monomannosylated ramoplanin aglycon could either be a Ram29 reaction product or the residual product from the α -mannosidase reaction. Nonetheless the dimannosylated ramoplanin aglycon product must stem from Ram29 activity as no dimannosylated ramoplanin was present after the α -mannosidase reaction.

Small amounts of ramoplanin aglycon and ramoplanin mannosyl aglycon can be observed in the mass spectrum (Figure 3.22) when the ramoplanin standard was purified by RP-HPLC and analysed by mass spectrometry. It is possible that the mannosyl group(s) was removed by hydrolysis in the RP-HPLC process because the mobile phase contains 0.1 % formic acid in the buffer system [245]. Thus the analytical procedures still need to be optimised since it may cause our mannosylated products lose sugars again. As shown in Figure 3.23, there are many other new products which have not been observed in the control ramoplanin aglycon, which need to be addressed. Further characterisation of the glycosylated products will be carried out *via* peptide sequencing using tandem mass spectrometry when sufficient product quantities can be obtained. Presently due to the excess amount of membrane-derived phospholipids present in the reaction, the product signals in the tandem mass spectra

are obscured. Therefore further purification of the Ram29 reaction products is required.

In addition to optimising our analytical procedures, we can also optimise our protein expression and assay conditions to increase the yield of glycosylated products by significantly raising the product yield we can compensate any loss during analysis. We attempted to optimise the *in vitro* Ram29 assay conditions using different durations of enzyme reactions ranging from 1 to 7 days. In addition, we attempted to optimise the expression of the protein Ram29 by testing different induction times and inducer concentrations, since membrane proteins are notoriously difficult to over-express. However we did not increase the yield of glycosylated products through these attempts at optimisation. In the proposed mannosylation mechanism, Ram29 transfers mannose from polyprenyl phosphomannose to the antibiotic aglycon (Figure 3.17). If the protein Ram29 was over-expressed on the membrane of *S. fungicidicus*, the rate-limiting factor for Ram29 activity could be the availability of the substrate polyprenyl phosphomannose. Therefore we believe that further investigation into the co-expression of Ram29 and polyprenyl phosphomannose synthase may produce improvements in yield.

Chapter 4. Re-engineering PreQ₁ Riboswitches

4.1. Introduction

In 2004, Breaker and his colleagues found a new riboswitch candidate in the *Bacillus subtilis ykvJKLM* operon by bioinformatic analysis [175]. Subsequently, the genes in the *ykvJKLM* operon were found to be involved in queuosine biosynthesis and were re-named *queC*, *queD*, *queE* and *queF* [247]. The queuosine biosynthetic pathway has been postulated as shown in Figure 4.1; with GTP serving as a starting material. *QueC* and *queF* were demonstrated to be involved in the synthesis of the queuosine precursor 7-cyano-7-deazaguanine (preQ₀) and 7-aminomethyl-7-deazaguanine (preQ₁), respectively [248, 249]. PreQ₁ is then transferred to appropriate tRNAs by a tRNA-guanine transglycosylase (TGT), followed by the addition of an epoxycyclopentandiol ring, whose epoxide group is reduced to produce queuosine *in situ* [170]. Following these findings, in 2007 the Breaker group demonstrated preQ₁ as the ligand for the *queCDEF* operon riboswitch candidate, re-named the preQ₁ riboswitch [170].

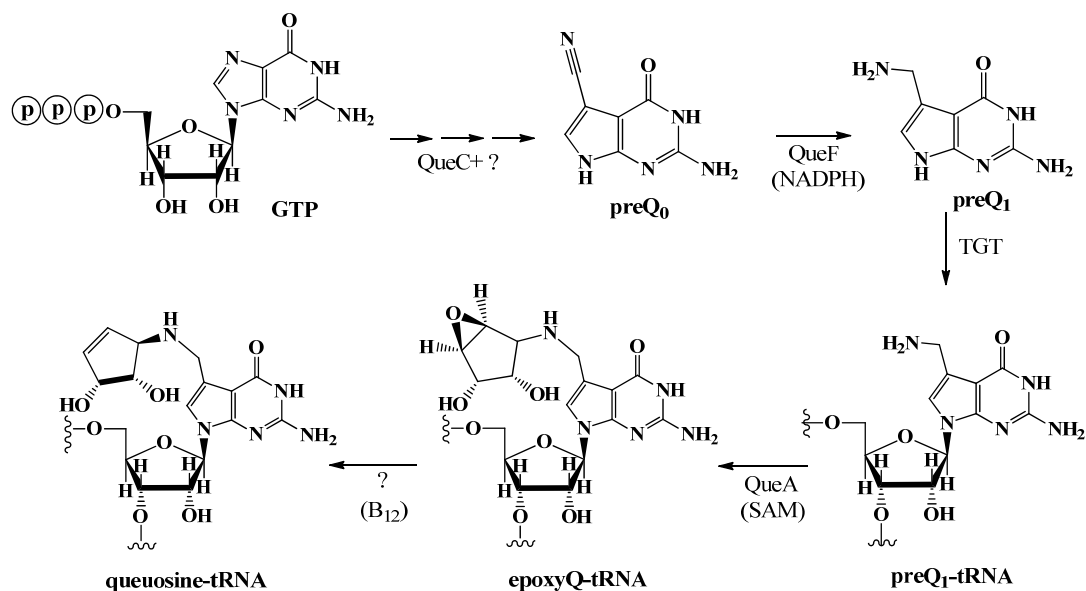


Figure 4.1. The queuosine biosynthetic pathway.

Enzymes known to participate in this pathway are indicated with their cofactors (brackets). Several transformation steps still remain to be characterised and their enzymes are indicated by question marks. Adapted from Roth *et al.* (2007) and Dirk Iwata-Reuyl (2003) [170, 250].

Structurally, the preQ₁ riboswitches can be grouped into two classes, class I and II [170, 251]. Class I preQ₁ riboswitches consist simply of a stem-loop followed by a short single-stranded segment, and can be further classified into two sub-types based on sequence differences (Figure 4.2A). They share a highly conserved cytidine residue (C18 and C17 in sub-type I and II respectively), which is adjacent to the loop-closing base pair (G5-C19 and G5-C18 of stem P1 in sub-type I and II respectively) and have been shown to have the same ligand specificity towards preQ₁ [170]. Class II preQ₁ riboswitches were discovered in the *Streptococcaceae* bacteria and are assumed to control gene expression at the translational level [251]. The aptamer domain is larger, and has a different sequence and structure to that of class I preQ₁ riboswitches (Figure 4.2B). Although it also binds preQ₁, it has a different ligand molecular recognition profile. For example, class I preQ₁ riboswitches can bind both ligands 7-carboxamide-7-deazaguanine and preQ₁ with similar efficiency, whilst class II preQ₁ riboswitches only bind preQ₁.

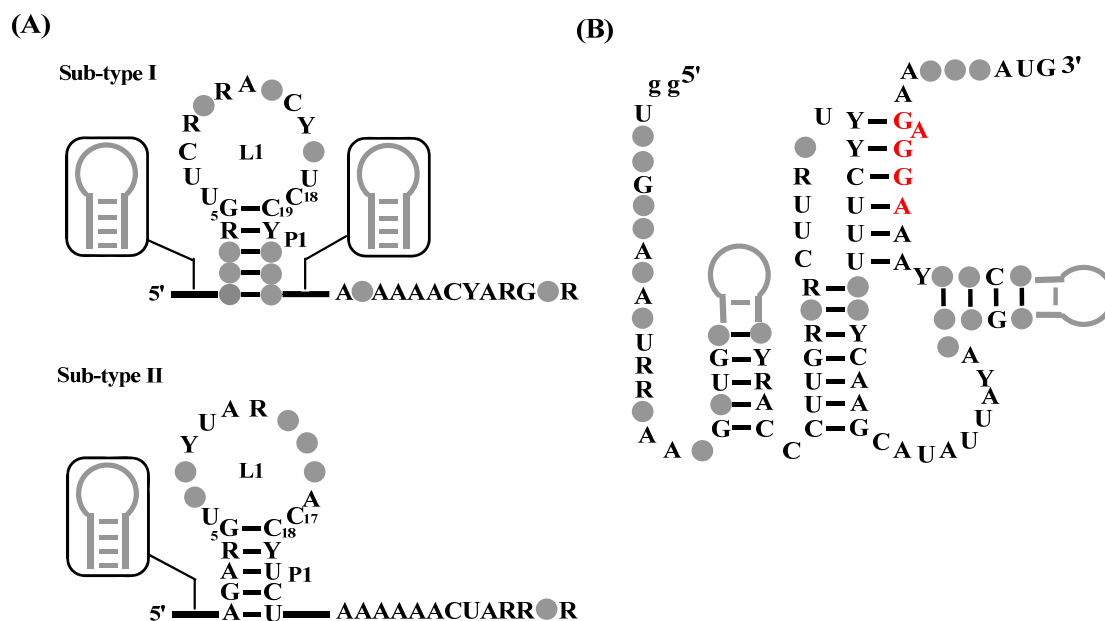


Figure 4.2. PreQ₁ riboswitches.

(A) Class I preQ₁ riboswitches showing the consensus sequence and secondary structure models for the two sub-types. (B) Class II preQ₁ riboswitch showing the consensus sequence and secondary structure model for the RNA motif. The red nucleotides comprise a conserved SD sequence. Less conserved regions are represented by grey circles or heavy black lines and putative stem elements are indicated in grey. R denotes A or G; Y denotes C or U.

The preQ₁ riboswitch from the 5'-UTR of the *B. subtilis* gene *queC* is well characterised and belongs to sub-type II of class I. It contains the smallest aptamer domain (34 nucleotides, nt) discovered to date [170]. Its cocrystal structure with preQ₁ was published in 2009 (3K1V.pdb; Figure 4.3A) and showed that the conserved residue C17 can form a Watson-Crick base pair with the ligand preQ₁ [252] (Figure 4.3C). Additionally, preQ₁ can make a *trans* sugar edge•Watson-Crick pair with A30 and has a hydrogen bond with its N-9 to U6 [252]. Furthermore, the aminomethyl group of preQ₁ can form hydrogen bonds with G5 and the *pro-Rp* phosphate oxygen of G11 [252].

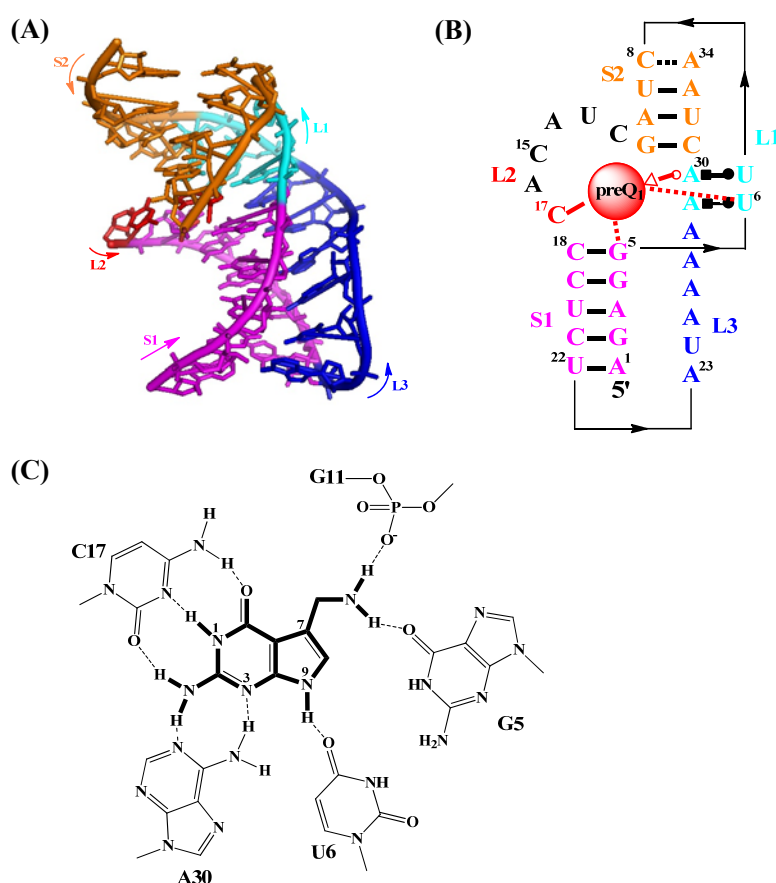


Figure 4.3. The preQ₁ riboswitch crystal structure.

(A) Representation of the tertiary crystal structure of the 34 nt preQ₁ riboswitch aptamer domain in complex with the ligand preQ₁ (3K1V.pdb). (B) Representation of the secondary structure of the 34 nt preQ₁ riboswitch aptamer. Thin arrows denote connectivity; base-pairing symbols follow the nomenclature rules [253]: Watson-Crick pair (—), *Trans* Watson-Crick/Sugar Edge (O→), *Cis* Watson-Crick/Hoogsteen (■—●) and single hydrogen bond interaction (···). L and S stand for loop and stem respectively. (C) The preQ₁ binding pocket. The preQ₁ ligand is shown with covalent bonds in bold, and H-bond interactions as dashed lines.

The *B. subtilis queC* preQ₁ riboswitch is assumed to regulate gene expression at the transcriptional level due to the existence of the transcription terminator [170, 252]. A study of the same class of preQ₁ riboswitch from *Fusobacterium nucleatum* showed that the binding of preQ₁ with C17 stabilises the formation of a pseudoknot, resulting in the formation of a transcriptional terminator to dissociate the RNA polymerase, and therefore repress gene transcription (Figure 4.4) [252, 254]. In the absence of preQ₁ binding, an antiterminator stem is formed at the 3'-end of the preQ₁ riboswitch, allowing the full-length mRNA to be synthesised by RNA polymerase (Figure 4.4) [252, 254].

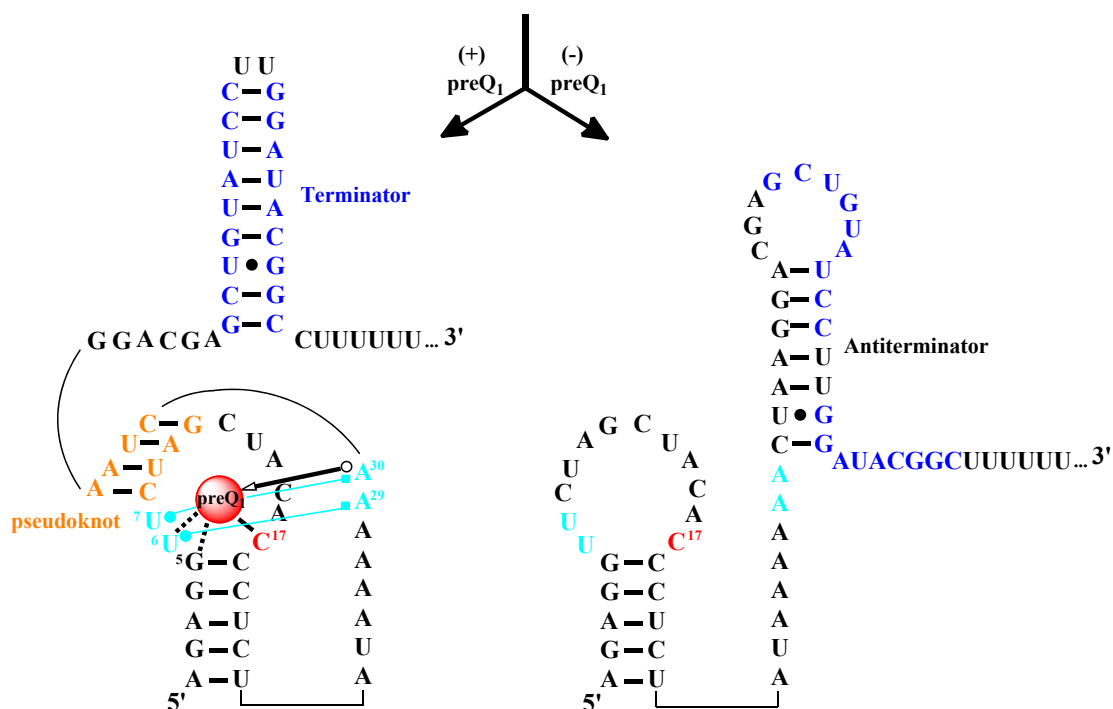


Figure 4.4. The preQ₁ riboswitch gene regulation mechanism.

In the presence of preQ₁, a pseudoknot is formed, allowing the formation of a transcription terminator. The conserved residue C17 is labelled in red, the pseudoknot is labelled in orange, the terminator is labelled in blue. Thin lines denote connectivity; base-pairing symbols follow the nomenclature rules [253], described in Figure 4.3.

In this study, we attempt to engineer the preQ₁ riboswitch as a novel tool for gene expression regulation, by structural mutagenesis of proximal and distal ligand-binding sites and synthesis of alternative ligands that bind to this engineered riboswitch. We plan to make conditional mutant bacteria, which can express genes without responding to the native ligands, but which repress gene expression in the presence of the small molecules that we have synthesised.

4.2. Mutational Analysis of PreQ₁ Riboswitches

The Watson-Crick base pair between the conserved cytidine and preQ₁ is involved in preQ₁ ligand recognition. In “In-line probing” analyses, a significant structural change was observed when C17 was replaced with uridine, which resulted in the change of its ligand selectivity from preQ₁ towards 2,6-diaminopurine [170]. This is very similar to guanine riboswitches, whose ligand selectivity is also based on a Watson-Crick base pair with a conserved cytidine residue [167, 255]. Therefore we are going to engineer the preQ₁ riboswitch by swapping the conserved cytidine to uridine, and by mutating the other ligand-binding residues to change its specificity towards the alternative ligands we have synthesised. Several mutation sites are taken into consideration, such as residues G5, U6 and C17 in the preQ₁ binding pocket (Figure 4.3C) and also residues C18 and A29 which have hydrogen bond interactions with G5 and U6 respectively (A29 has a hydrogen bond with U6 when the aptamer binds with preQ₁, Figure 4.3B).

4.2.1. Construction of PreQ₁ Riboswitch Mutants with the Reporter Gene *lacZ*

The vector pDG1661 containing the preQ₁ riboswitch construct amplified from the upstream region of *queC* from *B. subtilis* 1A40 was kindly supplied by the Breaker group [170]. To generate sequence variants, two flanking master primers (pRSW109f and pRSW110r) that mark the 5'-ends of both strands, and two internal primers that introduce the mutations of interest, were designed to amplify the 5'-fragment and 3'-fragment of the preQ₁ riboswitch (Figure 4.5). Then an overlapping PCR was employed to join both fragments together to get the complete mutant preQ₁ riboswitch sequence, which was cloned into the *EcoRI* and *BamHI* sites of pDG1661, upstream of the reporter gene of *lacZ*. The resulting plasmids were verified by sequencing and transformed into *B. subtilis*. Due to the existence of flanking sequences from *amyE*, the pDG1661 constructs can be integrated into the *amyE* locus of the chromosome by homologous recombination. The following mutant preQ₁ riboswitches were created: the single mutants: G5A, U6C, C17U and C18U; the double mutants: (C17U, C18U),

(G5A, C18U) and (U6C, C17U); the triple mutants (G5A, C17U, C18U) and (U6C, C17U, A29G); and the quadruple mutant (G5A, U6C, C17U, C18U).

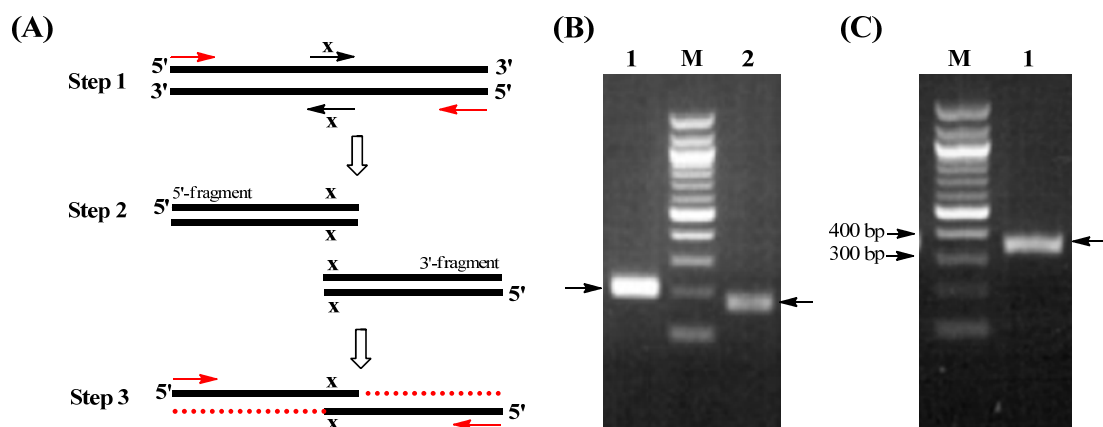


Figure 4.5. Introducing mutation sites into the preQ₁ riboswitch sequence.

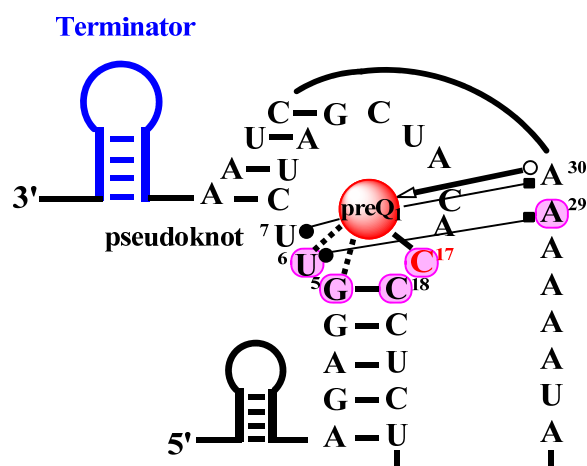
(A) Steps for introducing mutations by overlapping PCR. Two flanking master primers are labelled as red arrows and two internal primers with a mutation site ‘X’ are labelled as black arrows. (B) The products of step 1 PCR, Lane 1: 5'-fragment; Lane 2: 3'-fragment; Lane M: 100 bp ladder. (C) The product of Step 3 PCR, the preQ₁ riboswitch sequence with mutations. Lane M: 100 bp ladder; Lane 1: mutated preQ₁ riboswitch gene fragment.

4.2.2. The Effects of Variant PreQ₁ Riboswitches on the Expression of *lacZ*

The *Bacillus* transformants including the wild-type and mutants of the preQ₁-*lacZ* constructs were cultured in Nutrient Broth at 37 °C without any supplement of ligands. Cells were lysed when cell density reaches adequate OD₅₉₅, following by measuring β-galactosidase activity (OD₄₀₅). The details of assay procedures and the definition of miller unit were given in section 6.27. The wild-type transformant showed low levels of β-galactosidase activity (below 10 miller units), this suggests that the preQ₁ riboswitch acts to repress *lacZ* expression due to the presence of naturally produced preQ₁ ligand (Figure 4.6). In contrast, the mutant transformants have different *lacZ* expression profiles, presumably since they have different levels of binding affinity towards the native ligand or other cellular metabolites. The *lacZ* expression levels are presented as “derepression ratios” and provide a direct correlation between aptamer function and genetic control.

The double mutant (C17U, C18U) strain showed the highest derepression level amongst the ten mutant strains, which was then defined as 100 %. The quadruple mutant (G5A, U6C, C17U, C18U) also presents a nearly 100 % derepression ratio. It suggests that the ligand binding pocket of both mutants is completely disrupted, resulting in high expression of the reporter gene. Among the four single mutant strains, we found that the U6C mutant strain has the highest derepression level of about 85 %, which indicates that this residue plays an important role in ligand binding or aptamer structure maintenance. Although residue C17 has the most important role in ligand recognition, the C17U mutant strain only showed approximately 70 % derepression, the same as that of the single mutant strain C18U and the triple mutant strain (U6C, C17U, A29G). When the C17U mutation is coupled with the mutation of U6C, the derepression ratio is no higher than that of the single U6C mutant. It suggests that there are other factors such as folding kinetics can help single and double mutants of the preQ₁ riboswitch repress the reporter gene expression. Interestingly, the single mutant G5A only showed 34 % derepression, even though this mutation causes the disruption of stem P1. Comparison of the derepression levels observed for the single mutation strains G5A and C18U, suggests that the C18 residue plays a more important role than that of the G5 residue in maintaining preQ₁ function. The compensatory mutation (G5A, C18U) was designed to restore stem integrity, but still elicited an approximately 25 % derepression level, and also could not repress *lacZ* expression to the same extent as that of the wild-type strain. This is probably because of the strong bias in the nucleotide identities at this position (Figure 4.2A). Additionally the derepression ratio of the C17U mutation dropped from 74 % to 38 % when it was present in the triple mutant strain (G5A, C17U, C18U). Taken together, these data indicate that not only the ligand recognition residue (C17) is important for regulation of gene expression, but also the U6 residue plays an important role in functionality. Other factors such as folding kinetics are probably involved in the regulation mechanism since most mutants did not show the full 100 % derepression observed with the double mutant (C17U, C18U).

(A)



(B)

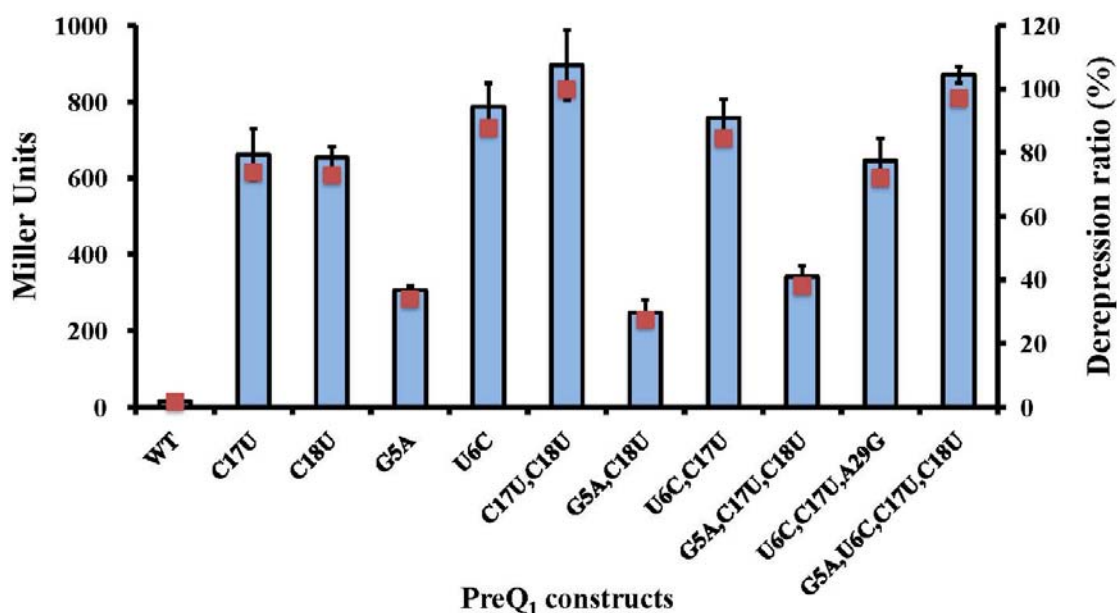


Figure 4.6. The *lacZ* expression profiles of variant preQ₁ riboswitches.

(A) Sequence of the wild-type preQ₁ riboswitch construct. The mutation sites are labelled as pink colored boxes. The hydrogen bond interactions between preQ₁ binding residues and ligand preQ₁ are presented as described in Figure 4.3. (B) The regulation of β -galactosidase activity by the wild-type preQ₁ riboswitch and by the preQ₁ riboswitch mutants. The β -galactosidase activities are presented as Miller Units (left axis indicated by the blue bars) and the error bars represent the mean \pm SD of four replicates. The derepression ratios (right axis indicated as brown squares) are calculated as a % of the expression level observed for the C17U, C18U double mutant strain.

4.3. The Molecular Recognition Determinants for the PreQ₁ Riboswitch

4.3.1. PreQ₁ Riboswitch Interactions With Guanine-Derived Molecules

A more detailed understanding of the molecular contacts between the preQ₁ ligand and riboswitch was obtained when Roth *et al.* (2007) determined the apparent *K_d* values for a series of purine analogues [170] (Figure 4.7). They found that the preQ₁ riboswitch binds preQ₁ and preQ₀ with a *K_d* of 20 nM and 100 nM respectively [170]. The 5-fold reduction of affinity for preQ₀ was due to the structural difference of a nitrile group attached to the atom C-7. The *sp*-hybridized nitrile group of preQ₀ cannot donate a hydrogen bond and must remain linear, whilst the *sp*³ C-NH₂ hybridization of preQ₁ at the aminomethyl moiety allows it to freely rotate and donate a hydrogen bond to the O-6 of G5. Interestingly the binding affinity for the ligand 7-carboxamide-7-deazaguanine was similar to that of the preQ₁ ligand, indicating that some steric tolerance exists at the methylene group of the C-7 position.

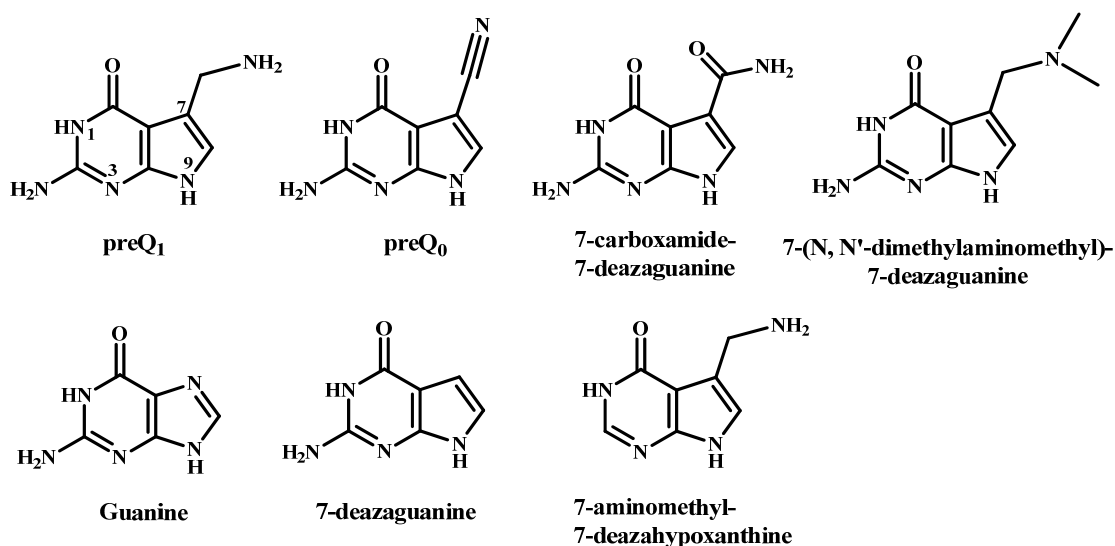


Figure 4.7. The purine analogues tested by Roth *et al.* [170] for preQ₁ riboswitch binding affinity.

Guanine showed a 25-fold lower binding affinity towards the preQ₁ riboswitch than that of the ligand preQ₁ [170]. Comparing the molecular structures of guanine and preQ₁ (Figure 4.7), the absence of an aminomethyl group and a nitrogen atom

substitution at the 7 position in the guanine structure seems to contribute to the reduction of affinity for the preQ₁ riboswitch. The ligand 7-deazaguanine, which is different from guanine only by a carbon atom substitution at the 7 position (Figure 4.7), does not improve the binding affinity, which is even lower than that of guanine [170]. This indicates that the aminomethyl group is the major molecular recognition determinant, rather than the C-7 atom. Additionally the placement of two methyl groups on the nitrogen of the aminomethyl moiety (7-(N, N'-dimethylaminomethyl)-7-deazaguanine, Figure 4.7) decreases the binding affinity to the same level as that of the 7-deazaguanine. This implies that steric interference or the absence of hydrogen-bonding are responsible for this poor affinity.

In addition to the importance of the aminomethyl group, the amino group at the 2 position of preQ₁ is also a crucial determinant of ligand recognition by the preQ₁ riboswitch. It can be demonstrated by comparing the riboswitch affinity of preQ₁ and 7-aminomethyl-7-deazahypoxanthine, whose only structural difference from preQ₁ is the absence of the 2-amino group (Figure 4.7). The binding affinity for 7-aminomethyl-7-deazahypoxanthine is at least four orders of magnitude lower than that of preQ₁ [170].

Here we examine the effect of the ligand preQ₀ on our ten preQ₁ riboswitch mutant strains. We expected that LacZ repression would be observed for the two single mutation strains (G5A and U6C), and the double mutation strain (G5A, C18U), due to the Watson-Crick base pair forming cytidine residue being preserved in these three constructs. However we did not observe any repression phenomena for all 10 mutants when cells were grown in medium with 2.5 mM of preQ₀ (Figure 4.8). It is reasonable that no repression was found for the U6C mutant, which showed a high derepression ratio (see Figure 4.6) indicating that the U6 residue plays an important role in the preQ₁ riboswitch. Furthermore preQ₀ could not repress the expression of *lacZ* to wild-type levels in the G5A mutant and double mutant (G5A, C18U) strains respectively, even though the nitrile group of preQ₀ does not have a direct hydrogen bond interaction with the residue G5. This indicates that the G5 residue plays a distinct role in the riboswitch regulatory mechanism, in addition to aiding the binding of preQ₁.

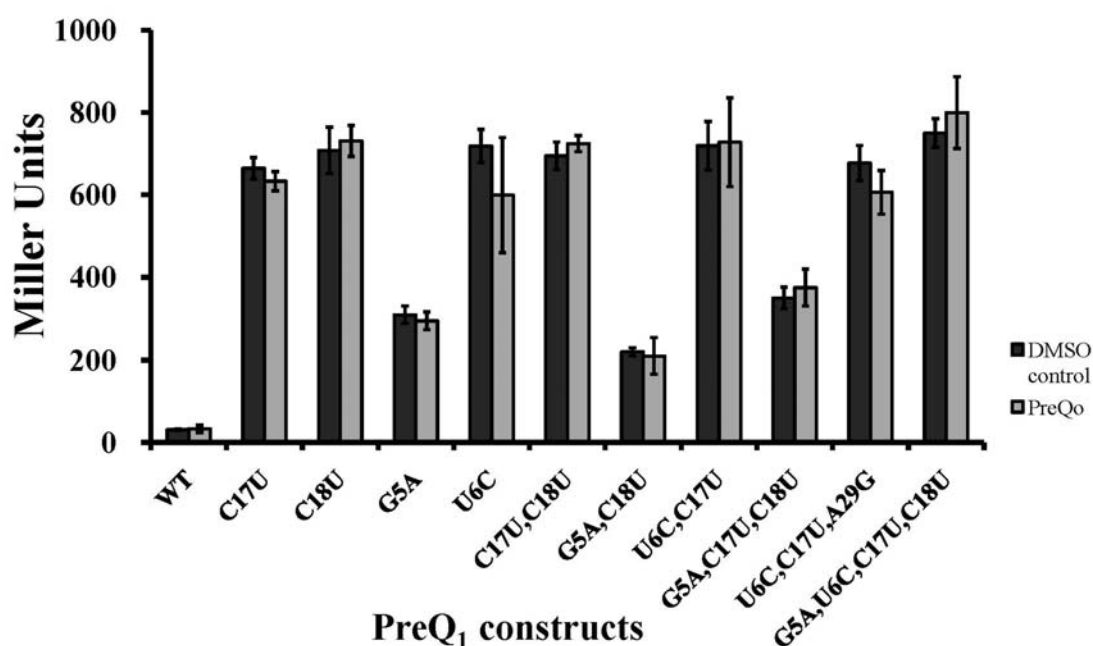


Figure 4.8. The effect of the ligand preQ₀ on *lacZ* expression controlled by the variant preQ₁ riboswitches.

A solution of preQ₀ in DMSO was administered to a final concentration of 2.5 mM, with control cells receiving DMSO alone. The β -galactosidase activities are presented as Miller Units and the error bars represent the mean \pm SD of three replicates.

4.3.2. Testing PreQ₁ Analogues With the Variant PreQ₁ Constructs

A comparison of the ligand binding affinities of preQ₁ and 7-aminomethyl-7-deazahypoxanthine, shows that the exocyclic amine at the 2 position of preQ₁ is an important determinant for ligand recognition [170]. Additionally Roth *et al.* (2007) showed that 2,6-diaminopurine, which has 2-amino moiety, can cause structural modulation of a C17U mutant *in vitro* with a K_d of about 300 μ M [170]. However adenine, which does not have a 2-amino moiety, cannot induce structural modulation in this mutant, emphasising the importance of this moiety. Therefore the ligands we designed and synthesised for the mutant preQ₁ riboswitches included this moiety (The ligand synthesis was carried out by Phillip Lowe (Micklefield group)). These synthetic ligands were: 5-(aminomethyl)furo[2,3-d]pyrimidine-2,4-diamine (**D6**), (2,4-diaminofuro[2,3-d]pyrimidin-5-yl)methanol (**D8**) and 2,4-diamino-7H-pyrrolo[2,3-d]pyrimidine-5-carbonitrile (**D9**), which are shown in Figure 4.9A. As an example, the binding mode for the preQ₁ analogues with the preQ₁

riboswitch mutants is illustrated using the ligand D8 with the preQ₁ riboswitch triple mutant (C17U, U6C, G5A) in Figure 4.9B. The ligand D8 can form a Watson-Crick base pair with the C17U residue. Additionally the O atom at position 9 and the hydroxymethyl group on the C-7 atom can provide hydrogen bonding contacts to U6C and G5A respectively, to maintain ligand binding for this triple mutant.

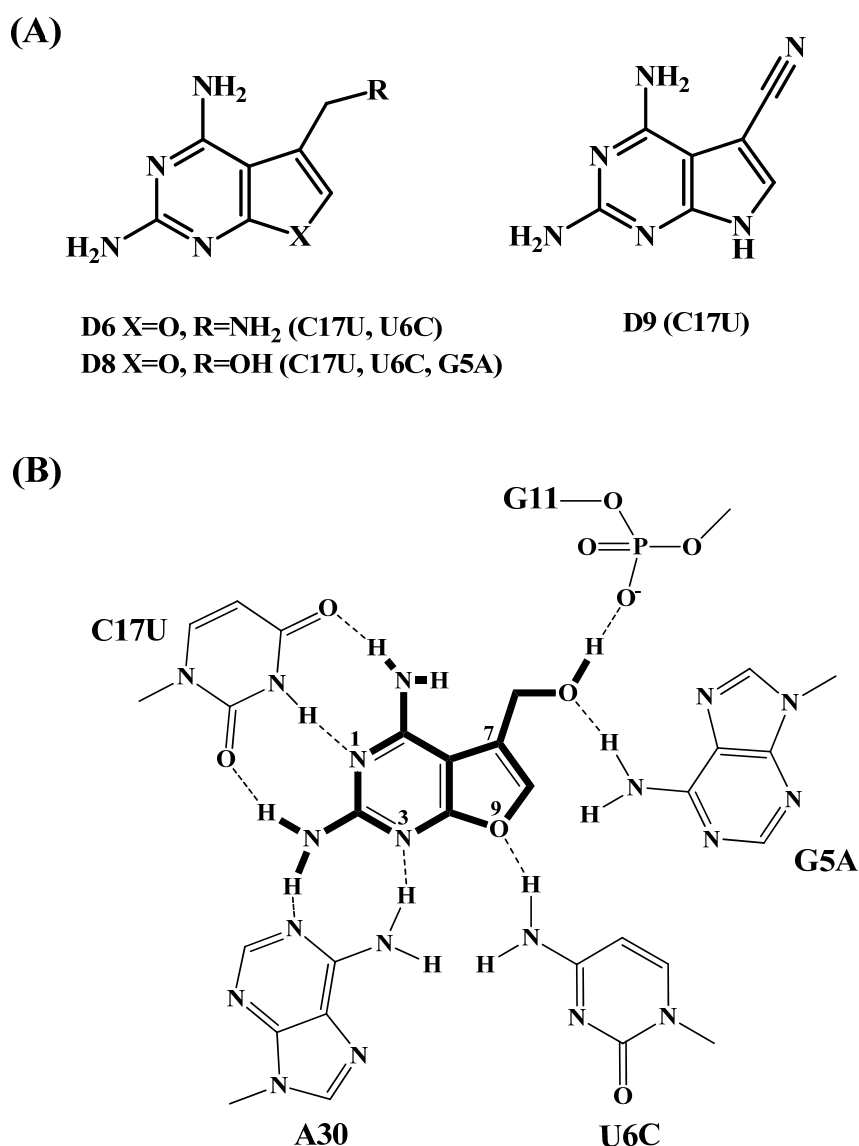


Figure 4.9. Synthetic analogues of preQ₁ and preQ₀.

(A) Analogues with complementary aptamer domain mutations in brackets. (B) The proposed binding mode of analogue D8 to a preQ₁ riboswitch triple mutant (C17U, U6C, G5A).

Due to the observed structural modulation of the C17U mutant with 2,6-diaminopurine *in vitro* [170], we tested 2,6-diaminopurine first with our ten mutants. However we did not observe any repression of *lacZ* expression amongst the

ten mutants, even at 2.5 mM concentration (Figure 4.10). The ligand 2,6-diaminopurine and guanine, which have similar structures and both lack the 7-deaza-7-aminomethyl moiety, can form Watson-Crick base-pairing interactions with the C17U and C17/C17U residues of preQ₁ riboswitch respectively. Interestingly the K_d for 2,6-diaminopurine with the C17U mutant is much lower (600-fold) than that for guanine with the wild-type preQ₁ riboswitch [170]. Therefore it seems that the lower binding affinity of 2,6-diaminopurine for the mutant riboswitches may be responsible for the lack of repression observed (Figure 4.10).

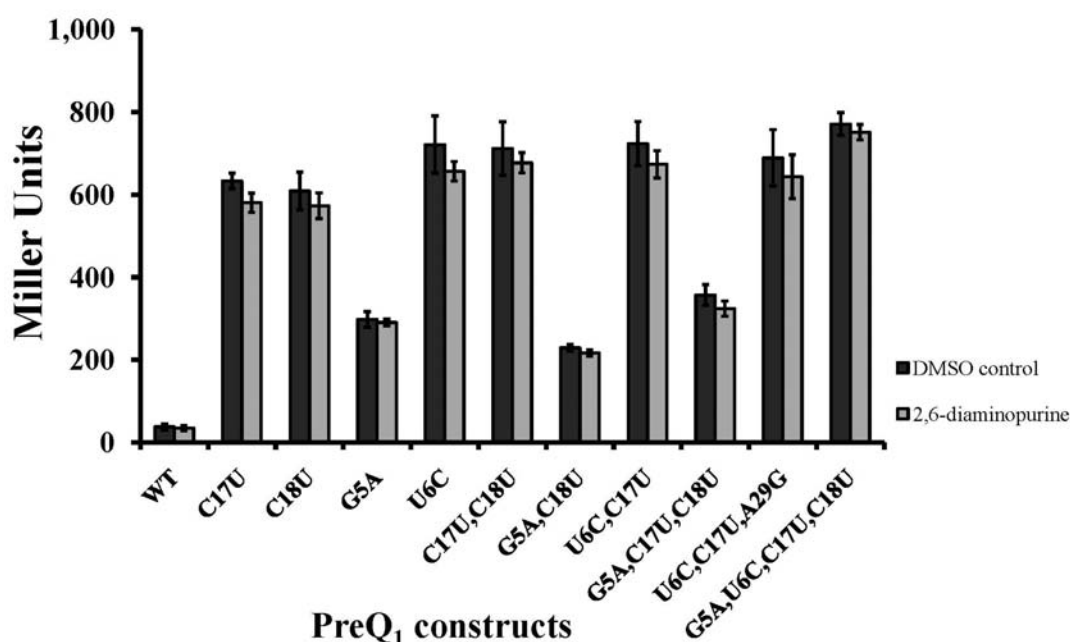


Figure 4.10. The effect of the ligand 2,6-diaminopurine on *lacZ* expression controlled by variant preQ₁ riboswitches.

A solution of 2,6-diaminopurine in DMSO was administered to a final concentration of 2.5 mM, with control cells receiving DMSO alone. The β -galactosidase activities are presented as Miller Units and the error bars represent the mean \pm SD of three replicates.

Subsequently the other three synthetic preQ₁ analogue ligands (D6, D8 and D9) were tested with the ten mutant riboswitch constructs. We expected that the double mutation (C17U, U6C) would favour binding of the analogue **D6**, whilst the triple mutant (C17U, U6C, G5A) would favour binding of the analogue **D8** (Figure 4.9). Additionally the single mutation C17U was predicted to bind the analogue **D9** (Figure 4.9).

The ligand D6 was examined with the ten riboswitch mutants under normal bacterial

growth conditions at 37 °C. As shown in Figure 4.11, the *lacZ* expression for the ten mutants was not repressed, even using 2.5 mM concentrations. However a small degree of repression with the C17U mutant was found when we reduced the growth rate of the bacteria by using a 30 °C incubation temperature and nutrient broth (which has less nutrient content compared to LB or 2YT broth). Interestingly the repression levels could be increased to 59 % when the cells were grown at the even lower temperature of 16 °C (Figure 4.12, Table 4.1). At this temperature repression was observed in a dose-dependent manner ranging from 27 % to 59 % between the concentrations 500 μ M to 2000 μ M of D6 (Table 4.1). Additionally the repression seems to be temperature-dependent, for example the repression at 22 °C was reduced to 19 % and 44 % at D6 concentrations of 500 μ M and 2000 μ M respectively (Table 4.1). Due to the proposed D6 binding model, the double mutant (C17U, U6C) was predicted to be the ideal strain for response to the D6 ligand. Therefore this strain was also grown with the ligand D6 at 16 °C and 22 °C to test whether repression can be observed at these lower temperatures, when we could not observe any repression at 30 °C. Encouragingly, repression was also observed in a temperature and dose-dependent manner, but lower than that observed with the C17U mutant (Figure 4.12, Table 4.1).

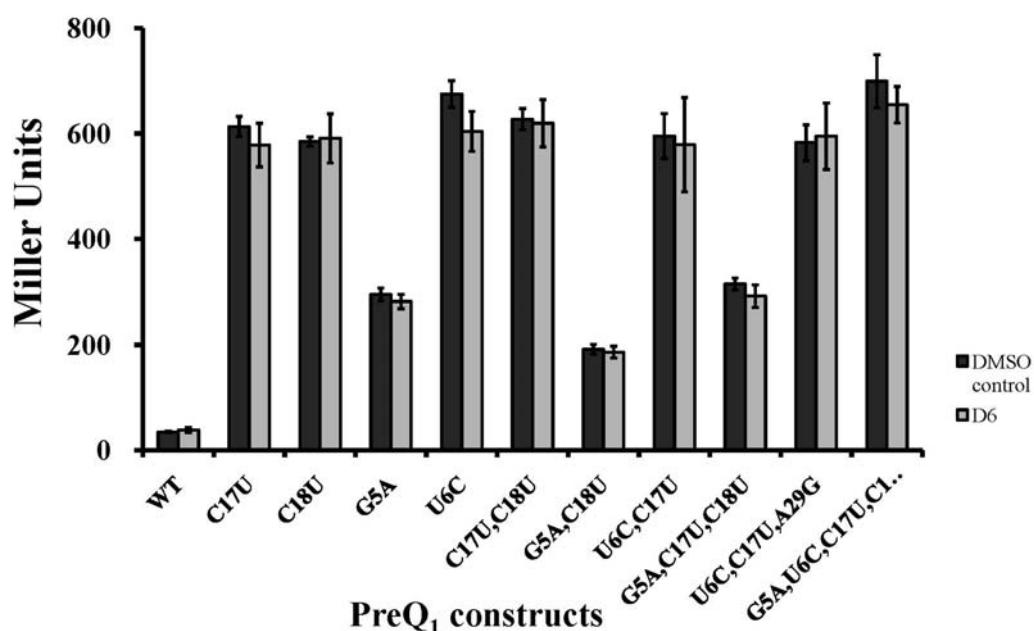


Figure 4.11. The effect of the ligand D6 on *lacZ* expression controlled by variant preQ₁ riboswitches at 37 °C.

A solution of D6 in DMSO was administered to a final concentration of 2.5 mM, with control cells receiving DMSO alone. The β -galactosidase activities are presented as Miller Units and the error bars represent the mean \pm SD of three replicates.

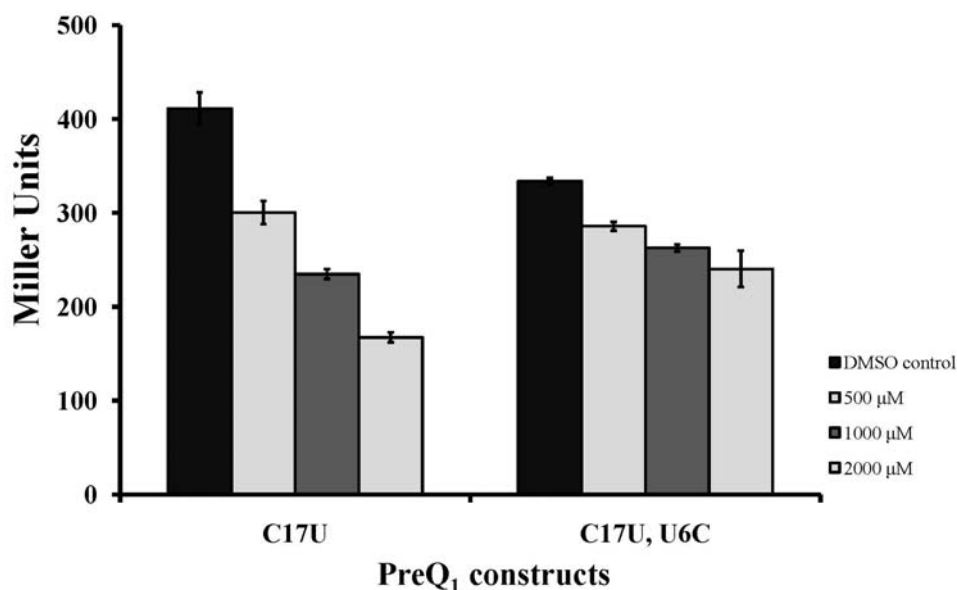


Figure 4.12. The effect of the ligand D6 on *lacZ* expression controlled by variant preQ₁ riboswitches at 16 °C.

A solution of D6 in DMSO was administered to a final concentration of 500 μM, 1000 μM and 2000 μM, with control cells receiving DMSO alone. The β-galactosidase activities are presented as Miller Units and the error bars represent the mean±SD of four replicates.

Table 4.1. *lacZ* repression ratios observed for the ligand D6 at lower temperatures.

D6 concentrations\Strains	16 °C		22 °C	
	C17U	C17U, U6C	C17U	C17U, U6C
500 μM	27.0	14.4	19.3	7.6
1000 μM	42.9	21.3	33.3	13.4
2000 μM	59.3	28.0	44.5	16.4

The ligand D8 could not be tested with the riboswitch mutants due to its insolubility in water. The ligand D9 was tested with the ten mutants and found to repress *lacZ* expression in the single mutant C17U by about 70 %, using a D9 concentration of 2.5 mM under normal bacterial growth conditions at 37 °C (Figure 4.13). The repression for the C17U mutant with the ligand D9 is dose-dependent over the tested range of concentrations between 6.25 μM and 4 mM and has an IC₅₀ value of 0.498 mM (Figure 4.14). However D9 has a toxic effect on cells at concentrations over 1 mM. The growth inhibition on *Bacillus* reached 60 % when the concentration of D9 was 4 mM (Figure 4.14).

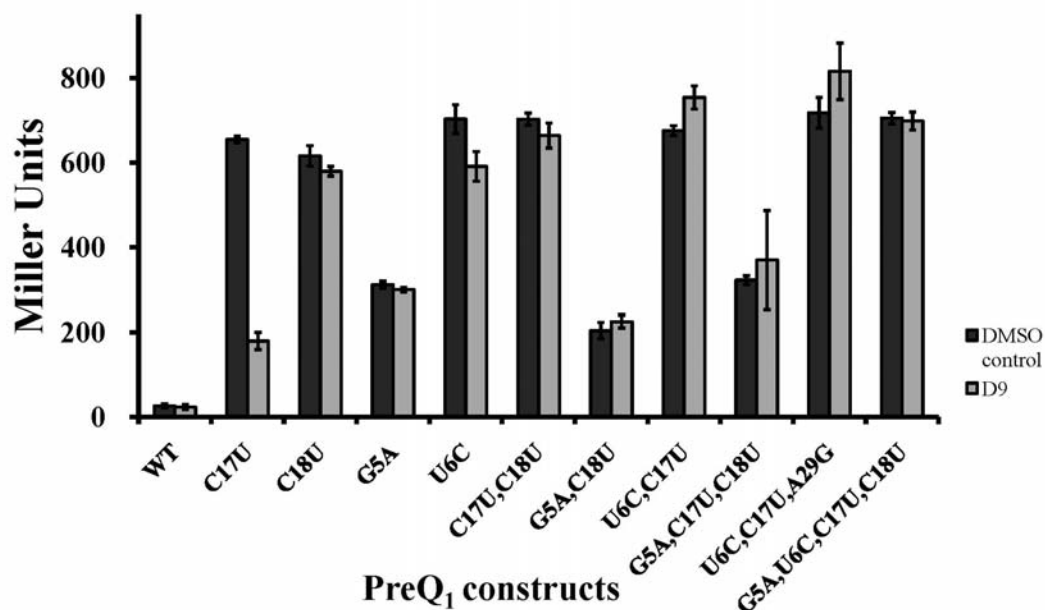


Figure 4.13. The effect of the ligand D9 on *lacZ* expression controlled by variant *preQ1* riboswitches.

A solution of D9 in DMSO was administered to a final concentration of 2.5 mM, with control cells receiving DMSO alone. The β -galactosidase activities are presented as Miller Units and the error bars represent the mean \pm SD of three replicates.

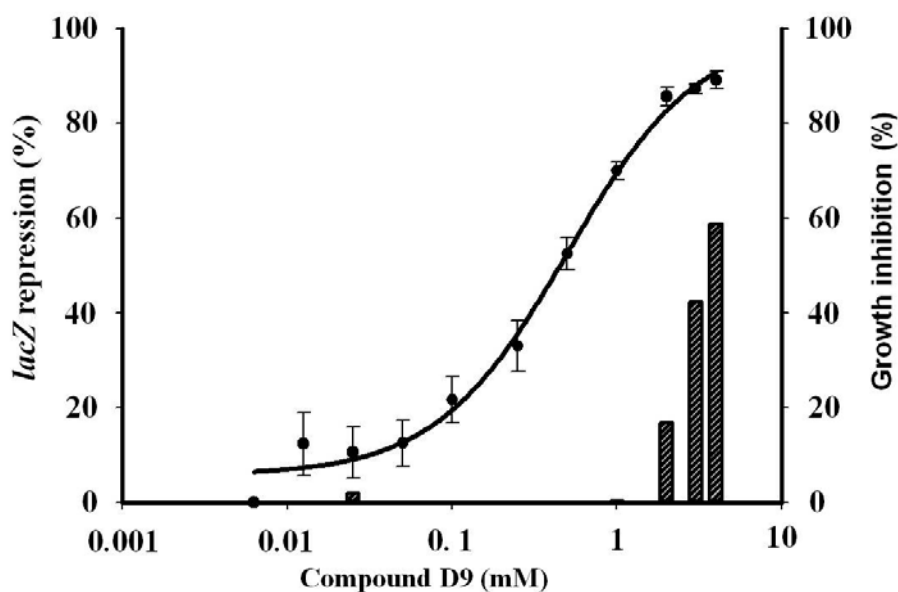


Figure 4.14. The dose-response curve and growth inhibition of the compound D9.

lacZ repression on the left axis is presented as a scatter plot with each point representing the mean of four repeats with the error bars showing mean \pm SD. Growth inhibition of *Bacillus* is presented as a bar chart on the right axis. The dose-response curve (with IC_{50} value 0.498 mM) was obtained by the *in vivo* LacZ activity assay using a D9 concentration range of 6.25 μ M to 4 mM. The IC_{50} value was calculated by non-linear regression analysis (four parameter logistic equation) using Sigmaplot 12.0.

4.4. Discussion

In this study, attempts at re-engineering preQ₁ riboswitches for response to non-natural ligands have been made in two complementary ways: structural mutagenesis of the ligand-binding site and synthesis of alternative ligands. From the structural mutagenesis study, we found that U6, C17 and C18 are important residues for maintaining preQ₁ functionality, since mutations at these positions showed high derepression ratios. Based on the preQ₁ crystal structure, C17 plays an important role in ligand recognition and binding by forming a Watson-Crick pair with the ligand preQ₁. U6 only contributes one hydrogen bond to N-9 of preQ₁, but the single mutation U6C showed a higher derepression ratio than that of C17U, suggesting that U6 plays an important role in preQ₁ RNA folding, likely through its interaction with A29. C18 forms a base pair with G5, which forms a hydrogen bond with the aminomethyl group of preQ₁. The C18U mutation showed the same level of derepression as the C17U mutation, it seems because of disruption of the base pair with G5, resulting in interfering with the G5 hydrogen bond to preQ₁ and possibly altering the position of C17 in binding site.

In light of the ligand tests with the ten preQ₁ mutants, we can further understand the relation between ligand binding and preQ₁ structure. In the preQ₀ ligand test, the G5A mutant could not respond to preQ₀ to repress reporter gene expression, even though the nitrile group of preQ₀ does not have any interaction with the residue G5. Therefore we suggest that the G5 residue plays an important role in preQ₁ riboswitch functionality, perhaps through organising the binding site architecture. In the ligand D6 test, we found repression with the single mutant C17U and the double mutant (C17U and U6C) at temperatures below 30 °C. It is possible that the O atom at position 9 of ligand D6 cannot form a stable interaction with the U6 residue, resulting in a loss of repression at normal temperature. However lower temperatures seem to promote both of the preQ₁ riboswitch mutants binding to D6 to repress reporter gene expression. Earlier we suggested that the U6 residue plays an important role in preQ₁ RNA folding due to the higher derepression ratio in the U6C mutant strain (Figure 4.6). Here the repression from ligand D6 with the double mutant (C17U and U6C) was lower than that with the C17U mutant alone, even though U6C is proposed to have a

hydrogen bond with the O atom at position 9 of the D6 ligand. This indicates that the mutation at residue U6 causes the preQ₁ riboswitch to lose its functionality, perhaps through compromising interactions with A29 which occur upon ligand binding. So far the ligand D9 shows the best repression of C17U mutant-controlled reporter gene expression. It seems that because this ligand can form a Watson-Crick pair with the C17U residue, and other functional moieties of this ligand mimic the preQ₀ structure such as the nitrogen atom at position 9 and the nitrile group attached to atom C-7. This ligand supports a better binding stability with the single mutant C17U. This result again demonstrates indirectly the previous finding that the G5 and U6 residues are important residues for functionality. When these residues are mutated in the C17U mutant, such as in the double mutant (U6C and C17U) and triple mutant (G5A, C17U and C18U), the resulting constructs are unable to repress reporter gene expression. The future work on preQ₁ riboswitch is given in the Chapter 5.

Chapter 5. Conclusions and Future Work

The work undertaken for this thesis has contributed to the delineation of the biosynthesis of OmAsp in the antibiotic A54145, the characterisation of two modifications, *i.e.* the chlorination and mannosylation of the antibiotic enduracidin, and the development of preQ₁ riboswitches as useful candidates for gene-expression tools.

The study of the biosynthesis of OmAsp demonstrates that the enzyme LptL from the A54145 biosynthetic gene cluster hydroxylates the free amino acid L-Asn. Additionally we over-expressed in *S. coelicolor* the *lptJ* and *lptK* genes, encoding putative hydroxylase and *O*-MTase enzymes, with the aim of producing CDA containing OmAsp. However the transformant *S. coelicolor* produced only wild-type CDAs. We were unable to obtain a clear conclusion as to whether CDA synthetase cannot interact with the OmAsp biosynthetic system, or whether LptJ and LptK are not sufficiently expressed by *S. coelicolor*. We have listed several factors which might explain this result in section 2.6. Nevertheless, Alexander *et al.* (2011) demonstrated our assumption of *lptJ* and *lptK* functionality in the biosynthesis of OmAsp by gene knock-out experiments [215].

Our work regarding the chlorination and mannosylation of ramoplanin was carried out by over-expression of the putative chlorinase (Ram20) and mannosyl transferase (Ram29) in *S. fungicidicus*, for the purpose of producing chlorinated and mannosylated enduracidins. In this project, we employed the optimised tetracycline-inducible vectors pMS17 and i-pAV11b for protein expression, and incorporated an additional 23 bp sequence upstream of the cloned genes which encompassed the SD sequence and ATG start codon from *eGFP*. With regards to Ram20 characterisation, we produced chlorinated enduracidin when we over-expressed Ram20 in *S. fungicidicus*. At the same time, Yin *et al.* (2010) published a similar study, demonstrating that Ram20 can monochlorinate Hpg-11 and Hpg-13 of enduracidin. However they did not observe chlorination on the Hpg-17 residue of enduracidin. Therefore further work could be conducted on the production

of altered chlorination patterns on other Hpg residues, such as Hpg-17 in enduracidin. Regarding the production of chlorinated-Hpg-17 enduracidin, we would like to replace module 8 of the End C synthetase with module 8 from the Ram14 synthetase, due to the possibility of substrate selectivity in the terminal C and TE domains of these modules, this could allow for the cyclisation of Chp-17-containing enduracidin. Additionally, we can characterise other halogenases such as ComH, which dichlorinates three Hpg residues during the synthesis of complestatin, and over-express them in *S. fungicidicus* with the goal of producing different Hpg chlorination patterns in enduracidin.

Regarding the characterisation of Ram29, the transformant SfpMS17-ram29 was created and successfully produced monomannosylated enduracidin following a 21-day fermentation with induction by aTC. Further characterisation showed that mannosylation occurs at Hpg-11 of enduracidin. Additionally, mannosylated products from substrate ramoplanin algycons can be produced by an *in vitro* system using the membrane fraction of the SfpMS17-ram29 strain. However the yield of mannosylated products is still low, despite various attempts at optimisation. We propose two factors that may account for this: **firstly**, demannosylation from endogenous mannosidase activity on the membrane of *S. fungicidicus*; **secondly**, the low solubility of enduracidin and ramoplanin under these assay conditions. Future work should be conducted on over-expressing Ram29 in a strain which produces mannosylated peptides, *e.g.* mannopeptimycin from *S. hygroscopicus*, since presumably mannosidase activity must be absent on the membrane of such strains. To overcome the problems with solubility, the *in vitro* membrane assays can be supplemented with 2 % DMSO, as Cudic *et al.* (2002) found that this can improve enduracidin and ramoplanin solubility [231].

We have successfully re-engineered the ligand specificity of the preQ₁ riboswitch by synthesising preQ₁ ligand analogues and mutating the ligand-binding residues of the riboswitch. The candidate ligand D9 had a maximal repression response of approximately 70 % with an effective dose of 1 mM, when used to inhibit β -galactosidase activity controlled by the preQ₁ C17U mutant. Further work is ongoing to synthesise a new compound with the nitrile group of D9 replaced with an aminomethyl moiety. From structural mutagenesis studies and *in vivo* assays with the

ligands D6 and D9, we believe that this new compound will afford a higher level of repression when used in conjunction with the preQ₁ riboswitch C17U mutant, than that observed for the ligand D9.

Chapter 6. Materials and Methods

General methods were taken from Sambrook *et al.* (2001) [256] and Kieser *et al.* (2000) [14] unless otherwise stated. All chemicals, solvents, enzymes and molecular biology kits were purchased from New England BioLabs, Pharmacia, Qiagen, Roche or Sigma-Aldrich-Fluka unless otherwise stated.

6.1. General information

Table 6.1. Bacterial strains used in this work

Strain/plasmid	Properties	Source/ reference
<i>Bacillus subtilis</i> 1A40	Riboswitch study	Roth <i>et al.</i> (2007) [170]
<i>Escherichia coli</i> DH5 α	General cloning host	Promega
<i>E. coli</i> BL21(DE3)	Protein over-expression host	Promega
<i>E. coli</i> ET12567/pUZ8002	<i>dam</i> ⁻ <i>dcm</i> ⁻ <i>hsdM</i> ⁻ , Kan ^R , Cm ^R ; Strain used for <i>E. coli</i> / <i>Streptomyces</i> spp. conjugation	Kieser <i>et al.</i> (2000) [14]
<i>B. subtilis</i> PY79	Indicator strain used for enduracidin bioassays	Nogami <i>et al.</i> (1984) [216]
<i>Micrococcus luteus</i>	Indicator strain used for CDA bioassays	Powell <i>et al.</i> (2007) [257]
<i>Streptomyces coelicolor</i> MT1110	Wild-type strain	Hindle and Smith (1994) [258]
<i>Sstreptomyces fungicidicus</i> ATCC31731 (Strain No. Emt 2-140)	Mutant strain with properties of enduracidin-excreting ability and high enduracidin-productivity	Nogami <i>et al.</i> (1984) [216]

Table 6.2. Antibiotics used in this work

Antibiotic	Stock concentration (mg/ml)	Final concentration ($\mu\text{g/ml}$)	
		Solid media	Liquid media
Apramycin (Apr) ¹	100	100	100
Chloramphenicol (Cm) ²	35	35	35
Hygromycin B (Hyg) ¹	50	50	50
Kanamycin (Kan) ¹	50	50	50

¹Stock solution prepared in dH₂O and filter-sterilised.

²Stock solution prepared in 100 % ethanol.

Table 6.3. Plasmids used in this work

Plasmids	Size (bp)	Characteristics	References
<i>E. coli</i> expression vector:			
pET-30b	5422	Carrying a C-terminal His ₆ -tag, Kan ^R	Novagen
<i>Streptomyces</i> expression/integration vectors:			
pMT3226	8453	Integrative glycerol inducible vector (containing phageΦC31 integrase), Apr ^R , bifunctional vector	Smith and Chater (1988) [209]
pIJ86	5795	Non-integrative vector with constitutive <i>ermE*</i> promoter, Apr ^R , bifunctional vector	Bibb <i>et al.</i> (1985) [211] and Bibb <i>et al.</i> (1994) [212]
pMS17	5894	Integrative tetracycline-inducible vector (containing phageΦC31 integrase), Apr ^R , bifunctional vector	Rodriguez-Garcia <i>et al.</i> (2005) [220]
i-pAV11b	7097	Integrative tetracycline-inducible vector (containing phageΦBT1 integrase), Hyg ^R , bifunctional vector	Rodriguez-Garcia <i>et al.</i> (2005) [220]
<i>Bacillus</i> integration vector:			
pDG1661	10156	Integrative vector (integration site is <i>amyE</i>), Cm ^R , bifunctional vector, contains <i>lacZ</i> reporter gene	Guerout-Fleury <i>et al.</i> (1996) [259]

Table 6.4. Target genes in this work

Genes	Size (bp)
<i>lptL</i> (hydroxylase)	960
<i>lptK</i> (methyl transferase)	789
<i>lptJ</i> (hydroxylase)	996
<i>ram20</i> (chlorinase)	1476
<i>ram29</i> (mannosyl transferase)	1860
<i>ppm1</i> (polyprenyl phosphomannose synthase)	909
<i>preQ₁</i> (riboswitch)	330

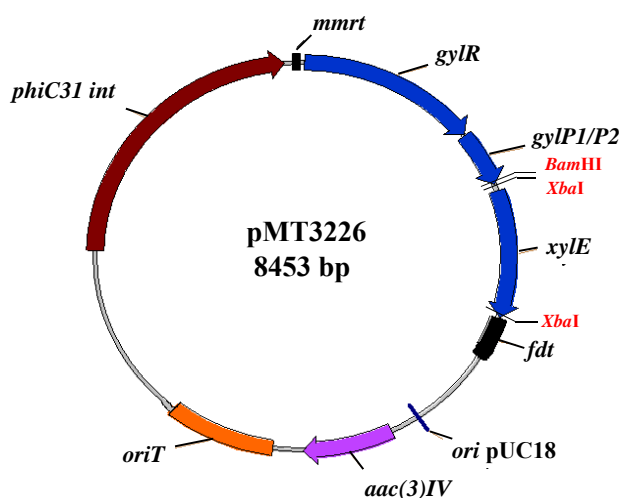
Table 6.5. Plasmids constructed in this work

Plasmid	Size (bp)	Cloning Sites
<i>E. coli</i> protein expression vectors:		
pET30b-lptL	6219	<i>NdeI</i> and <i>XhoI</i>
pET30b-lptK	6048	<i>NdeI</i> and <i>XhoI</i>
pET30b-lptJ	6237	<i>NdeI</i> and <i>XhoI</i>
<i>Streptomyces</i> protein expression/integration vectors:		
pMT3226-lptJ	9462	<i>BamHI</i>
pMT3226-lptJK	10254	<i>BamHI</i>
pIJ86-lptJ	6804	<i>BglII</i>
pIJ86-lptJK	7596	<i>BglII</i>
pIJ86-ram20	7306	<i>BglII</i>
pIJ86-ram29	7695	<i>BglII</i>
pMS17-ram20	7396	<i>EcoRV</i> and <i>NsiI</i>
i-pAV11b-ram20	8573	<i>EcoRV</i>
pMS17-ram29	7780	<i>EcoRV</i> and <i>NsiI</i>
i-pAV11b-Scppm1	8006	<i>EcoRV</i>
<i>Bacillus</i> integration vector:		
pDG1661-preQ ₁	10474	<i>EcoRI</i> and <i>BamHI</i>

6.2. Protein Expression Vector Maps

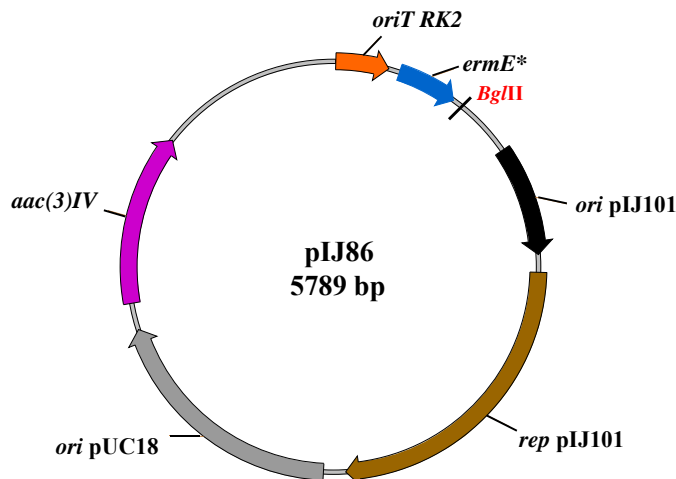
6.2.1. pMT3226

A glycerol-inducible integrative expression vector with the *gylP* cassette for gene expression. Two transcription terminators, *mmrt* and *fdt*, flank the *gylR* cassette. The presence of the PhiC31 Integrase enables this plasmid to integrate into the PhiC31 *attB* site of the *Streptomyces spp.* chromosome. Additional genes and elements include the reporter gene *xylE*, *E.coli* origin of replication (*ori* pUC18), selective marker (*aac(3)IV*; Apr^R) and the element *oriT* for plasmid transfer into *Streptomyces spp.*



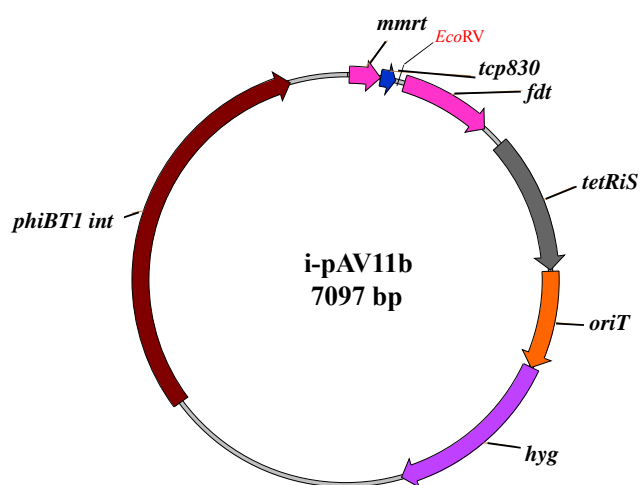
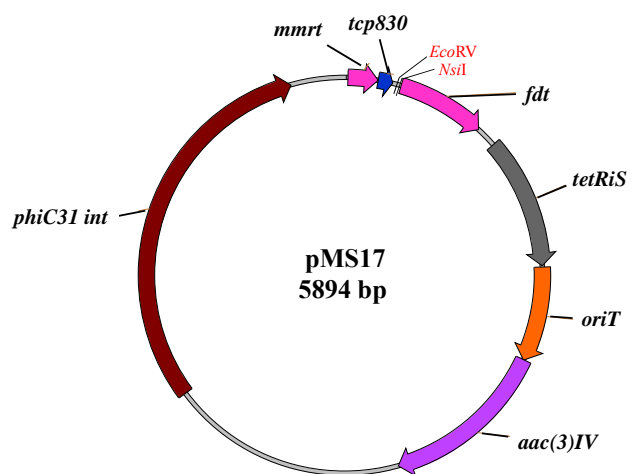
6.2.2. pIJ86

A non-integrative expression vector with a constitutive promoter *ermE** for gene expression. Additional genes and elements include the *E.coli* origin of replication (*ori* pUC18), *Streptomyces spp.* origin of replication (*ori* pIJ101 and *rep* pIJ101), selective marker (*aac(3)IV*; Apr^R) and the element *oriT* for plasmid transfer into *Streptomyces spp.*



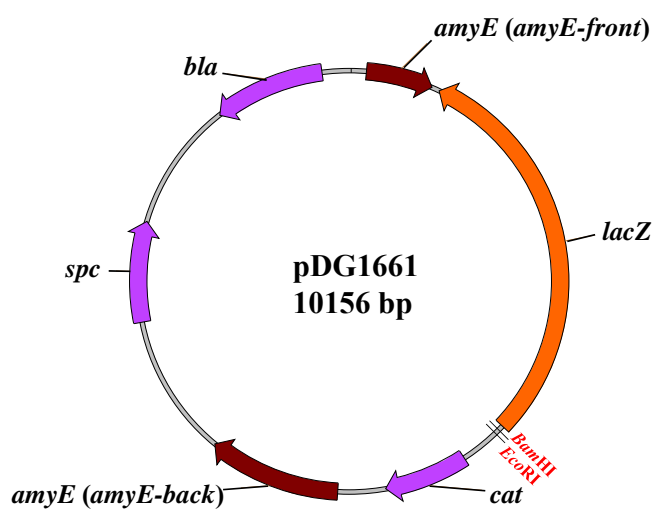
6.2.3. **pMS17 and i-pAV11b**

Both vectors are integrative expression vectors utilising a tetracycline (Tc) repressor/operator (TetR/*tetO*) regulatable system for gene expression. Two transcription terminators, *mmrt* and *fdt*, flank the promoter *tcp830* cassette. The *tetRis* encodes for a modified Tet repressor suitable for expression in *Streptomyces spp.*. Both vectors contain the gene element *oriT* for plasmid transfer. Additional genes include the selective marker *aac(3)IV* (Apr^R) and *phiC31 int* (PhiC31 Integrase) in pMS17 and the selective marker *hyg* (Hyg^R) and *phiBT1 int* (phiBT1 Integrase) in i-pAV11b.



6.2.4. pDG1661

A cloning vector (accession number U46196) for gene integration into the *amyE* locus of the *B. subtilis* chromosome *via* homologous recombination using the two gene fragments *amyE-front* and *amyE-back*. A reporter gene *lacZ* is located downstream of two cloning sites *Bam*HI and *Eco*RI. Selective markers are provided by three antibiotic resistance genes, *i.e.* *bla* (Ampicillin^R), *cat* (Cm^R) and *spc* (Spectinomycin^R).



6.3. Media and Buffers

All media and buffers used for bacterial studies are sterilised by autoclaving before use.

6.3.1. *E. coli* Culture Media

- **LB Broth (agar)**

Difco Bacto tryptone 10.0 g, Difco Yeast Extract 5.0 g, MgCl₂ 0.95 g, NaCl 10.0 g, dH₂O (distilled H₂O) to 1.0 litre (15.0 g Difco Bacto agar).

- **2YT Broth**

Difco Bacto tryptone 16.0 g, Difco Yeast Extract 10.0 g, NaCl 5.0 g, dH₂O to 1.0 litre (15.0 g Difco Bacto agar).

- **NB Broth (agar)**

Difco Nutrient Broth 8.0 g, dH₂O to 1.0 litre (15.0 g Difco Bacto agar).

6.3.2. *Streptomyces* Culture Media

- **Mannitol soya flour medium (MS medium)**

Soya flour 20.0 g, Mannitol 20.0 g, Agar 20.0 g, tap water to 1.0 litre.

- **ISP Medium 1 broth (agar)**

Pancreatic Digest of Casein 5.0 g, Yeast Extract 3.0 g, dH₂O to 1.0 litre (20.0 g Difco Bacto agar).

- **ISP Medium 2 broth (agar)**

Yeast Extract 4.0 g, Malt Extract 10.0 g, Dextrose 4.0 g, dH₂O to 1.0 litre (20.0 g Difco Bacto agar).

- **Oxoid Nutrient Agar (ONA plate; CDA bioassay plate)**

Negative control plates used to set up without calcium. No background CDA activity was observed, due to low calcium content of ONA.

- **Bouillon broth (Seed medium for *S. coelicolor* and *S. fungicidicus*)**

Meat Extract 10.0 g, Peptone 10.0 g, NaCl 5.0 g, K₂HPO₄ 2.0 g, dH₂O to 1.0 litre, pH 7.0.

- **Glycerin Bouillon plate (Enduracidin bioassay plate)**

Meat Extract 10.0 g, Peptone 10.0 g, NaCl 5.0 g, Difco Bacto agar 15.0 g, Glycerol 50 ml, dH₂O to 1.0 litre, pH 8.0.

- **SV2 medium**

Glucose 15.0 g, Glycerol 15.0 ml, Soy Peptone 3.0 g, CaCO₃ 1.0 g, dH₂O to 1.0 litre, pH 7.0.

- **Enduracidin fermentation medium**

Corn Flour 80.0 g, Corn gluten meal 30.0 g, Corn-steep-liquor 5.0 ml, NaCl 1.0 g, (NH₄)₂SO₄ 3.0 g, ZnCl₂ 0.1 g, Lactose 10.0 g, Vegetable oil 14.0 ml, Potassium lactate 5.0 ml, dH₂O to 1.0 litre, pH 7.0.

6.3.3. General Buffers

6.3.3.1. Buffers and Media for Preparation of *E.coli* Competent Cell

- **0.5 M PIPES (piperazine-1,2-bis[2-ethanesulfonic acid]) solution**

15.1 g PIPES in 80.0 ml dH₂O, KOH added until a solution is formed. The pH should be approximately pH 6.7. Add dH₂O up to 100.0 ml.

- **Inoue solution**

MnCl₂ 10.9 g, CaCl₂ 2.2 g, KCl 18.7 g, 0.5 M PIPES solution 20.0 ml, dH₂O to 1.0 litre and filter-sterilize.

- **SOB medium**

NaCl 0.5 g, tryptone 20.0 g, yeast extract 5.0 g, KCl 0.186 g. dH₂O to 1.0 litre.

- **SOC medium**

NaCl 0.5 g, tryptone 20.0 g, yeast extract 5.0 g, KCl 0.186 g. dH₂O to 1.0 litre. After autoclaving, add 20.0 ml of sterile 1.0 M glucose.

6.3.3.2. Buffers and Media for Preparation of *Bacillus* Competent Cell

■ **10xT-Base**

(NH₄)₂SO₄ 4.95 g, K₂HPO₄ 34.8 g, KH₂PO₄ 14.96 g, sodium citrate 2.57 g. dH₂O to 250.0 ml.

■ **SpC (minimal culture media)**

100.0 ml of 1x T-Base, 1.0 ml of 50 % Glucose, 1.5 ml of 1.2 % MgSO₄, 2.0 ml of 10 % yeast extract, 2.5 ml of 1 % casamino acids (Each component is autoclaved separately before mixing).

■ **SpII (starvation media)**

100.0 ml of 1x T-Base, 1.0 ml of 50 % Glucose, 7.0 ml of 1.2 % MgSO₄, 1.0 ml of 10 % yeast extract, 1.0 ml of 1 % casamino acids, 0.5 ml of 100 mM CaCl₂ (Each component is autoclaved separately before mixing).

6.3.3.3. DNA Electrophoresis Buffer

- 10x TBE buffer containing 890 mM Tris-Cl, 890 mM Boric acid and 20 mM EDTA, pH 8.2.

6.3.3.4. Protein Electrophoresis Solutions

■ **5x Denaturing sample buffer**

0.312 M Tris-Cl pH 6.8, 0.05 % bromophenol blue, 50 % glycerol, 10 % SDS, 25 % β-mercaptoethanol.

■ **1x Running buffer**

25.0 mM Tris-base pH 8.8, 192.0 mM glycine, 0.1 % SDS.

■ **Coomassie Blue staining solution**

50 % methanol, 10 % glacial acetic acid, 0.1 % Coomassie Brilliant Blue R-250 in dH₂O.

■ **Destaining solution**

1. 50 % methanol, 10 % glacial acetic acid in dH₂O.
2. 5 % methanol, 7 % glacial acetic acid in dH₂O.

Table 6.6. SDS-PAGE gel preparation

Stock solution	Separating gel (12.50 %)	Stacking gel (5 %)
Tris-Cl (1.5M, pH 8.8)	3.0	-
Tris-Cl (0.5M, pH 6.8)	-	1.67
Deionized water	2.60	3.14
30 % Acrylamide/ Bis (29/1)	4.16	1.00
10 % (w/v) SDS	0.07	0.07
10 % (w/v) APS	0.10	0.10
TEMED	0.01	0.01
Total volume (ml)	9.94	6.00

6.4. Preparation of Chemically Competent *E.coli* Cells

This method was modified from that described in Inoue *et al.* (1990) [260] with the buffers and media used as listed in section 6.3.3.1. A single colony of *E.coli* (DH5 α) from an LB agar plate was used to inoculate 3 ml SOB medium and cultured overnight at 37 °C. Three ml of overnight culture were then used to inoculate 250 ml SOB, and grown at 18 °C with shaking until OD₆₀₀ 0.4. The culture was put on ice for 10 min, and then centrifuged (10 min, 4000 g, 4 °C). The cell pellet was re-suspended in 80 ml ice-cold Inoue solution and pelleted by centrifugation (10 min, 4000 g, 4 °C). The cells were re-suspended in 20 ml cold Inoue solution with DMSO (final concentration 7 %) and incubated on ice for 10 min. The cell suspension was then dispensed in 200 μ l aliquots into 1.6 ml eppendorf tubes and immediately frozen in liquid nitrogen and stored at -80 °C.

6.5. Preparation of *Bacillus* Competent Cells

This method was modified from that described in Spizizen, (1958) [261] with the buffers and media used as listed in section 6.3.3.2. A single *Bacillus* colony from an NB agar plate was used to inoculate 3 ml SpC medium and cultured overnight at 37 °C. One ml of the overnight culture was used to inoculate 8 ml of SpC and was incubated at 37 °C for 3 h with shaking until OD₄₅₀ 1.2-1.6. The culture was again diluted with the same volume of SpII medium and further incubated for 2 h at 37 °C with shaking. After incubation, the culture maintains maximum competence for approximately 1 h; competent cells were dispensed in 500 µl aliquots into 1.6 ml eppendorf tubes. At this stage, cells can be either frozen (containing 10 % glycerol and stored at -80 °C) for future use or used directly for transformation.

6.6. *E. coli* Transformation Procedures

Twenty µl of ligation product were added into 200 µl of *E.coli* (DH5a) competent cells, gently mixed and incubated on ice for 20 min. The cells were then heat shocked at 42 °C for 1 min and incubated on ice for a further 5 min. Three hundred µl of SOC medium were added to cells and incubated at 37 °C with shaking for 1 h before 300 µl culture were spread onto an LB agar plate supplemented with the appropriate antibiotic. The plate was incubated at 37 °C overnight.

6.7. *Bacillus* Transformation Procedures

This method was modified from that described in Spizizen (1958) [261]. Four µl of pDG1661 construct were added to 500 µl of *Bacillus* competent cells and incubated at 37 °C with shaking for 30 min. Three hundred µl of LB was then added for stimulating growth of competent cells and incubated for another 30 min at 37 °C shaker before 500 µl culture were plated out on NB agar plate containing 5 µg/ml Cm.

6.8. Growth of *Streptomyces* Cultures

S. fungicidicus (ATCC31731) and *Actinoplanes* sp. (ATCC33076) were purchased from ATCC bioresource centre and cultured in ISP Medium 2 and ISP Medium 1 broth respectively.

6.8.1. Plating out *Streptomyces* Spore Suspensions

The *Streptomyces* spores from ATCC were suspended in appropriate media and pipetted onto a solid media plate. The spores were spread over the surface of the plate with an ethanol flamed glass spreader and incubated at 30 °C. Spores would appear as a grey pigment over the surface of the colonies after 5 to 7 days.

6.8.2. Harvesting of Spores from *Streptomyces* mycelia

Sterile water (9 ml) was added to the plates and a sterile wire loop was used to gently scrape off the spores. The spore suspension was pipetted into a 15 ml falcon tube and vortexed thoroughly to separate the spores from the mycelia.

Spore filter tubes were constructed using small glass (15 x 125 mm) tubes with a 0.5 cm diameter hole in the bottom, containing a plug of sterile cotton wool [14]. The small tubes were housed into larger tubes (15 x 150 mm), sealed with aluminium foil and autoclaved. The spore suspension was then filtered through the small tube into a 50 ml falcon tube and pelleted by centrifugation (15 min, 4000 g, 4 °C). The supernatant was discarded. One ml of sterile 20 % glycerol was added to the spore pellet, vortexed thoroughly and transferred to a screw cap tube and stored at -80 °C.

6.8.3. Growth of *Streptomyces* Mycelia in Liquid Media

Ten µl of spore stock were used to inoculate 20 ml of the appropriate liquid medium. The cultures were grown for 5 days at 30 °C with shaking (230 rpm). The mycelia were then collected for DNA extraction or further fermentation for antibiotic production.

6.9. Extraction of *Streptomyces* Genomic DNA

One ml of mycelial culture (section 6.8.3) was spun down and the medium discarded. The genomic DNA was extracted using a QIAGEN Genomic DNA extraction kit and following the manufacturer's protocol for bacterial genomic DNA extraction. The genomic DNA was re-dissolved in 100 µl of 10 mM Tris-Cl buffer pH 8.5.

6.10. Cloning of Target Genes (*lptL*, *lptK*, *lptJ*, *lptJK*, *ram20*, *ram29*, *ppm1*, *preQ₁*)

The gene sequences for *lptL*, *lptK* and *lptJ* from the antibiotic A54145 biosynthetic gene cluster of *S. fradiae*, *ram20* and *ram29* from the antibiotic ramoplanin biosynthetic gene cluster of *Actinoplanes sp.* and *ppm1* from *S. coelicolor* were taken from the NCBI database. The vector pDG1661 containing the preQ₁ riboswitch sequence amplified from the upstream region of the gene *queC* from *B. subtilis* 1A40 was kindly supplied by the Breaker group [170]. The primers were designed using Vector NTI 11.5 software (InforMax Inc., USA). The software enabled analyses of the secondary structures of the primers thus avoiding potential problems that may inhibit the PCR efficiency. After cloning of the target genes into suitable vectors, their sequences were verified by sequencing.

6.10.1. PCR Amplification Reaction

Genes were amplified by PCR with Phusion DNA polymerase (Finnzymes, Finland). The reaction components and conditions are listed in Table 6.7 and Table 6.8.

Table 6.7. PCR reaction components

PCR reaction components	Volume
5x Phusion TM HF Buffer	10
10 mM dNTPs	1
10 mM Forward primer	1
10 mM Reverse primer	1
Template DNA	2
DMSO	1.5
Phusion TM DNA Polymerase	0.5
Total volume (µl)	50

Note: The gene *ram20* PCR reaction requires 7 % DMSO and the gene *ram29* PCR reaction requires 2 mM MgCl₂.

Table 6.8. PCR amplification conditions

Cycle step	Temperature (°C)	Time
Initial denaturation	98	2 min
Denaturation	98	20 sec
Annealing ^a	42-60	20 sec
Extension ^b	72	dependent on gene size
Final extension	72	5 min

^aAnnealing temperature: 50 °C for *lptL*, 60 °C for *lptK*, 40 °C for *lptJ*, *ram20* and *ram29*.

^bPhusionTM DNA polymerase activity: 1 Kb/30 sec.

6.10.2. Primers for the gene *lptL*, *lptK*, *lptJ*, *lptJK*, *ram20*, *ram29*, *eGFP* and *ppm1*

Table 6.9. List of primers for antibiotic engineering work

Primer name ^a	Sequence (5'→3')
<i>In vitro</i> protein expression primers (Cloning into pET30b expression vector)	
lptL-F1	AAAAAACATATGGAACCCGAGAACACCTT
lptL-R1	AAAAAACTCGAGGCTCAGTACACGCGAATGGT
lptK-F1	AAAAAACATATGACCATCGCCCTCGCCGA
lptK-R1	AAAAAACTCGAGGCCGCGACGCGGTTCGG
lptJ-F1	AAAAAACATATGGAACCGACCACCGCCT
lptJ-R1	AAAAAACTCGAGTCGGTGGCCTTCCTGCGG
<i>In vivo</i> protein expression primers (Cloning into <i>Streptomyces</i> expression vector)	
sclptJK-F1	AAAAAAAGGATCCGAAAGAGCCCCGCATGGAA
sclptJ-R1	AAAAAAAGGATCCTCATCGGTGGCCTTCCTGC
sclptJK-R1	AAAAAAAGGATCCACGTGGCTCAGCCGCGA
sfram20-F1	AAAAAAAGATATCTGACCAAAGGAGGCGGACATATGGCCA TGGCTGCTCAACCGGAAGA
sfram20-R1	AAAAAAAATGCATCTCTGTTCGTGTGCTCAGTCGG
sfram29-F3	AAAAAAAGATATCTGACCAAAGGAGGCGGACATATGGCCA TGGATGTCCCGAGGGTGC
sfram29-R1	AAAAAAAATGCATTCAGGGCTTGTCGCATCG
Scppm1F	AAAAAAAGATATCTGACCAAAGGAGGCGGACATATGGCCA TGAACGACGGCGACGGGA
Scppm1R	AAAAAAAGATATCTCAGGCCTTGGCGTCGGA
eGFP-F1	AAAAAAAGATATCTGACCAAAGGAGGCGGACAT
eGFP-R1	AAAAAAAGGTACCTTACTTGTACAGCTCGTCCATGC
Sequencing primers	
glyP	CTTTCGATAACACAGCTCTTGA
xyle	ATCCTCATCCACAACCTTGAA
ermP-F1	TGCACGCGGTTCGATCTTGA
ermP-R1	GTTGGGGTCCTGGACCAGGTT
pMS17-F1	GAGCTCGGTACTGCTGTTGGCTA
pMS17-R1	CAGGGAGTCAAAGGCCGCTT
pAV11-F1	GAGCTCGGTACTGCTGTTGGCTA
pAV11-R1	CTCATCAAGGTCAAGGCGTAGGT

^aF and R indicate forward and reverse primers, respectively.

6.10.3. Primers for Constructing PreQ₁ Riboswitch Variants

Table 6.10. List of primers for riboswitch engineering work

Primers	Purposes	Sequences (5'→3')
pRSW104f	pDG1661 sequencing primer	AGTAACTTCCACAGTAGTTCA
pRSW105	C17U, C18U mutagenic primer forward	AGAGGTTCTAGCTACATTCTCTATA
pRSW106	C17U, C18U mutagenic primer reverse	TATAGAGAATGTAGCTAGAACCTCT
pRSW107	U6C, C17U mutagenic primer forward	AGAGGCTCTAGCTACATCCTCTATA
pRSW108	U6C, C17U mutagenic primer reverse	TATAGAGGATGTAGCTAGAGCCTCT
pRSW109f	PreQ1 cloning forward	AAAAAAAGAATTCATAATGAAACGAACC GTCACTATAG
pRSW110r	PreQ1 cloning reverse	AAAAAAAGGATCCGACAATTGCTTTTTCT TTTTTCATG
pRSW113	C17U, C18U, G5A mutagenic primer forward	AGAGATTCTAGCTACATTCTCTATA
pRSW114	C17U, C18U, G5A mutagenic primer reverse	TATAGAGAATGTAGCTAGAATCTCT
pRSW115	G5A mutagenic primer forward	AGAGATTCTAGCTACACCCTCTATA
pRSW116	G5A mutagenic primer reverse	TATAGAGGGTGTAGCTAGAATCTCT
pRSW117	U6C mutagenic primer forward	AGAGGCTCTAGCTACACCCTCTATA
pRSW118	U6C mutagenic primer reverse	TATAGAGGGTGTAGCTAGAGCCTCT
pRSW119	C18U mutagenic primer forward	AGAGGTTCTAGCTACACTCTCTATA
pRSW120	C18U mutagenic primer reverse	TATAGAGAGTGTAGCTAGAACCTCT
pRSW121	C18U, G5A mutagenic primer forward	AGAGATTCTAGCTACACTCTCTATA
pRSW122	C18U, G5A mutagenic primer reverse	TATAGAGAGTGTAGCTAGAATCTCT
pRSW123	U6C, C17U, A29G mutagenic primer forward	CCTCTATAAAAGACTAAGGACGA
pRSW124	U6C, C17U, A29G mutagenic primer reverse	TCGTCCTTAGTCTTTTATAGAGG
pRSW125	C17U, C18U, G5A, U6C mutagenic primer forward	AGAGACTCTAGCTACATTCTCTATA
pRSW126	C17U, C18U, G5A, U6C mutagenic primer reverse	TATAGAGAATGTAGCTAGAGTCTCT

6.11. DNA Electrophoresis

A 1.2 % agarose gel was made by dissolving agarose in 50 ml of 1x TBE buffer by heating. Ethidium bromide (2 μ l from 10 mg/ml stock) was added, mixed well before pouring into a gel caster with a comb. After setting the gel, it was placed in a gel running tank and immersed in TBE buffer. Samples were mixed with DNA loading dye, loaded into the well and electrophoresised at 100 V for approximately 30 min. The DNA migration through the gel was observed using a UV transilluminator.

6.12. Extraction of DNA from Agarose Gels

The bands corresponding in size to the target DNA fragments were excised from the agarose gel with a scalpel. The DNA was extracted using a QIAquick[®] Gel Extraction Kit (Qiagen), following the manufacturer's protocol. The DNA was eluted with 50 μ l of 10 mM Tris-Cl pH 8.5.

6.13. Digestion, Ligation and Transformation

Restriction enzymes and T4 DNA ligase were purchased from New England Biolabs. Target genes were digested with the appropriate restriction enzymes and cloned into the vectors. Digestion and ligation conditions were carried out according to the manufacturer's protocols. Ligation products were used to transform *E. coli* DH5 α cells by heat shocking at 42 °C and incubating the cells on LB-agar plates with selective antibiotics overnight at 37 °C. The colonies obtained by antibiotic selection were screened by colony PCR. Verified colonies containing the correct gene construct were used to inoculate 3 ml of LB with appropriate antibiotics for selection, grown overnight and the plasmids extracted using a QIAprep[®] Spin Miniprep Kit (Qiagen). The target gene sequences in the constructs were verified by DNA sequencing. The constructs were subsequently used to transform *E. coli* BL21 (DE3) cells for protein over-expression, *E. coli* strain ET12567 for *Streptomyces spp.* conjugation or *Bacillus* strain for riboswitch studies.

6.14. Recombinant Protein Expression and Purification

A 3 ml LB liquid culture was incubated overnight. The culture was diluted 1:100 in fresh LB with appropriate antibiotics and grown at 37 °C with shaking until the OD₆₀₀ reached 0.6 (usually after 3 h). IPTG was added to the culture at a final concentration of 1 mM for induction of protein over-expression and the culture further incubated at 30 °C for 3 h. Cells were then pelleted by centrifugation (15 min, 6000 g, 4 °C) and re-suspended in 50 ml of protein purification lysis buffer (20 mM KH₂PO₄, 500 mM NaCl, 20 mM Imidazole, pH 7.4). The cells were lysed using a Cell disrupter (Constant systems Ltd) with an aperture pressure of 22 K p.s.i. The cell lysates were spun down and the supernatant was applied to a Fast Protein Liquid Chromatography system (FPLC; AKTA prime, Amersham).

Recombinant protein purification was performed using FPLC in conjunction with a HiTrap[™] Chelating HP 5 ml column (GE Healthcare, UK). The column was equilibrated with lysis buffer and activated with a 0.1 M Ni²⁺ ion solution before re-equilibrating with lysis buffer. The cell lysate was loaded, and the column was washed with lysis buffer until the absorbance read approximately zero. The column was then washed again with 6 column volumes of wash buffer (20 mM KH₂PO₄, 500 mM NaCl, 60 mM Imidazole, pH 7.4) and recombinant protein was eluted with 6 column volumes of elution buffer (20 mM KH₂PO₄, 500 mM NaCl, 250 mM Imidazole, pH 7.4). The column was washed and stripped of nickel ions with EDTA-buffer (20 mM KH₂PO₄, 500 mM NaCl, 500 mM EDTA, pH 7.4).

Elate fractions containing the target protein were pooled before proceeding to buffer exchange using a Vivaspin centrifugal concentrator (Sartorius, UK) to replace the elution buffer with storage buffer (100 mM HEPES, 100 mM NaCl, 10 % glycerol, pH 7.5). The protein was analysed by SDS-PAGE and stored at -80 °C for further analysis.

6.15. Introduction of DNA into *Streptomyces* by Conjugation

A broad host range mobilisation system was established by Simon *et al.* (1983) demonstrating the intergeneric transfer of plasmids from *E. coli* to *Streptomyces* strains [262]. Most *E. coli* strains contain site-specific MTases, *e.g.* Dam and Dcm which methylate vectors. These methylated vectors are recognised as foreign genetic material by the recipient *Streptomyces* strains and are digested by endonucleases. Therefore the methylation-deficient *E. coli* strain ET12567 is employed for transferring plasmids to *Streptomyces* strains [263]. Additionally the ET12567 strain carries the non-transmissible pUZ8002 which encodes a transfer protein Tra enabling the transfer of expression vector containing *oriT* (from the IncP-group plasmid RP4) to *Streptomyces* strains [14].

The conjugation procedures are based on the protocol developed by Kieser *et al.* (2000) [14]. An overnight culture (10 ml) of *E. coli* ET12567/pUZ8002 transformed with *Streptomyces* expression vectors such as pMT3226, pIJ86, pMS17 or i-pAV11b was set up. The culture was pelleted and washed twice with an equal volume of LB medium to remove antibiotics that might inhibit the growth of *Streptomyces spp.* and re-suspended in 500 µl LB. Approximately 10^8 *Streptomyces* spores were added to 500 µl 2xYT Broth, heat shocked at 50 °C for 10 min to induce germination, and cooled to room temperature. The *E. coli* cells were mixed with 500 µl heat-shocked spores and spun down. The supernatant was discarded and the residual liquid was used to re-suspend the cell pellet before plating out on a MS medium plate (containing 10 mM MgCl₂) and incubating at 30 °C for 16-20 h. Five hundred µl water containing 0.5 mg nalidixic acid and 1.25 mg appropriate antibiotic for plasmid selection was added and overlaid by spreading. Transformant colonies can usually be observed after 3-6 days incubation at 30 °C. The exoconjugants were picked off to selective MS agar plates containing nalidixic acid (25 µg/ml) to inhibit growth of the *E. coli* strain. Spores appeared from exoconjugant colonies after 2 weeks and were used to inoculate Bouillon broth and incubated at 30 °C for 2-4 days. The mycelia were collected for genomic DNA extraction or further antibiotic fermentation.

6.16. Catechol Dioxygenase Assay

Streptomyces transformants (pMT3226 constructs) were selected on MS-apramycin plates containing 1 % glycerol. After incubation at 30 °C for 2 days, plates were overlaid with an aqueous solution of 300 µl of 0.5 M catechol, wrapped in foil and incubated for 1 h at 30 °C. The expression of catechol dioxygenase is indicated by yellow pigmentation of the mycelia and surrounding media caused by the formation of hydroxymuconic semialdehyde [14, 210].

6.17. Antibiotic Activity Plate Bioassay

Bioassays were performed using plates containing antibiotic producing strains or paper discs infused with antibiotics. Approximately 10^5 - 10^6 spores from *S. fungicidicus* and *S. coelicolor* were used to inoculate Glycerin Bouillon and ONA plates respectively and incubated at 30 °C for 4 days for production of antibiotics. Additionally, 20 µl of antibiotic samples from cultures or different purification steps were added to small filter paper discs, placed onto plates and incubated for 1 day at 30 °C to allow antibiotics to diffuse.

B. subtilis PY79 and *M. luteus* were used as indicator strains for the enduracidin and CDA activity assays respectively. Glycerol stocks of the indicator strains were streaked out on LB plates and left to grow overnight at 30 °C. A single colony was used to inoculate 3 ml LB medium and incubated with shaking at 30 °C overnight. The cell density of the overnight culture was measured and diluted to OD₆₀₀ 0.1. The culture was further diluted by 50-fold and 100 µl added to 7 ml of melted soft agar of LB and ON for enduracidin and CDA activity assays respectively. The melted soft agar containing the indicator strain was poured onto plates which contained 4 day-old antibiotic producing strains or paper discs with antibiotics and incubated at 30 °C overnight. The inhibition zones can usually be observed after 24 h.

CDA activity assays were performed on ONA plates with and without 16 mM CaNO₃ to ensure the inhibition zone observed are the results of Calcium Dependant Antibiotics and not the other possible antibiotics produced by *S. coelicolor*.

6.18. CDA Purification from *S. coelicolor* Fermentation Broth

The spores of wild-type or transformant *S. coelicolor* strains (10^5 - 10^7) were used to inoculate SV2 media and grown at 30 °C for 7 days (purple pigmentation will be observed in the media). The cultures were transferred to 50 ml falcon tubes and pelleted by centrifugation (15 min, 5000 g). The supernatant which contained CDA was collected and adjusted to pH 2.0, followed by filtration through standard Whatman filter paper. A Varian C₁₈ Bond Elut SPE cartridge was washed with 1 column volume of methanol followed by 2 column volumes of water. The CDA filtrate was then loaded onto the cartridge which was then washed with 2 column volumes of water. The CDA was eluted with 2 column volumes of 30 %, 50 % and 100 % methanol, with the majority of CDA found in the 100 % methanol eluate. The eluates were dried by rotary evaporation under vacuum and were further analysed by LC-MS (section 6.20). This part of work was done by Laura Nunns.

6.19. Enduracidin Purification from *S. fungicidicus* Fermentation Broth

The spores of wild-type or transformant *S. fungicidicus* strains (10^5 - 10^7) were used to inoculate 75 ml of seed medium (Bouillon medium) and cultured at 30 °C for 3 days. Ten ml of mycelia were then transferred to 300 ml of enduracidin fermentation medium containing antibiotics as appropriate and cultured at 30 °C with shaking for 14-21 days (for pMS17/i-pAV11b transformant strains tetracycline inducer was added to a final concentration of 0.1 µg/ml every 3 days). After fermentation, the cultures were transferred to 50 ml falcon tubes and pelleted by centrifugation (15 min, 10,000 g). Mycelia were collected, re-suspended in 70 % acidic methanol (pH 2.0), and stirred for 3 h. The methanol extract was then centrifuged at 10,000 g for 10 min and the supernatant was mixed with an equal volume of ethyl acetate for removal of proteins and lipids. The aqueous partition was separated and adjusted to pH 8.2. Enduracidin in the aqueous layer was then transferred into 300 ml (100 ml x 3) of butanol by liquid-liquid extraction. After washing with water, enduracidin in the

butanol layer was partitioned into 300 ml (100 ml x 3) of 0.005 N HCl aqueous solution (pH 2.0). When the acidic aqueous layer was adjusted to pH 8.2, enduracidin was partitioned back to 300 ml (100 ml x 3) of butanol. The butanol layer containing enduracidin was separated and concentrated by evaporation under reduced pressure, leaving a white crude powder. The crude enduracidin was re-dissolved in 5 % acetonitrile and further purified by RP-HPLC using a Gemini RP column (Phenomenex, C₁₈, particle size of 5 µm; 250 x 10 mm) with the following gradients: 10 min linear gradient from A/B (95:5 %) to A/B (80:20 %), 30 min gradient to A/B (60:40 %), 5 min linear gradient to A/B (5:95 %) and held at A/B (5:95 %) for 4 min, 3 min linear gradient to A/B (95:5 %) before re-equilibrating at A/B (95:5 %) over 8 min. A flow rate of 3 ml/ min was used and the elution of enduracidin was monitored at A₂₃₀. The following solvents were used; solvent A: water with 0.1 % formic acid; solvent B: acetonitrile with 0.1 % formic acid. Fractions were collected and further analysed by LC-MS. This work was done with the help from Laura Nunns and Matthew Styles.

6.20. Analysis of CDA and Enduracidin extracts using LC-MS

LC-MS analysis was carried out on a Micromass LCT TOF mass spectrometer (Waters Ltd., UK), equipped with an electrospray ionisation source run in positive mode (scanning from *m/z* 700 to 2500) combined with a Waters 2790 Separation module. Samples were dissolved in acetonitrile/water/formic acid (50/50/0.1). Gradient elution was carried out using a Gemini RP column (Phenomenex, C₁₈, particle size of 3 µm; 150 x 4.6 mm) at a flow rate of 1 ml/min with the following gradients: 10 min linear gradient from A/B (80:20 %) to A/B (30:70 %), 1 min gradient to A/B (0: 100 %) and held at A/B (0:100 %) for further 4 min. Solvent A: water with 1 % acetonitrile and 0.1 % formic acid; solvent B: acetonitrile with 1 % water and 0.1 % formic acid. This work was done by Laura Nunns.

6.21. Tandem Mass Spectrometry (MS)

Tandem MS analysis was carried out on a Q-TOF Ultima Global instrument (Waters Ltd., UK) by direct infusion of a purified sample dissolved in acetonitrile/water/formic acid (50/50/0.1). Borosilicate nanoemitters (Proxeon, UK) were loaded with sample using gel loading pipette tips. The tips were mounted onto the source and then activated at 0.4 kV capillary voltage by applying pressure against the sample cone, snapping the glass fibre. Once activated, the capillary voltage was ramped steadily until a consistent ion beam was observed, typically between 1.6 and 1.8 kV. Each nanoemitter holds approximately 20 μL and provides a spray for approximately 30 min, giving a flow rate of approximately 0.66 $\mu\text{L}/\text{min}$. The instrument was calibrated using the tandem MS of [Glu]-Fibrinopeptide B.

Once an ion beam was achieved, the MS source could be tuned to the product by ramping the capillary voltage and the cone voltage until the signal intensity peaks. An MS scan was carried out to determine a suitable m/z value to fragment (ideally the $[\text{M}+2\text{H}]^{2+}$ ion, as this yields singly charged fragments). An MSMS scan would then be carried out by selecting the parent ion for analysis, and raising the collision energy within the collision cell until the parent ion was a minority peak on the resulting spectrum. The data was collected over 20 min, or until a sufficient spectrum was achieved for sequencing analysis. The subsequent spectrum was background subtracted (polynomial order= 4, 40 % below curve, tolerance= 0.010), smoothed (Savitzky Golay smoothing, smooth channels= ± 5 , number of smooths= 2), and centred (minimum peak width at half height= 4, 80 % centroid top) by the MassLynx software (Version 4.0, Waters) prior to analysis. This work was carried out by Matthew Styles (Micklefield group).

6.22. Enduracidin Sequencing

A database of theoretical m/z values was set up based on all the possible fragmentation patterns of enduracidin, assuming two cleavages yielding a linear peptide ion. Additionally a further database was set up assuming that the first cleavage occurs at the ester bond linkage, thus additional fragmentation events yield a, b, y and z ions at either one or both ends (as both the C- and N-termini have undergone fragmentation). Tandem MS data were then compared with these theoretical values to generate ion series that can be used to characterise the peptide, and signature losses were identified by observation.

6.23. Bradford Assay (Protein Quantitation)

BioRad[®] Dye reagent concentrate was used to perform protein quantitation. The assay dye reagent was prepared by diluting 1 part of the reagent concentrate with 4 parts of dH₂O. One ml of assay dye reagent was then added to each tube containing 20 μ l of protein samples such as target proteins and protein standard BSA. Protein samples were incubated at room temperature for at least 5 min and the absorbances at OD₆₀₀ were measured for each sample using a Varian UV-Vis Spectrophotometer. The protein standard BSA was used to produce a series of dilutions in the concentration range 0.2 to 0.9 mg/ml for generating a standard curve for deriving concentrations of unknown protein samples *via* their absorbances.

6.24. Sodium-Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

SDS-PAGE gels were prepared in accordance with the SDS-PAGE gel preparation table (section 6.3.3.4). Protein samples were mixed with protein loading buffer and denatured by heating to 100 °C for 5 min before loading. The gel was run at 200 V until the dye front reached the bottom of the gel and visualised by staining with coomassie brilliant blue. The BioLabs[®] prestain protein marker was used as a protein marker.

6.25. Amino Acid Pre-column Derivatisation

Phenylisothiocyanate (PITC) was used to derivatise amino acids in this study due to the stability of the PITC derivatives and easy removal of excess PITC by evaporation [264]. Here we modified the PITC derivative method from that described in Heinrikson and Meredith (1984) [264]. The amino acid samples were dried by rotary evaporation and re-dissolved in 100 μ l of coupling buffer (acetonitrile: pyridine: triethylamine: H₂O [10:5:2:3]). The derivatisation reaction was started by mixing with 5 μ l of PITC and incubating at room temperature for 5 min. The reaction sample was then dried by rotary evaporation to remove excess PITC and coupling buffer. The dry derivatised samples (PTC-amino acids) were re-dissolved in water and analysed by RP-HPLC using a Gemini RP column (Phenomenex, C₁₈, particle size of 5 μ m; 250 x 10 mm) with the following gradients: 12 min linear gradient from A/B (91:9 %) to A/B (80:20 %), 1 min gradient to A/B (65:35 %), 17 min linear gradient to A/B (60:40 %), and 2 min linear gradient to A/B (5:95 %) and held at A/B (5:95 %) for 5 min, and 3 min linear gradient to A/B (91:9%) before re-equilibrating at A/B (91:9%) over 5 min. A flow rate of 2 ml/min was used and the elution of PITC-derivatised products was monitored at A₂₅₄. The following solvents were used; solvent A: 0.05 M ammonium acetate pH 6.8; solvent B: 0.1 M ammonium acetate pH 6.8 in methanol: water (80:20; v/v).

6.26. Enzyme Assays

6.26.1. Hydroxylase Activity Assay

The assay conditions are modified from that described in the study of AsnO by Strieker *et al.* (2007) [111]. The reaction was conducted in 50 mM HEPES buffer pH 7.5 containing 500 μ M amino acid substrates, 600 μ M α -KG, 100 μ M FeSO₄, 100 μ M ascorbate, 100 μ M DTT and 6.25 μ g of enzyme sample (LptL or LptJ) in a total volume of 150 μ l. Reactions were initiated by adding enzyme and incubating at 28 °C for 18 h. The reaction was stopped by boiling and removing the enzyme *via* 3 KDa membrane filtration, then derivatising by PITC.

6.26.2. Methyl Transferase Activity Assay

The assay conditions are modified from that described in the study of *O*-methyl transferase which is involved in biosynthesis of the antibiotic FK-506 [265]. The reaction was conducted in 50 mM HEPES buffer pH 7.5 containing 500 μ M amino acid substrates, 600 μ M SAM, 4 mM MgSO₄ and 5.5 μ g of enzyme sample (LptK) in a total volume of 150 μ l. Reactions were initiated by adding enzyme and incubating at 28 °C for 18 h. The reaction was stopped by boiling and removing the enzyme *via* 3 KDa membrane filtration and derivatising by PITC.

6.26.3. Alpha-mannosidase Activity Assay

The assay conditions are modified from that described in Di Palo *et al.* (2007) [246]. The ramoplanin deglycosylation reaction was carried out in 0.1 M phosphate buffer pH 7.0 containing 250 μ g of ramoplanin and 0.68 unit of α -mannosidase (Jack bean, Sigma-Aldrich, Inc) in a total volume of 200 μ l at 30 °C for 18 h. The ramoplanin aglycon was purified by butanol and dried under rotary evaporation. Additional purification was carried out by RP-HPLC (Phenomenex, C₁₈, particle size of 5 μ m; 250 x 10 mm) using the following gradients: 10 min linear gradient from A/B (95:5 %) to A/B (80:20 %), 30 min gradient to A/B (60:40 %), 5 min linear gradient to A/B (5:95 %) and held at A/B (5:95 %) for 4 min, and 3 min linear gradient to A/B (95:5 %)

before re-equilibrating at A/B (95:5 %) over 8 min. A flow rate of 3 ml/min was used and the elution of ramoplanin or ramoplanin aglycon was monitored at A_{285} . The following solvents were used: solvent A: water with 0.1 % formic acid; solvent B: acetonitrile with 0.1 % formic acid. Fractions were collected and further analysed by LC-MS. This work was done with the help from Laura Nunns and Matthew Styles.

6.26.4. Mannosyl Transferase Assay

The assay conditions are modified from that described in the study of polyprenyl phosphomannose synthase by Wehmeier *et al.* (2009) [244]. Mycelia from SfpMS17-ram29 and Sfram29ppm1 cultures were harvested by centrifugation (5 min, 5000 g), washed with MOPS buffer (50 mM MOPS pH 7.9, 5 mM β -mercaptoethanol, 10 mM $MgCl_2$, 4 °C) and stored at -80 °C. For the preparation of the membrane fractions, mycelial pellets were re-suspended in twice the pellet volume of lysis buffer at 4 °C (50 mM MOPS pH 7.9, 5 mM β -mercaptoethanol, 10 mM $MgCl_2$, protease inhibitor tablet (Roche)). The mycelia were lysed using a Cell disrupter with aperture pressure 22K p.s.i. and the cell debris was removed by centrifugation (20 min, 10,000 g, 4 °C). The membrane fragments in the supernatant were pelleted by ultracentrifugation (1 h, 100,000 g, 4°C) and re-suspended and homogenised in the lysis buffer. Mannosylation was carried out in the MOPS buffer containing 500 mg of total membrane protein, 20 mM GDP-mannose, 0.1 mM DTT and 62.5 mM ATP with either 50 μ g of enduracidin or ramoplanin aglycon in a total volume of 1 ml and incubated at 30 °C for 48 h. The reaction products were then purified by butanol and dried under a stream of nitrogen. The dry sample was re-dissolved in 50 % acetonitrile with 0.1 % formic acid and analysed by LC-MS.

6.27. Pre-Q₁ Riboswitch Functional Assay

The pre-Q₁ riboswitch (approximately 330 bp) with a *lacZ* reporter gene was integrated into the locus of *amyE* of *Bacillus* by the vector pDG1661 [170]. Its functional assay is carried out by measuring β -galactosidase activity with a high-throughput microtitreplate method modified from that described in Lynch *et al.* (2007) [266]. The *Bacillus* overnight culture was diluted 50-fold into 190 μ l NB with 5 μ g/ml Cm and grown in a 96-well plate with either 10 μ l of 50 mM ligand in DMSO or pure DMSO alone, and left to grow at 37 °C. After the OD₅₉₅ reached 0.085-0.14 (corresponding to an OD₅₉₅ of 0.3-0.5 with a 1 cm path length cuvette), the value was recorded by the plate reader (Anthos Zenyth 3100) and 200 μ l of *Bacillus* cultures in each well of 96-well plate were lysed by PopCulture reagent (Novagen, 20 μ l, 10:1, PopCulture: lysozyme [4 U/ml]) at room temperature for at least 10 min. Fifteen μ l of each lysed culture were moved to a new plate and combined with 132.5 μ l of Z buffer (60 mM Na₂HPO₄, 40 mM NaH₂PO₄, 10 mM KCl, 1 mM MgSO₄, 50 mM β -mercaptoethanol, pH 7.0). The β -galactosidase reaction was started by adding 30 μ l of *O*-nitrophenyl- β -D-galactoside (ONPG, 4 mg/ml in Z buffer) and left at 30 °C until a yellow colour developed. The reaction was then quenched by adding 75 μ l of 1 M Na₂CO₃. The length of time (**Time**) between substrate addition and quenching and the **OD₄₀₅** for β -galactosidase reaction was recorded. The Miller units were calculated using the following formula:

$$\text{Miller units} = 1000 \times (\text{OD}_{405} / (\text{OD}_{595} \times \text{Time} \times [\text{volume of cell lysate} / \text{total volume}]))$$

Chapter 7. References

1. Moellering, R.C., *NDM-1-- A Cause for Worldwide Concern*. N. Engl. J. Med., 2010. **363**(25): p. 2377-2379.
2. Miao, V., R. Brost, J. Chapple, K. She, M.-F. Gal, and R. Baltz, *The lipopeptide antibiotic A54145 biosynthetic gene cluster from Streptomyces fradiae*. J. Ind. Microbiol. Biotechnol., 2006. **33**(2): p. 129-140.
3. Hojati, Z., C. Milne, B. Harvey, L. Gordon, M. Borg, F. Flett, B. Wilkinson, P.J. Sidebottom, B.A. Rudd, M.A. Hayes, C.P. Smith, and J. Micklefield, *Structure, biosynthetic origin, and engineered biosynthesis of calcium-dependent antibiotics from Streptomyces coelicolor*. Chem. Biol., 2002. **9**(11): p. 1175-1187.
4. Miao, V., M.-F. Coeffet-LeGal, P. Brian, R. Brost, J. Penn, A. Whiting, S. Martin, R. Ford, I. Parr, M. Bouchard, C.J. Silva, S.K. Wrigley, and R.H. Baltz, *Daptomycin biosynthesis in Streptomyces roseosporus: cloning and analysis of the gene cluster and revision of peptide stereochemistry*. Microbiology, 2005. **151**(5): p. 1507-1523.
5. Yin, X. and T.M. Zabriskie, *The enduracidin biosynthetic gene cluster from Streptomyces fungicidicus*. Microbiology, 2006. **152**(10): p. 2969-2983.
6. McCafferty, D.G., P. Cudic, B.A. Frankel, S. Barkallah, R.G. Kruger, and W. Li, *Chemistry and biology of the ramoplanin family of peptide antibiotics*. Biopolymers, 2002. **66**(4): p. 261-284.
7. Raja, A., J. LaBonte, J. Lebbos, and P. Kirkpatrick, *Daptomycin*. Nat. Rev. Drug Discov., 2003. **2**(12): p. 943-944.
8. János, B., *Bioactive Microbial Metabolites*. J. Antibiot., 2005. **58**: p. 1-26.
9. Hanson, J.R., *Natural Products: The Secondary Metabolites*. 2003: Royal Society of Chemistry.
10. Demain, A.L. and S. Sanchez, *Microbial drug discovery: 80 years of progress*. J. Antibiot., 2009. **62**(1): p. 5-16.
11. Tirosh, I. and N. Barkai, *Two strategies for gene regulation by promoter nucleosomes*. Genome Res. , 2008. **18**(7): p. 1084-1091.
12. Nakamoto, T., *A unified view of the initiation of protein synthesis*. Biochem. Biophys. Res. Commun., 2006. **341**(3): p. 675-678.
13. Roth, A. and R.R. Breaker, *The Structural and Functional Diversity of Metabolite-Binding Riboswitches*. Annu. Rev. Biochem., 2009. **78**(1): p. 305-334.
14. Kieser, T., M.J. Bibb, M.J. Buttner, K.F. Chater, and D.A. Hopwood, *Practical*

- Streptomyces Genetics*. 2000, The John Innes Foundation: Norwich.
15. Barabás, G., G. Vargha, I. Szabó, A. Penyige, S. Damjanovich, J. Szöllösi, J. Matkó, T. Hirano, and A. Mátyus, *n-Alkane uptake and utilisation by Streptomyces strains*. *Antonie van Leeuwenhoek*, 2001. **79**(3): p. 269-276.
 16. Martinkova, L., B. Uhnakova, M. Patek, J. Nesvera, and V. Kren, *Biodegradation potential of the genus Rhodococcus*. *Environ. Int.*, 2009. **35**(1): p. 162-177.
 17. Walsh, C.T. and G. Wright, *Introduction: Antibiotic Resistance*. *Chem. rev.*, 2005. **105**(2): p. 391-394.
 18. Gold, H.S. and R.C. Moellering, *Antimicrobial-Drug Resistance*. *N. Engl. J. Med.*, 1996. **335**(19): p. 1445-1453.
 19. Eiland, E.H., III and D. Gatlin, *Forecast of Antibiotic Development in an Era of Increasing Bacterial Resistance*. *J. Pharm. Pract.*, 2008. **21**(5): p. 313-318.
 20. Koehn, F.E. and G.T. Carter, *The evolving role of natural products in drug discovery*. *Nat. Rev. Drug Discov.*, 2005. **4**(3): p. 206-220.
 21. Walsh, C.T. and M.A. Fischbach, *Natural Products Version 2.0: Connecting Genes to Molecules*. *J. Am. Chem. Soc.*, 2010.
 22. Gross, H., *Strategies to unravel the function of orphan biosynthesis pathways: recent examples and future prospects*. *Appl. Microbiol. Biotechnol.*, 2007. **75**(2): p. 267-277.
 23. Fleischmann, R.D., M.D. Adams, O. White, R.A. Clayton, E.F. Kirkness, A.R. Kerlavage, C.J. Bult, J.F. Tomb, B.A. Dougherty, J.M. Merrick, and e. al, *Whole-genome random sequencing and assembly of Haemophilus influenzae Rd*. *Science*, 1995. **269**(5223): p. 496-512.
 24. Corre, C. and G.L. Challis, *Heavy Tools for Genome Mining*. *Chem. Biol.*, 2007. **14**(1): p. 7-9.
 25. Wilkinson, B. and J. Micklefield, *Mining and engineering natural-product biosynthetic pathways*. *Nat. Chem. Biol.*, 2007. **3**(7): p. 379-386.
 26. Fischbach, M.A. and C.T. Walsh, *Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: Logic, machinery, and mechanisms*. *Chem. rev.*, 2006. **106**(8): p. 3468-3496.
 27. Rokem, J.S., A.E. Lantz, and J. Nielsen, *Systems biology of antibiotic production by microorganisms*. *Nat. Prod. Rep.*, 2007. **24**(6): p. 1262-1287.
 28. Neu, H.C., *The Crisis in Antibiotic Resistance*. *Science*, 1992. **257**(5073): p. 1064-1073.
 29. Lambert, P.A., *Bacterial resistance to antibiotics: Modified target sites*. *Adv. Drug Deliv. Rev.*, 2005. **57**(10): p. 1471-1485.
 30. de la Cruz, F. and J. Davies, *Horizontal gene transfer and the origin of species:*

- lessons from bacteria*. Trends Microbiol., 2000. **8**(3): p. 128-133.
31. Eggert, U.S., N. Ruiz, B.V. Falcone, A.A. Branstrom, R.C. Goldman, T.J. Silhavy, and D. Kahne, *Genetic Basis for Activity Differences Between Vancomycin and Glycolipid Derivatives of Vancomycin*. Science, 2001. **294**(5541): p. 361-364.
 32. Mechali, M., *Eukaryotic DNA replication origins: many choices for appropriate answers*. Nat. Rev. Mol. Cell Biol., 2010. **11**(10): p. 728-738.
 33. Cunningham, E.L. and J.M. Berger, *Unraveling the early steps of prokaryotic replication*. Curr. Opin. Struct. Biol., 2005. **15**(1): p. 68-76.
 34. Hooper, D.C., *Mechanisms of Action of Antimicrobials: Focus on Fluoroquinolones*. Clin. Infect. Dis., 2001. **32**(s1): p. S9-S15.
 35. Vesely, J.J., F.D. Pien, and B.C. Pien, *Rifampin, a useful drug for nonmycobacterial infections*. Pharmacotherapy, 1998. **18**: p. 345-357.
 36. Dozzo, P. and H.E. Moser, *New aminoglycoside antibiotics*. Expert opin. ther. pat., 2010. **20**(10): p. 1321-1341.
 37. Zakeri, B. and G.D. Wright, *Chemical biology of tetracycline antibiotics*. Biochem. Cell Biol., 2008. **86**(2): p. 124-136.
 38. Iyer, R. and S.K. Tomar, *Folate: A Functional Food Constituent*. J. Food Sci., 2009. **74**(9): p. R114-R122.
 39. Olliaro, P., *Mode of action and mechanisms of resistance for antimalarial drugs*. Pharmacol. therapeut., 2001. **89**(2): p. 207-219.
 40. Bermingham, A. and J.P. Derrick, *The folic acid biosynthesis pathway in bacteria: evaluation of potential for antibacterial drug discovery*. BioEssays, 2002. **24**(7): p. 637-648.
 41. Breukink, E. and B. de Kruijff, *Lipid II as a target for antibiotics*. Nat. Rev. Drug Discov., 2006. **5**(4): p. 321-323.
 42. Silverman, J.A., N.G. Perlmutter, and H.M. Shapiro, *Correlation of Daptomycin Bactericidal Activity and Membrane Depolarization in Staphylococcus aureus*. Antimicrob. Agents Chemother., 2003. **47**(8): p. 2538-2544.
 43. Eschenburg, S., M. Priestman, and E. Schonbrunn, *Evidence That the Fosfomycin Target Cys115 in UDP-N-acetylglucosamine Enolpyruvyl Transferase (MurA) Is Essential for Product Release*. J. Biol. Chem., 2005. **280**(5): p. 3757-3763.
 44. Fang, X., K. Tiyanont, Y. Zhang, J. Wanner, D. Boger, and S. Walker, *The mechanism of action of ramoplanin and enduracidin*. Mol. Biosyst., 2006. **2**(1): p. 69-76.
 45. Ruzin, A., G. Singh, A. Severin, Y. Yang, R.G. Dushin, A.G. Sutherland, A.

- Minnick, M. Greenstein, M.K. May, D.M. Shlaes, and P.A. Bradford, *Mechanism of Action of the Mannopectimycins, a Novel Class of Glycopeptide Antibiotics Active against Vancomycin-Resistant Gram-Positive Bacteria*. *Antimicrob. Agents Chemother.*, 2004. **48**(3): p. 728-738.
46. Aharonowitz, Y., G. Cohen, and J.F. Martin, *Penicillin and Cephalosporin Biosynthetic Genes: Structure, Organization, Regulation, and Evolution*. *Annu. Rev. Microbiol.*, 1992. **46**(1): p. 461-495.
47. Cheng, T.-J.R., M.-T. Sung, H.-Y. Liao, Y.-F. Chang, C.-W. Chen, C.-Y. Huang, L.-Y. Chou, Y.-D. Wu, Y.-H. Chen, Y.-S.E. Cheng, C.-H. Wong, C. Ma, and W.-C. Cheng, *Domain requirement of moenomycin binding to bifunctional transglycosylases and development of high-throughput discovery of antibiotics*. *Proc. Natl. Acad. Sci. U.S.A.*, 2008. **105**(2): p. 431-436.
48. van Heijenoort, J., *Lipid Intermediates in the Biosynthesis of Bacterial Peptidoglycan*. *Microbiol. Mol. Biol. Rev.*, 2007. **71**(4): p. 620-635.
49. Ahmed, B., E.T. Amy, D.H.B. Timothy, and M.-L. Dominique, *The biosynthesis of peptidoglycan lipid-linked intermediates*. *FEMS Microbiol. Rev.*, 2008. **32**(2): p. 208-233.
50. Barreteau, H., A. Kovač, A. Boniface, M. Sova, S. Gobec, and D. Blanot, *Cytoplasmic steps of peptidoglycan biosynthesis*. *FEMS Microbiol. Rev.*, 2008. **32**(2): p. 168-207.
51. Brandish, P.E., K.I. Kimura, M. Inukai, R. Southgate, J.T. Lonsdale, and T.D. Bugg, *Modes of action of tunicamycin, liposidomycin B, and mureidomycin A: inhibition of phospho-N-acetylmuramyl-pentapeptide translocase from Escherichia coli*. *Antimicrob. Agents Chemother.*, 1996. **40**(7): p. 1640-1644.
52. Martin, J.-F. and P. Liras, *Engineering of regulatory cascades and networks controlling antibiotic biosynthesis in Streptomyces*. *Curr. Opin. Microbiol.* **13**(3): p. 263-273.
53. Bibb, M.J., *Regulation of secondary metabolism in streptomycetes*. *Curr. Opin. Microbiol.*, 2005. **8**(2): p. 208-215.
54. Arias, P., M.A. Fernandez-Moreno, and F. Malpartida, *Characterization of the Pathway-Specific Positive Transcriptional Regulator for Actinorhodin Biosynthesis in Streptomyces coelicolor A3(2) as a DNA-Binding Protein*. *J. Bacteriol.*, 1999. **181**(22): p. 6958-6968.
55. Takano, E., H.C. Gramajo, E. Strauch, N. Andres, J. White, and M.J. Bibb, *Transcriptional regulation of the redD transcriptional activator gene accounts for growth-phase-dependent production of the antibiotic undecylprodigiosin in Streptomyces coelicolor A3(2)*. *Mol. Microbiol.*, 1992. **6**(19): p. 2797-2804.
56. Ryding, N.J., T.B. Anderson, and W.C. Champness, *Regulation of the*

- Streptomyces coelicolor* Calcium-Dependent Antibiotic by *absA*, Encoding a Cluster-Linked Two-Component System. J. Bacteriol., 2002. **184**(3): p. 794-805.
57. McKenzie, N.L. and J.R. Nodwell, *Phosphorylated AbsA2 Negatively Regulates Antibiotic Production in Streptomyces coelicolor through Interactions with Pathway-Specific Regulatory Gene Promoters*. J. Bacteriol., 2007. **189**(14): p. 5284-5292.
58. Sanchez, S. and A.L. Demain, *Metabolic regulation of fermentation processes*. Enzyme Microb. Tech., 2002. **31**(7): p. 895-906.
59. Ruiz, B., A. Chavez, A. Forero, Y. Garcia-Huante, A. Romero, M. Sanchez, D. Rocha, B. Sanchez, R. Rodriguez-Sanoja, S. Sanchez, and E. Langley, *Production of microbial secondary metabolites: Regulation by the carbon source*. Crit. Rev. Microbiol., 2010. **36**(2): p. 146-167.
60. Marahiel, M.A., L.O. Essen, and A.H. David, *Chapter 13 Nonribosomal Peptide Synthetases: Mechanistic and Structural Aspects of Essential Domains*, in *Meth. Enzymol.* 2009, Academic Press. p. 337-351.
61. Grunewald, J. and M.A. Marahiel, *Chemoenzymatic and Template-Directed Synthesis of Bioactive Macrocyclic Peptides*. Microbiol. Mol. Biol. Rev., 2006. **70**(1): p. 121-146.
62. Marahiel, M.A., T. Stachelhaus, and H.D. Mootz, *Modular Peptide Synthetases Involved in Nonribosomal Peptide Synthesis*. Chem. Rev., 1997. **97**(7): p. 2651-2674.
63. Schwarzer, D., R. Finking, and M.A. Marahiel, *Nonribosomal peptides: from genes to products*. Nat. Prod. Rep., 2003. **20**(3): p. 275-287.
64. Mootz, H.D., D. Schwarzer, and M.A. Marahiel, *Ways of Assembling Complex Natural Products on Modular Nonribosomal Peptide Synthetases*. Chembiochem, 2002. **3**(6): p. 490-504.
65. Keating, T.A., C.G. Marshall, and C.T. Walsh, *Reconstitution and characterization of the Vibrio cholerae vibriobactin synthetase from VibB, VibE, VibF, and VibH*. Biochemistry, 2000 **39**: p. 15522-15530.
66. Stachelhaus, T. and M.A. Marahiel, *Modular Structure of Peptide Synthetases Revealed by Dissection of the Multifunctional Enzyme GrsA*. J. Biol. Chem., 1995. **270**(11): p. 6163-6169.
67. Turgay, K., M. Krause, and M.A. Marahiel, *Four homologous domains in the primary structure of GrsB are related to domains in a superfamily of adenylate-forming enzymes*. Mol. Microbiol., 1992. **6**(4): p. 529-546.
68. Arnez, J.G. and D. Moras, *Structural and functional considerations of the*

- aminoacylation reaction*. Trends Biochem. Sci., 1997. **22**(6): p. 211-216.
69. Conti, E., T. Stachelhaus, M.A. Marahiel, and P. Brick, *Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S*. EMBO J., 1997. **16**(14): p. 4174-4183.
70. Stachelhaus, T., H.D. Mootz, and M.A. Marahiel, *The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases*. Chem. Biol., 1999. **6**(8): p. 493-505.
71. Challis, G.L., J. Ravel, and C.A. Townsend, *Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains*. Chem. Biol., 2000. **7**(3): p. 211-224.
72. Eppelmann, K., T. Stachelhaus, and M.A. Marahiel, *Exploitation of the Selectivity-Conferring Code of Nonribosomal Peptide Synthetases for the Rational Design of Novel Peptide Antibiotics*. Biochemistry, 2002. **41**(30): p. 9718-9726.
73. Sieber, S.A. and M.A. Marahiel, *Molecular Mechanisms Underlying Nonribosomal Peptide Synthesis: Approaches to New Antibiotics*. Chem. Rev., 2005. **105**(2): p. 715-738.
74. Lambalot, R.H., A.M. Gehring, R.S. Flugel, P. Zuber, M. LaCelle, M.A. Marahiel, R. Reid, C. Khosla, and C.T. Walsh, *A new enzyme superfamily -- the phosphopantetheinyl transferases*. Chem. Biol., 1996. **3**(11): p. 923-936.
75. Stein, T., J. Vater, V. Kruff, A. Otto, B. Wittmann-Liebold, P. Franke, M. Panico, R. McDowell, and H.R. Morris, *The multiple carrier model of nonribosomal peptide biosynthesis at modular multienzymatic templates*. J. Biol. Chem., 1996. **271**(26): p. 15428-15435.
76. Stachelhaus, T., A. Huser, and M.A. Marahiel, *Biochemical characterization of peptidyl carrier protein (PCP), the thiolation domain of multifunctional peptide synthetases*. Chem. Biol., 1996. **3**(11): p. 913-921.
77. Stachelhaus, T., H.D. Mootz, V. Bergendahl, and M.A. Marahiel, *Peptide Bond Formation in Nonribosomal Peptide Biosynthesis*. J. Biol. Chem., 1998. **273**(35): p. 22773-22781.
78. Belshaw, P.J., C.T. Walsh, and T. Stachelhaus, *Aminoacyl-CoAs as Probes of Condensation Domain Selectivity in Nonribosomal Peptide Synthesis*. Science, 1999. **284**(5413): p. 486-489.
79. Kohli, R.M. and C.T. Walsh, *Enzymology of acyl chain macrocyclization in natural product biosynthesis*. Chem. Commun., 2003(3): p. 297-307.
80. Doekel, S. and M.A. Marahiel, *Biosynthesis of Natural Products on Modular Peptide Synthetases*. Metab. Eng., 2001. **3**(1): p. 64-77.
81. Keating, T.A., D.E. Ehmann, R.M. Kohli, C.G. Marshall, J.W. Trauger, and

- C.T. Walsh, *Chain Termination Steps in Nonribosomal Peptide Synthetase Assembly Lines: Directed Acyl-S-Enzyme Breakdown in Antibiotic and Siderophore Biosynthesis*. *ChemBiochem*, 2001. **2**(2): p. 99-107.
82. Trauger, J.W., R.M. Kohli, H.D. Mootz, M.A. Marahiel, and C.T. Walsh, *Peptide cyclization catalysed by the thioesterase domain of tyrocidine synthetase*. *Nature*, 2000. **407**(6801): p. 215-218.
83. Vallari, D.S., S. Jackowski, and C.O. Rock, *Regulation of pantothenate kinase by coenzyme A and its thioesters*. *J. Biol. Chem.*, 1987. **262**(6): p. 2468-2471.
84. Quadri, L.E.N., P.H. Weinreb, M. Lei, M.M. Nakano, P. Zuber, and C.T. Walsh, *Characterization of Sfp, a Bacillus subtilis Phosphopantetheinyl Transferase for Peptidyl Carrier Protein Domains in Peptide Synthetases*. *Biochemistry*, 1998. **37**(6): p. 1585-1595.
85. Schwarzer, D., H.D. Mootz, U. Linne, and M.A. Marahiel, *Regeneration of misprimed nonribosomal peptide synthetases by type II thioesterases*. *Proc. Natl. Acad. Sci. U.S.A.*, 2002. **99**(22): p. 14083-14088.
86. Walsh, C.T., H. Chen, T.A. Keating, B.K. Hubbard, H.C. Losey, L. Luo, C.G. Marshall, D.A. Miller, and H.M. Patel, *Tailoring enzymes that modify nonribosomal peptides during and after chain elongation on NRPS assembly lines*. *Curr. Opin. Chem. Biol.*, 2001. **5**(5): p. 525-534.
87. Konz, D. and M.A. Marahiel, *How do peptide synthetases generate structural diversity?* *Chem. Biol.*, 1999. **6**: p. R39-R48.
88. Stachelhaus, T. and C.T. Walsh, *Mutational Analysis of the Epimerization Domain in the Initiation Module PheATE of Gramicidin S Synthetase*. *Biochemistry*, 2000. **39**(19): p. 5775-5787.
89. Stein, T., B. Kluge, J. Vater, P. Franke, A. Otto, and B. Wittmann-Liebold, *Gramicidin S synthetase I (phenylalanine racemase), a prototype of amino acid racemases containing the cofactor 4'-phosphopantetheine*. *Biochemistry*, 1995. **34**(14): p. 4633-4642.
90. Linne, U. and M.A. Marahiel, *Control of Directionality in Nonribosomal Peptide Synthesis: Role of the Condensation Domain in Preventing Misinitiation and Timing of Epimerization*. *Biochemistry*, 2000. **39**(34): p. 10439-10447.
91. Hoffmann, K., E. Schneider-Scherzer, H. Kleinkauf, and R. Zocher, *Purification and characterization of eucaryotic alanine racemase acting as key enzyme in cyclosporin biosynthesis*. *J. Biol. Chem.*, 1994. **269**(17): p. 12710-12714.
92. Weber, G., K. Schörgendorfer, E. Schneider-Scherzer, and E. Leitner, *The peptide synthetase catalyzing cyclosporine production in Tolypocladium*

- niveum* is encoded by a giant 45.8-kilobase open reading frame. *Curr. Genet.*, 1994. **26**(2): p. 120-125.
93. Chen, H., M.G. Thomas, S.E. O'Connor, B.K. Hubbard, M.D. Burkart, and C.T. Walsh, *Aminoacyl-S-enzyme intermediates in beta-hydroxylations and alpha,beta-desaturations of amino acids in peptide antibiotics*. *Biochemistry*, 2001. **40**(39): p. 11651-11659.
94. Li, T.-L., F. Huang, S.F. Haydock, T. Mironenko, P.F. Leadlay, and J.B. Spencer, *Biosynthetic Gene Cluster of the Glycopeptide Antibiotic Teicoplanin: Characterization of Two Glycosyltransferases and the Key Acyltransferase*. *Chem. Biol.*, 2004. **11**(1): p. 107-119.
95. Neary, J.M., A. Powell, L. Gordon, C. Milne, F. Flett, B. Wilkinson, C.P. Smith, and J. Micklefield, *An asparagine oxygenase (AsnO) and a 3-hydroxyasparaginyl phosphotransferase (HasP) are involved in the biosynthesis of calcium-dependent lipopeptide antibiotics*. *Microbiology*, 2007. **153**(3): p. 768-776.
96. Chen, H., B.K. Hubbard, S.E. O'Connor, and C.T. Walsh, *Formation of [beta]-Hydroxy Histidine in the Biosynthesis of Nikkomycin Antibiotics*. *Chem. Biol.*, 2002. **9**(1): p. 103-112.
97. van Wageningen, A.M., P.N. Kirkpatrick, D.H. Williams, B.R. Harris, J.K. Kershaw, N.J. Lennard, M. Jones, S.J.M. Jones, and P.J. Solenberg, *Sequencing and analysis of genes involved in the biosynthesis of a vancomycin group antibiotic*. *Chem. Biol.*, 1998. **5**(3): p. 155-162.
98. Hammond, S.J., M.P. Williamson, D.H. Williams, L.D. Boeck, and G.G. Marconi, *On the biosynthesis of the antibiotic vancomycin*. *J. Chem. Soc.*, 1982(6): p. 344-346.
99. Chen, H. and C.T. Walsh, *Coumarin formation in novobiocin biosynthesis: beta-hydroxylation of the aminoacyl enzyme tyrosyl-S-NovH by a cytochrome P450 NovI*. *Chem. Biol.*, 2001. **8**(4): p. 301-312.
100. Wang, Z.-X., S.-M. Li, and L. Heide, *Identification of the Coumermycin A1 Biosynthetic Gene Cluster of Streptomyces rishiriensis DSM 40489*. *Antimicrob. Agents Chemother.*, 2000. **44**(11): p. 3040-3048.
101. Kato, T., H. Hinoo, Y. Terui, J. Kikuchi, and J. Shoji, *The structures of katanosins A and B*. *J. Antibiot.*, 1988. **41**: p. 719-725.
102. Traber, R., C. Keller-Juslén, H.-R. Loosli, M. Kuhn, and A. Von Wartburg, *Cyclopeptid-Antibiotika aus Aspergillus-Arten. Struktur der Echinocandine C und D*. *Helv. Chim. Acta*, 1979. **62**(4): p. 1252-1267.
103. Kempter, C., D. Kaiser, S. Haag, G. Nicholson, V. Gnau, T. Walk, K.H. Gierling, H. Decker, H. Zähner, G. Jung, and J.W. Metzger, *CDA:*

- Calcium-Dependent Peptide Antibiotics from Streptomyces coelicolor A3(2) Containing Unusual Residues*. Angew. Chem. Int. Ed. Engl., 1997. **36**(5): p. 498-501.
104. Shoji, J., H. Hino, T. Katayama, Y. Nakagawa, Y. Ikenishi, K. Iwatani, and T. Yoshida, *Structures of new peptide antibiotics, plusbacins A1-A4 and B1-B4*. J. Antibiot., 1992. **45**: p. 824-831.
105. Singh, G.M., P.D. Fortin, A. Koglin, and C.T. Walsh, *β -Hydroxylation of the Aspartyl Residue in the Phytotoxin Syringomycin E: Characterization of Two Candidate Hydroxylases AspH and SyrP in Pseudomonas syringae*. Biochemistry, 2008. **47**(43): p. 11310-11320.
106. Shen, B., L. Du, C. Sanchez, D.J. Edwards, M. Chen, and J.M. Murrell, *The biosynthetic gene cluster for the anticancer drug bleomycin from Streptomyces verticillus ATCC15003 as a model for hybrid peptide-polyketide natural product biosynthesis*. J. Ind. Microbiol. Biotechnol., 2001. **27**(6): p. 378-385.
107. Yoshikawa, Y., K. Ikai, Y. Umeda, A. Ogawa, K. Takesako, I. Kato, and H. Naganawa, *Isolation, structures, and antifungal activities of new aureobasidins*. J. Antibiot., 1993. **46**: p. 1347-1354.
108. Gould, S.J., J. Lee, and J. Wityak, *Biosynthesis of streptothricin F : 7. The fate of the arginine hydrogens*. Bioorg. Chem., 1991. **19**(3): p. 333-350.
109. Hausinger, R.P., *Fe (II)/ α -Ketoglutarate-Dependent Hydroxylases and Related Enzymes*. Crit. Rev. Biochem. Mol. Biol., 2004. **39**(1): p. 21-68.
110. Prescott, A.G. and M.D. Lloyd, *The iron(II) and 2-oxoacid-dependent dioxygenases and their role in metabolism*. Nat. Prod. Rep., 2000. **17**(4): p. 367-383.
111. Strieker, M., F. Kopp, C. Mahlert, L.O. Essen, and M.A. Marahiel, *Mechanistic and Structural Basis of Stereospecific C β -Hydroxylation in Calcium-Dependent Antibiotic, a Daptomycin-Type Lipopeptide*. ACS Chem. Biol., 2007. **2**: p. 187-196.
112. Urlacher, V.B. and S. Eiben, *Cytochrome P450 monooxygenases: perspectives for synthetic application*. Trends Biotechnol., 2006. **24**(7): p. 324-330.
113. Bischoff, D., B. Bister, M. Bertazzo, V. Pfeifer, E. Stegmann, G.J. Nicholson, S. Keller, S. Pelzer, W. Wohlleben, and R.D. Süssmuth, *The Biosynthesis of Vancomycin-Type Glycopeptide Antibiotics—A Model for Oxidative Side-Chain Cross-Linking by Oxygenases Coupled to the Action of Peptide Synthetases*. ChemBiochem, 2005. **6**(2): p. 267-272.
114. Schubert, H.L., R.M. Blumenthal, and X. Cheng, *Many paths to methyltransfer: a chronicle of convergence*. Trends Biochem. Sci., 2003. **28**(6): p. 329-335.
115. Martin, J.L. and F.M. McMillan, *SAM (dependent) I AM: the*

- S-adenosylmethionine-dependent methyltransferase fold*. *Curr. Opin. Struct. Biol.*, 2002. **12**(6): p. 783-793.
116. Apweiler, R., A. Bairoch, C.H. Wu, W.C. Barker, B. Boeckmann, S. Ferro, E. Gasteiger, H. Huang, R. Lopez, M. Magrane, M.J. Martin, D.A. Natale, C. O'Donovan, N. Redaschi, and L.S. Yeh, *UniProt: the Universal Protein knowledgebase*. *Nucleic Acids Res.*, 2004. **32**: p. D115-119.
117. O'Donovan, C., M.J. Martin, A. Gattiker, E. Gasteiger, A. Bairoch, and R. Apweiler, *High-quality protein knowledge resource: SWISS-PROT and TrEMBL*. *Brief. Bioinform.*, 2002. **3**(3): p. 275-284.
118. Mahlert, C., F. Kopp, J. Thirlway, J. Micklefield, and M.A. Marahiel, *Stereospecific Enzymatic Transformation of α -Ketoglutarate to (2S,3R)-3-Methyl Glutamate during Acidic Lipopeptide Biosynthesis*. *J. Am. Chem. Soc.*, 2007. **129**(39): p. 12011-12018.
119. Huang, Y.-T., S.-Y. Lyu, P.-H. Chuang, N.-S. Hsu, Y.-S. Li, H.-C. Chan, C.-J. Huang, Y.-C. Liu, C.-J. Wu, W.-B. Yang, and T.-L. Li, *In vitro Characterization of Enzymes Involved in the Synthesis of Nonproteinogenic Residue (2S,3S)- β -Methylphenylalanine in Glycopeptide Antibiotic Mannopeptimycin*. *Chembiochem*, 2009. **10**(15): p. 2480-2487.
120. Blasiak, L.C. and C.L. Drennan, *Structural Perspective on Enzymatic Halogenation*. *Acc. Chem. Res.*, 2008. **42**(1): p. 147-155.
121. Vaillancourt, F.H., E. Yeh, D.A. Vosburg, S. Garneau-Tsodikova, and C.T. Walsh, *Nature's Inventory of Halogenation Catalysts: Oxidative Strategies Predominate*. *Chem. Rev.*, 2006. **106**(8): p. 3364-3378.
122. Winter, J.M. and B.S. Moore, *Exploring the Chemistry and Biology of Vanadium-dependent Haloperoxidases*. *J. Biol. Chem.*, 2009. **284**(28): p. 18577-18581.
123. Wagner, C., M. El Omari, and G.M. König, *Biohalogenation: Nature's Way to Synthesize Halogenated Metabolites*. *J. Nat. Prod.*, 2009. **72**(3): p. 540-553.
124. Murphy, C.D., *Recent developments in enzymatic chlorination*. *Nat. Prod. Rep.*, 2006. **23**: p. 147-152.
125. Dong, C., S. Flecks, S. Unversucht, C. Haupt, K.-H. van Pee, and J.H. Naismith, *Tryptophan 7-Halogenase (PrnA) Structure Suggests a Mechanism for Regioselective Chlorination*. *Science*, 2005. **309**(5744): p. 2216-2219.
126. Keller, S., T. Wage, K. Hohaus, M. Holzer, E. Eichhorn, and K.-H. van Pee, *Purification and Partial Characterization of Tryptophan 7-Halogenase (PrnA) from Pseudomonas fluorescens*. *Angew. Chem. Int. Ed. Engl.*, 2000. **39**(13): p. 2300-2302.
127. Blasiak, L.C., A. Koglin, C.L. Drennan, and C.T. Walsh, *Chlorination by a*

- Long-Lived Intermediate in the Mechanism of Flavin-Dependent Halogenases.* Biochemistry, 2007. **46**(5): p. 1284-1292.
128. Vaillancourt, F.H., J. Yin, and C.T. Walsh, *SyrB2 in syringomycin E biosynthesis is a nonheme FeII α -ketoglutarate- and O₂-dependent halogenase.* Proc. Natl. Acad. Sci. U.S.A., 2005. **102**(29): p. 10111-10116.
129. Blasiak, L.C., F.H. Vaillancourt, C.T. Walsh, and C.L. Drennan, *Crystal structure of the non-haem iron halogenase SyrB2 in syringomycin biosynthesis.* Nature, 2006. **440**(7082): p. 368-371.
130. Lommel, M. and S. Strahl, *Protein O-mannosylation: Conserved from bacteria to humans.* Glycobiology, 2009. **19**(8): p. 816-828.
131. Counter, F.T., N.E. Allen, D.S. Fukuda, J.N. Hobbs, J. Ott, P.W. Ensminger, J.S. Mynderse, D.A. Prestone, and C.Y.E. Wu, *A54145 A NEW LIPOPEPTIDE ANTIBIOTIC COMPLEX: MICROBIOLOGICAL EVALUATION.* J. Antibiot., 1990. **43**(6): p. 616-622.
132. Baltz, R.H., V. Miao, and S.K. Wrigley, *Natural products to drugs: daptomycin and related lipopeptide antibiotics.* Nat. Prod. Rep., 2005. **22**(6): p. 717-741.
133. Kopp, F., J. Grunewald, C. Mahlert, and M.A. Marahiel, *Chemoenzymatic Design of Acidic Lipopeptide Hybrids: New Insights into the Structure-Activity Relationship of Daptomycin and A54145.* Biochemistry, 2006. **45**(35): p. 10474-10481.
134. Uguru, G.C., C. Milne, M. Borg, F. Flett, C.P. Smith, and J. Micklefield, *Active-Site Modifications of Adenylation Domains Lead to Hydrolysis of Upstream Nonribosomal Peptidyl Thioester Intermediates.* J. Am. Chem. Soc., 2004. **126**(16): p. 5032-5033.
135. Nguyen, K.T., D. Kau, J.Q. Gu, P. Brian, S.K. Wrigley, R.H. Baltz, and V. Miao, *A glutamic acid 3-methyltransferase encoded by an accessory gene locus important for daptomycin biosynthesis in Streptomyces roseosporus.* Mol. Microbiol., 2006. **61**(5): p. 1294-307.
136. Rudd, B.A.M., *Genetics of pigmented secondary metabolites in Streptomyces coelicolor* 1978, PhD thesis University of East Anglia: Norwich, UK.
137. Hopwood, D.A. and H.M. Wright, *CDA is a New Chromosomally-determined Antibiotic from Streptomyces coelicolor A3(2).* J. Gen. Microbiol. , 1983. **129**: p. 3575-3579.
138. Bentley, S.D., K.F. Chater, A.M. Cerdeno-Tarraga, G.L. Challis, N.R. Thomson, K.D. James, D.E. Harris, M.A. Quail, H. Kieser, D. Harper, A. Bateman, S. Brown, G. Chandra, C.W. Chen, M. Collins, A. Cronin, A. Fraser, A. Goble, J. Hidalgo, T. Hornsby, S. Howarth, C.H. Huang, T. Kieser, L. Larke, L. Murphy, K. Oliver, S. O'Neil, E. Rabbinowitsch, M.A. Rajandream, K.

- Rutherford, S. Rutter, K. Seeger, D. Saunders, S. Sharp, R. Squares, S. Squares, K. Taylor, T. Warren, A. Wietzorrek, J. Woodward, B.G. Barrell, J. Parkhill, and D.A. Hopwood, *Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2)*. *Nature*, 2002. **417**(6885): p. 141-147.
139. Tally, F.P. and M.F. DeBruin, *Development of daptomycin for Gram-positive infections*. *J. Antimicrob. Chemother.*, 2000. **46**(4): p. 523-526.
140. Hamill, R.L. and M. Hoehn, *U.S. Patent 4399067*. 1980.
141. Huber, F.M., R.L. Pieper, and A.J. Tietz, *The formation of daptomycin by supplying decanoic acid to Streptomyces roseosporus cultures producing the antibiotic complex A21978C*. *J. Biotechnol.*, 1988. **7**(4): p. 283-292.
142. Penn, J., X. Li, A. Whiting, M. Latif, T. Gibson, C. Silva, P. Brian, J. Davies, V. Miao, S. Wrigley, and R. Baltz, *Heterologous production of daptomycin in Streptomyces lividans*. *J. Ind. Microbiol. Biotechnol.*, 2006. **33**(2): p. 121-128.
143. Fukuda, D.S., *A54145, A NEW LIPOPEPTIDE ANTIBIOTIC COMPLEX: ISOLATION AND CHARACTERIZATION*. *J. Antibiot.*, 1990. **43**(6): p. 594-600.
144. Boeck, L.D., H.R. Papiska, R.W. Wetzel, J.S. Mynderse, D.S. Fukuda, F.P. Mertz, and D.M. Berry, *A54145, A new lipopeptide antibiotic complex: discovery, taxonomy, fermentation and hplc*. *J. Antibiot.*, 1990. **43**(6): p. 587-593.
145. Higashide, E., K. Hatano, M. Shibata, and K. Nakazawa, *Enduracidin, a new antibiotic. I. Streptomyces fungicidicus No. B5477, an enduracidin producing organism*. *J. Antibiot.*, 1968. **21**(2): p. 126-137.
146. Asai, M., M. Muroi, N. Sugita, H. Kawashima, and K. Mizuno, *Enduracidin, a new antibiotic. II. Isolation and characterization*. *J. Antibiot.*, 1968. **21**(2): p. 138-146.
147. Cavalleri, B., H. Pagani, G. Volpe, E. Selva, and F. Parenti, *A-16686, a new antibiotic from Actinoplanes. I. Fermentation, isolation and preliminary physico-chemical characteristics*. *J. Antibiot.*, 1984. **37**(4): p. 309-317.
148. Castiglione, F., A. Marazzi, M. Meli, and G. Colombo, *Structure elucidation and 3D solution conformation of the antibiotic enduracidin determined by NMR spectroscopy and molecular dynamics*. *Magn. Reson. Chem.*, 2005. **43**(8): p. 603-610.
149. Brumfitt, W., M.R. Salton, and J.M. Hamilton-Miller, *Nisin, alone and combined with peptidoglycan-modulating antibiotics: activity against methicillin-resistant Staphylococcus aureus and vancomycin-resistant enterococci*. *J. Antimicrob. Chemother.*, 2002. **50**(5): p. 731-734.
150. Farver, D.K., D.D. Hedge, and S.C. Lee, *Ramoplanin: a lipoglycopeptide*

- antibiotic*. Ann. Pharmacother., 2005. **39**(5): p. 863-868.
151. Petersen, P.J., T.Z. Wang, R.G. Dushin, and P.A. Bradford, *Comparative In Vitro Activities of AC98-6446, a Novel Semisynthetic Glycopeptide Derivative of the Natural Product Mannopectimycin alpha, and Other Antimicrobial Agents against Gram-Positive Clinical Isolates*. Antimicrob. Agents Chemother., 2004. **48**(3): p. 739-746.
152. Tsuchiya, K., *ENDURACIDIN, A NEW ANTIBIOTIC. III*. J. Antibiot., 1968. **21**(2): p. 147-153.
153. Schmidt, J.W., A. Greenough, M. Burns, A.E. Luteran, and D.G. McCafferty, *Generation of ramoplanin-resistant Staphylococcus aureus*. FEMS Microbiology Letters, 2010. **310**(2): p. 104-111.
154. Hori, M., *Enduracidin, a New Antibiotic. VII. Primary Structure of the Peptide Moiety*. Chem. Pharm. Bull., 1973. **21**(6): p. 1175-1183.
155. Iwasaki, H., *Enduracidin, a New Antibiotic. VIII. Structures of Enduracidins A and B*. Chem. Pharm. Bull., 1973. **21**(6): p. 1184-1191.
156. Peromet, M., E. Schoutens, and E. Yourassowsky, *Clinical and Microbiological Study of Enduracidin in Infections Due to Methicillin-Resistant Strains of Staphylococcus aureus*. Chemotherapy, 1973. **19**(1): p. 53-61.
157. Balibar, C.J., F.H. Vaillancourt, and C.T. Walsh, *Generation of D Amino Acid Residues in Assembly of Arthrofactin by Dual Condensation/Epimerization Domains*. Chem. Biol., 2005. **12**(11): p. 1189-1200.
158. Gastaldo, L., R. Ciabatti, F. Assi, E. Restelli, J.K. Kettenring, L.F. Zerilli, G. Romano, M. Denaro, and B. Cavalleri, *Isolation, structure determination and biological activity of A-16686 factors A' 1, A' 2 and A' 3 glycolipodepsipeptide antibiotics*. J. Ind. Microbiol., 1992. **11**(1): p. 13-18.
159. Farnet, C.M., E. Zazopoulos, and A. Staffa, In PCT Int Appl (Ecopia Biosciences Inc., Can.), 2002. **Wo0231155**: p. 212.
160. Dambach, M.D. and W.C. Winkler, *Expanding roles for metabolite-sensing regulatory RNAs*. Curr. Opin. Microbiol., 2009. **12**(2): p. 161-169.
161. Garst, A.D. and R.T. Batey, *A switch in time: Detailing the life of a riboswitch*. Biochim. Biophys. Acta., 2009. **1789**(9-10): p. 584-591.
162. Wachter, A., *Riboswitch-mediated control of gene expression in eukaryotes*. RNA Biology, 2010. **7**: p. 67-76.
163. Blouin, S., J. Mulhbach, J.C. Penedo, and D.A. Lafontaine, *Riboswitches: Ancient and Promising Genetic Regulators*. ChemBioChem, 2009. **10**(3): p. 400-416.
164. Yarnell, W.S. and J.W. Roberts, *Mechanism of Intrinsic Transcription*

- Termination and Antitermination*. Science, 1999. **284**(5414): p. 611-615.
165. Gusarov, I. and E. Nudler, *The Mechanism of Intrinsic Transcription Termination*. Mol. Cell, 1999. **3**(4): p. 495-504.
166. Green, N.J., F.J. Grundy, and T.M. Henkin, *The T box mechanism: tRNA as a regulatory molecule*. FEBS Letters, 2010. **584**(2): p. 318-324.
167. Mandal, M. and R.R. Breaker, *Adenine riboswitches and gene activation by disruption of a transcription terminator*. Nat. Struct. Mol. Biol., 2004. **11**(1): p. 29-35.
168. Mandal, M., B. Boese, J.E. Barrick, W.C. Winkler, and R.R. Breaker, *Riboswitches Control Fundamental Biochemical Pathways in Bacillus subtilis and Other Bacteria*. Cell, 2003. **113**(5): p. 577-586.
169. Kim, J.N., A. Roth, and R.R. Breaker, *Guanine riboswitch variants from Mesoplasma florum selectively recognize 2'-deoxyguanosine*. Proc. Natl. Acad. Sci. U.S.A., 2007. **104**(41): p. 16092-16097.
170. Roth, A., W.C. Winkler, E.E. Regulski, B.W.K. Lee, J. Lim, I. Jona, J.E. Barrick, A. Ritwik, J.N. Kim, R. Welz, D. Iwata-Reuyl, and R.R. Breaker, *A riboswitch selective for the queuosine precursor preQ1 contains an unusually small aptamer domain*. Nat. Struct. Mol. Biol., 2007. **14**(4): p. 308-317.
171. Ames, T.D. and R.R. Breaker, *Bacterial aptamers that selectively bind glutamine*. RNA Biology, 2011. **8**(1): p. 82-89.
172. Mandal, M., M. Lee, J.E. Barrick, Z. Weinberg, G.M. Emilsson, W.L. Ruzzo, and R.R. Breaker, *A Glycine-Dependent Riboswitch That Uses Cooperative Binding to Control Gene Expression*. Science, 2004. **306**(5694): p. 275-279.
173. Sudarsan, N., J.K. Wickiser, S. Nakamura, M.S. Ebert, and R.R. Breaker, *An mRNA structure in bacteria that controls gene expression by binding lysine*. Genes Dev., 2003. **17**(21): p. 2688-2697.
174. Borsuk, P., A. Przykorska, K. Blachnio, M. Koper, J.M. Pawlowicz, M. Pekala, and P. Weglenski, *L-Arginine influences the structure and function of arginase mRNA in Aspergillus nidulans*. Biol. Chem., 2007. **388**(2): p. 135-144.
175. Barrick, J.E., K.A. Corbino, W.C. Winkler, A. Nahvi, M. Mandal, J. Collins, M. Lee, A. Roth, N. Sudarsan, I. Jona, J.K. Wickiser, and R.R. Breaker, *New RNA motifs suggest an expanded scope for riboswitches in bacterial genetic control*. Proc. Natl. Acad. Sci. U.S.A., 2004. **101**(17): p. 6421-6426.
176. Sudarsan, N., E.R. Lee, Z. Weinberg, R.H. Moy, J.N. Kim, K.H. Link, and R.R. Breaker, *Riboswitches in Eubacteria Sense the Second Messenger Cyclic Di-GMP*. Science, 2008. **321**(5887): p. 411-413.
177. Groisman, E.A., M.J. Cromie, Y. Shi, and T. Latifi, *A Mg²⁺-responding RNA That Controls the Expression of a Mg²⁺ Transporter*. Cold Spring Harb. Symp.

- on *Quant. Biol.*, 2006. **71**: p. 251-258.
178. Nahvi, A., J.E. Barrick, and R.R. Breaker, *Coenzyme B12 riboswitches are widespread genetic control elements in prokaryotes*. *Nucleic Acids Res.*, 2004. **32**(1): p. 143-150.
179. Winkler, W., A. Nahvi, and R.R. Breaker, *Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression*. *Nature*, 2002. **419**(6910): p. 952-956.
180. Winkler, W.C., S. Cohen-Chalamish, and R.R. Breaker, *An mRNA structure that controls gene expression by binding FMN*. *Proc. Natl. Acad. Sci. U.S.A.*, 2002. **99**: p. 15908-15913.
181. Winkler, W.C., A. Nahvi, N. Sudarsan, J.E. Barrick, and R.R. Breaker, *An mRNA structure that controls gene expression by binding S-adenosylmethionine*. *Nat. Struct. Mol. Biol.*, 2003. **10**(9): p. 701-707.
182. Wang, J.X., E.R. Lee, D.R. Morales, J. Lim, and R.R. Breaker, *Riboswitches that Sense S-adenosylhomocysteine and Activate Genes Involved in Coenzyme Recycling*. *Mol. Cell.*, 2008. **29**: p. 691-702.
183. Regulski, E.E., R.H. Moy, Z. Weinberg, J.E. Barrick, Z. Yao, W.L. Ruzzo, and R.R. Breaker, *A widespread riboswitch candidate that controls bacterial genes involved in molybdenum cofactor and tungsten cofactor metabolism*. *Mol. Microbiol.*, 2008. **68**(4): p. 918-932.
184. Serganov, A. and D.J. Patel, *Amino acid recognition and gene regulation by riboswitches*. *Biochim. Biophys. Acta.*, 2009. **1789**(9-10): p. 592-611.
185. Corbino, K., J. Barrick, J. Lim, R. Welz, B. Tucker, I. Puskarz, M. Mandal, N. Rudnick, and R. Breaker, *Evidence for a second class of S-adenosylmethionine riboswitches and other regulatory RNA motifs in alpha-proteobacteria*. *Genome Biol.*, 2005. **6**(8): p. R70.1-R70.10.
186. Fuchs, R.T., F.J. Grundy, and T.M. Henkin, *The SMK box is a new SAM-binding RNA for translational regulation of SAM synthetase*. *Nat. Struct. Mol. Biol.*, 2006. **13**(3): p. 226-233.
187. Weinberg, Z., J.E. Barrick, Z. Yao, A. Roth, J.N. Kim, J. Gore, J.X. Wang, E.R. Lee, K.F. Block, N. Sudarsan, S. Neph, M. Tompa, W.L. Ruzzo, and R.R. Breaker, *Identification of 22 candidate structured RNAs in bacteria using the CMfinder comparative genomics pipeline*. *Nucleic Acids Res.*, 2007. **35**: p. 4809-4819.
188. Poiata, E., M.M. Meyer, T.D. Ames, and R.R. Breaker, *A variant riboswitch aptamer class for S-adenosylmethionine common in marine bacteria*. *RNA*, 2009. **15**(11): p. 2046-2056.
189. Kim, J.N. and R.R. Breaker, *Purine sensing by riboswitches*. *Biol. Cell*, 2008.

- 100**(1): p. 1-11.
190. Chowdhury, S., C. Maris, F.H.T. Allain, and F. Narberhaus, *Molecular basis for temperature sensing by an RNA thermometer*. EMBO J., 2006. **25**(11): p. 2487-2497.
191. Johansson, J., P. Mandin, A. Renzoni, C. Chiaruttini, M. Springer, and P. Cossart, *An RNA Thermosensor Controls Expression of Virulence Genes in Listeria monocytogenes*. Cell, 2002. **110**(5): p. 551-561.
192. Blount, K.F. and R.R. Breaker, *Riboswitches as antibacterial drug targets*. Nat. Biotech., 2006. **24**(12): p. 1558-1564.
193. Forsyth, R.A., R.J. Haselbeck, K.L. Ohlsen, R.T. Yamamoto, H. Xu, J.D. Trawick, D. Wall, L. Wang, V. Brown-Driver, J.M. Froelich, K.G. C, P. King, M. McCarthy, C. Malone, B. Misiner, D. Robbins, Z. Tan, Z.-y. Zhu, G. Carr, D.A. Mosca, C. Zamudio, J.G. Foulkes, and J.W. Zyskind, *A genome-wide strategy for the identification of essential genes in Staphylococcus aureus*. Mol. Microbiol., 2002. **43**(6): p. 1387-1400.
194. Dunman, P.M., E. Murphy, S. Haney, D. Palacios, G. Tucker-Kellogg, S. Wu, E.L. Brown, R.J. Zagursky, D. Shlaes, and S.J. Projan, *Transcription Profiling-Based Identification of Staphylococcus aureus Genes Regulated by the agr and/or sarA Loci*. J. Bacteriol., 2001. **183**(24): p. 7341-7353.
195. Epshtein, V., A.S. Mironov, and E. Nudler, *The riboswitch-mediated control of sulfur metabolism in bacteria*. Proc. Natl. Acad. Sci. U.S.A., 2003. **100**(9): p. 5052-5056.
196. Robbins, W.J., *The Pyridine Analog of Thiamin and the Growth of Fungi*. Proc. Natl. Acad. Sci. U.S.A., 1941. **27**(9): p. 419-422.
197. Woolley, D.W. and A.G.C. White, *Selective reversible inhibition of microbial growth with pyrithiamine*. J. Exp. Med., 1943. **78**(6): p. 489-497.
198. Sudarsan, N., S. Cohen-Chalamish, S. Nakamura, G.M. Emilsson, and R.R. Breaker, *Thiamine Pyrophosphate Riboswitches Are Targets for the Antimicrobial Compound Pyrithiamine*. Chem. Biol., 2005. **12**(12): p. 1325-1335.
199. Reading, C. and M. Cole, *Clavulanic acid: a beta-lactamase-inhibiting beta-lactam from Streptomyces clavuligerus* Antimicrob. Agents Chemother., 1977. **11**(5): p. 852-857.
200. Baggaley, K.H., A.G. Brown, and C.J. Schofield, *Chemistry and biosynthesis of clavulanic acid and other clavams* Nat. Prod. Rep., 1997. **14**(4): p. 309-333
201. Notredame, C., D.G. Higgins, and J. Heringa, *T-Coffee: A novel method for fast and accurate multiple sequence alignment*. J. Mol. Biol., 2000. **302**(1): p. 205-217.

202. Hogan, D.A., S.R. Smith, E.A. Saari, J. McCracken, and R.P. Hausinger, *Site-directed Mutagenesis of 2,4-Dichlorophenoxyacetic Acid/alpha-Ketoglutarate Dioxygenase*. J. Biol. Chem., 2000. **275**(17): p. 12400-12409.
203. Schwede, T., J. Kopp, N. Guex, and M.C. Peitsch, *SWISS-MODEL: an automated protein homology-modeling server*. Nucleic Acids Res., 2003. **31**(13): p. 3381-3385.
204. Arnold, K., L. Bordoli, J. Kopp, and T. Schwede, *The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling*. Bioinformatics, 2006. **22**(2): p. 195-201.
205. Guex, N. and M.C. Peitsch, *SWISS-MODEL and the Swiss-Pdb Viewer: An environment for comparative protein modeling*. Electrophoresis, 1997. **18**: p. 2714-2723.
206. Molnar, I., J.F. Aparicio, S.F. Haydock, L. Ee Khaw, T. Schwecke, A. Konig, J. Staunton, and P.F. Leadlay, *Organisation of the biosynthetic gene cluster for rapamycin in Streptomyces hygroscopicus: Analysis of genes flanking the polyketide synthase*. Gene, 1996. **169**(1): p. 1-7.
207. Motamedi, H., A. Shafiee, S.J. Cai, S.L. Streicher, B.H. Arison, and R.R. Miller, *Characterization of methyltransferase and hydroxylase genes involved in the biosynthesis of the immunosuppressants FK506 and FK520*. J. Bacteriol., 1996. **178**(17): p. 5243-5248.
208. Kagan, R.M. and S. Clarke, *Widespread Occurrence of Three Sequence Motifs in Diverse S-Adenosylmethionine-Dependent Methyltransferases Suggests a Common Structure for These Enzymes*. Arch. Biochem. Biophys., 1994. **310**(2): p. 417-427.
209. Smith, C.P. and K.F. Chater, *Structure and regulation of controlling sequences for the Streptomyces coelicolor glycerol operon*. J. Mol. Biol., 1988. **204**(3): p. 569-580.
210. Ingram, C., M. Brawner, P. Youngman, and J. Westpheling, *xylE functions as an efficient reporter gene in Streptomyces spp.: use for the study of galP1, a catabolite-controlled promoter*. J. Bacteriol., 1989. **171**: p. 6617-6624.
211. Bibb, M.J., G.R. Janssen, and J.M. Ward, *Cloning and analysis of the promoter region of the erythromycin resistance gene (ermE) of Streptomyces erythraeus*. Gene, 1985. **38**(1-3): p. 215-226.
212. Bibb, M.J., J. White, J.M. Ward, and G.R. Janssen, *The mRNA for the 23S rRNA methylase encoded by the ermE gene of Saccharopolyspora erythraea is translated in the absence of a conventional ribosome-binding site*. Mol. Microbiol., 1994. **14**(3): p. 533-545.
213. Strieker, M., L.-O. Essen, C.T. Walsh, and M.A. Marahiel, *Non-Heme*

- Hydroxylase Engineering For Simple Enzymatic Synthesis of L-threo-Hydroxyaspartic Acid*. *Chembiochem*, 2008. **9**(3): p. 374-376.
214. Strieker, M. and M.A. Marahiel, *The Structural Diversity of Acidic Lipopeptide Antibiotics*. *Chembiochem*, 2009. **10**(4): p. 607-616.
215. Alexander, D.C., J. Rock, J.-Q. Gu, C. Mascio, M. Chu, P. Brian, and R.H. Baltz, *Production of novel lipopeptide antibiotics related to A54145 by Streptomyces fradiae mutants blocked in biosynthesis of modified amino acids and assignment of lptJ, lptK and lptL gene functions*. *J. Antibiot.*, 2011. **64**(1): p. 79-87.
216. Nogami, I., H. Shirafuji, and S. Matsumura, *Production of enduracidin and microorganisms therefor*. United States Patent, 1984. **4465771**.
217. Sun, J., G.H. Kelemen, J.M. Fernandez-Abalos, and M.J. Bibb, *Green fluorescent protein as a reporter for spatial and temporal gene expression in Streptomyces coelicolor A3(2)*. *Microbiology*, 1999. **145**(9): p. 2221-2227.
218. Chevillotte, M., R. Menges, G. Muth, W. Wohlleben, and E. Stegmann, *A quick and reliable method for monitoring gene expression in actinomycetes*. *J. Biotechnol.*, 2008. **135**(3): p. 262-265.
219. McKenzie, N.L. and J.R. Nodwell, *Transmembrane topology of the AbsA1 sensor kinase of Streptomyces coelicolor*. *Microbiology*, 2009. **155**: p. 1812-1818.
220. Rodriguez-Garcia, A., P. Combes, R. Perez-Redondo, M.C.A. Smith, and M.C.M. Smith, *Natural and synthetic tetracycline-inducible promoters for use in the antibiotic-producing bacteria Streptomyces*. *Nucl. Acids Res.*, 2005. **33**(9): p. e87-e94.
221. Ringquist, S., S. Shinedling, D. Barrick, L. Green, J. Binkley, G.D. Stormo, and L. Gold, *Translation initiation in Escherichia coli: sequences within the ribosome-binding site*. *Mol. Microbiol.*, 1992. **6**: p. 1219-1229.
222. Chiu, H.-T., B.K. Hubbard, A.N. Shah, J. Eide, R.A. Fredenburg, C.T. Walsh, and C. Khosla, *Molecular cloning and sequence analysis of the complestatin biosynthetic gene cluster*. *Proc. Natl. Acad. Sci. U.S.A.*, 2001. **98**(15): p. 8548-8553.
223. Guenzi, E., G. Galli, I. Grgurina, D.C. Gross, and G. Grandi, *Characterization of the Syringomycin Synthetase Gene Cluster. A LINK BETWEEN PROKARYOTIC AND EUKARYOTIC PEPTIDE SYNTHETASES*. *J. Biol. Chem.*, 1998. **273**(49): p. 32857-32863.
224. Gerhard, U., J.P. Mackay, R.A. Maplestone, and D.H. Williams, *The role of the sugar and chlorine substituents in the dimerization of vancomycin antibiotics*. *J. Am. Chem. Soc.*, 1993. **115**(1): p. 232-237.

225. Malabarba, A., F. Spreafico, P. Ferrari, J. Kettenring, P. Strazzolini, G. Tarzia, R. Pallanza, M. Berti, and B. Cavalleri, *Dechloro teicoplanin antibiotics*. *J. Antibiot.*, 1989. **42**: p. 1684-1697.
226. Beauregard, D.A., D.H. Williams, M.N. Gwynn, and D.J. Knowles, *Dimerization and membrane anchors in extracellular targeting of vancomycin group antibiotics*. *Antimicrob. Agents Chemother.*, 1995. **39**(3): p. 781-785.
227. Puk, O., P. Huber, D. Bischoff, J. Recktenwald, G. Jung, R.D. Submuth, K.-H. van Pee, W. Wohlleben, and S. Pelzer, *Glycopeptide Biosynthesis in Amycolatopsis mediterranei DSM5908: Function of a Halogenase and a Haloperoxidase/Perhydrolase*. *Chem. Biol.*, 2002. **9**(2): p. 225-235.
228. Yin, X., Y. Chen, L. Zhang, Y. Wang, and T.M. Zabriskie, *Enduracidin Analogues with Altered Halogenation Patterns Produced by Genetically Engineered Strains of Streptomyces fungicidicus*. *J. Nat. Prod.*, 2010. **73**(4): p. 583-589.
229. Mackay, J.P., U. Gerhard, D.A. Beauregard, R.A. Maplestone, and D.H. Williams, *Dissection of the Contributions toward Dimerization of Glycopeptide Antibiotics*. *J. Am. Chem. Soc.*, 1994. **116**(11): p. 4573-4580.
230. He, H., *Mannopectimycins, a novel class of glycopeptide antibiotics active against gram-positive bacteria*. *Appl. Microbiol. Biotechnol.*, 2005. **67**(4): p. 444-452.
231. Cudic, P., D.C. Behenna, J.K. Kranz, R.G. Kruger, A.J. Wand, Y.I. Veklich, J.W. Weisel, and D.G. McCafferty, *Functional Analysis of the Lipoglycodepsipeptide Antibiotic Ramoplanin*. *Chem. Biol.*, 2002. **9**(8): p. 897-906.
232. Cegelski, L., M. Preobrazhenskaya, and J. Schaefer, *Structures of Staphylococcus aureus Cell-Wall Complexes with Vancomycin, Eremomycin, and Chloroeremomycin Derivatives by $^{13}\text{C}\{^{19}\text{F}\}$ and $^{15}\text{N}\{^{19}\text{F}\}$ Rotational-Echo Double Resonance*. *Biochemistry*, 2006. **45**(16): p. 5235-5250.
233. Losey, H.C., J. Jiang, J.B. Biggins, M. Oberthur, X.-Y. Ye, S.D. Dong, D. Kahne, J.S. Thorson, and C.T. Walsh, *Incorporation of Glucose Analogs by GtfE and GtfD from the Vancomycin Biosynthetic Pathway to Generate Variant Glycopeptides*. *Chem. Biol.*, 2002. **9**(12): p. 1305-1314.
234. Lu, W., M. Oberthur, C. Leimkuhler, J. Tao, D. Kahne, and C.T. Walsh, *Characterization of a regiospecific epivancosaminyl transferase GtfA and enzymatic reconstitution of the antibiotic chloroeremomycin*. *Proc. Natl. Acad. Sci. U.S.A.*, 2004. **101**(13): p. 4390-4395.
235. Pelzer, S., R. Sussmuth, D. Heckmann, J. Recktenwald, P. Huber, G. Jung, and

- W. Wohlleben, *Identification and Analysis of the Balhimycin Biosynthetic Gene Cluster and Its Use for Manipulating Glycopeptide Biosynthesis in Amycolatopsis mediterranei DSM5908*. Antimicrob. Agents Chemother., 1999. **43**(7): p. 1565-1573.
236. Sosio, M., S. Stinchi, F. Beltrametti, A. Lazzarini, and S. Donadio, *The Gene Cluster for the Biosynthesis of the Glycopeptide Antibiotic A40926 by Nonomuraea Species*. Chem. Biol., 2003. **10**(6): p. 541-549.
237. Hofmann, K. and W. Stoffel, *TMbase - A database of membrane spanning proteins segments*. Biol. Chem. Hoppe-Seyler, 1993. **374**: p. 166.
238. Tusnady, G.E. and I. Simon, *The HMMTOP transmembrane topology prediction server*. Bioinformatics, 2001. **17**(9): p. 849-850.
239. Moller, S., M.D.R. Croning, and R. Apweiler, *Evaluation of methods for the prediction of membrane spanning regions*. Bioinformatics, 2001. **17**: p. 646-653.
240. Bernsel, A., H. Viklund, J. Falk, E. Lindahl, G.v. Heijne, and A. Elofsson, *Prediction of membrane-protein topology from first principles*. Proc. Natl. Acad. Sci. U.S.A., 2008. **105**: p. 7177-7181.
241. Viklund, H. and A. Elofsson, *Best alpha-helical transmembrane protein topology predictions are achieved using hidden Markov models and evolutionary information*. Protein Sci., 2004 **13**: p. 1908-1917.
242. Viklund, H. and A. Elofsson, *Improving topology prediction by two-track ANN-based preference scores and an extended topological grammar*. Bioinformatics 2008. **24** p. 1662-1668.
243. Borghi, A., P. Ferrari, G.G. Gallo, M. Zanol, L.F. Zerilli, and G.C. Lancini, *Microbial de-mannosylation and mannosylation of teicoplanin derivatives*. J. Antibiot. , 1991. **44**(12): p. 1444-1451.
244. Wehmeier, S., A.S. Varghese, S.S. Gurucha, B. Tissot, M. Panico, P. Hitchen, H.R. Morris, G.S. Besra, A. Dell, and M.C.M. Smith, *Glycosylation of the phosphate binding protein, PstS, in Streptomyces coelicolor by a pathway that resembles protein O-mannosylation in eukaryotes*. Mol. Microbiol., 2009. **71**(2): p. 421-433.
245. Wanner, J., D. Tang, C.C. McComas, B.M. Crowley, W. Jiang, J. Moss, and D.L. Boger, *A new and improved method for deglycosidation of glycopeptide antibiotics exemplified with vancomycin, ristocetin, and ramoplanin*. Bioorg. Med. Chem. Lett., 2003. **13**(6): p. 1169-1173.
246. Di Palo, S., R. Gandolfi, S. Jovetic, F. Marinelli, D. Romano, and F. Molinari, *A new bacterial mannosidase for the selective modification of ramoplanin and its derivatives*. Enzyme Microb. Technol., 2007. **41**(6-7): p. 806-811.

247. Reader, J.S., D. Metzgar, P. Schimmel, and V. de Crecy-Lagard, *Identification of Four Genes Necessary for Biosynthesis of the Modified Nucleoside Queuosine*. J. Biol. Chem., 2004. **279**(8): p. 6280-6285.
248. Gaur, R. and U. Varshney, *Genetic Analysis Identifies a Function for the queC (ybaX) Gene Product at an Initial Step in the Queuosine Biosynthetic Pathway in Escherichia coli*. J. Bacteriol., 2005. **187**(20): p. 6893-6901.
249. Van Lanen, S.G., J.S. Reader, M.A. Swairjo, V. de Crecy-Lagard, B. Lee, and D. Iwata-Reuyl, *From cyclohydrolase to oxidoreductase: Discovery of nitrile reductase activity in a common fold*. Proc. Natl. Acad. Sci. U.S.A., 2005. **102**(12): p. 4264-4269.
250. Dirk, I.-R., *Biosynthesis of the 7-deazaguanosine hypermodified nucleosides of transfer RNA*. Bioorg. Chem., 2003. **31**(1): p. 24-43.
251. Meyer, M.M., A. Roth, S.M. Chervin, G.A. Garcia, and R.R. Breaker, *Confirmation of a second natural preQ1 aptamer class in Streptococcaceae bacteria*. RNA, 2008. **14**(4): p. 685-695.
252. Klein, D.J., T.E. Edwards, and A.R. Ferre-D'Amare, *Cocrystal structure of a class I preQ1 riboswitch reveals a pseudoknot recognizing an essential hypermodified nucleobase*. Nat. Struct. Mol. Biol., 2009. **16**(3): p. 343-344.
253. Leontis, N.B. and E. Westhof, *Geometric nomenclature and classification of RNA base pairs*. RNA, 2001. **7**: p. 499-512.
254. Rieder, U., K. Lang, C. Kreutz, N. Polacek, and R. Micura, *Evidence for Pseudoknot Formation of Class I preQ1 Riboswitch Aptamers*. ChemBioChem, 2009. **10**(7): p. 1141-1144.
255. Gilbert, S.D., C.D. Stoddard, S.J. Wise, and R.T. Batey, *Thermodynamic and Kinetic Characterization of Ligand Binding to the Purine Riboswitch Aptamer Domain*. J. Mol. Biol., 2006. **359**(3): p. 754-768.
256. Sambrook, J. and E.F. Firtsch, *Molecular Cloning, A Laboratory Manual (3rd ed)*. 2001: Cold Spring Harbor Laboratory Press, New York.
257. Powell, A., M. Borg, B. Amir-Heidari, J.M. Neary, J. Thirlway, B. Wilkinson, C.P. Smith, and J. Micklefield, *Engineered Biosynthesis of Nonribosomal Lipopeptides with Modified Fatty Acid Side Chains*. J. Am. Chem. Soc., 2007. **129**(49): p. 15182-15191.
258. Hindle, Z. and C. Smith, P. , *Substrate induction and catabolite repression of the Streptomyces coelicolor glycerol operon are mediated through the GyIR protein*. Mol. Microbiol., 1994. **12**(5): p. 737-745.
259. Guerout-Fleury, A.-M., N. Frandsen, and P. Stragier, *Plasmids for ectopic integration in Bacillus subtilis*. Gene, 1996. **180**(1-2): p. 57-61.
260. Inoue, H., H. Nojima, and H. Okayama, *High efficiency transformation of*

- Escherichia coli* with plasmids. *Gene*, 1990. **96**(1): p. 23-28.
261. Spizizen, J., *Transformation of biochemically deficient strains of bacillus subtilis by deoxyribonucleate*. *Proc. Natl. Acad. Sci. U.S.A.* , 1958. **44**: p. 1072-1078.
262. Simon, R., U. Priefer, and A. Puhler, *A Broad Host Range Mobilization System for In Vivo Genetic Engineering: Transposon Mutagenesis in Gram Negative Bacteria*. *Nat. Biotech.*, 1983. **1**(9): p. 784-791.
263. MacNeil, D.J., K.M. Gewain, C.L. Ruby, G. Dezeny, P.H. Gibbons, and T. MacNeil, *Analysis of Streptomyces avermitilis genes required for avermectin biosynthesis utilizing a novel integration vector*. *Gene*, 1992. **111**(1): p. 61-68.
264. Heinrikson, R.L. and S.C. Meredith, *Amino acid analysis by reverse-phase high-performance liquid chromatography: Precolumn derivatization with phenylisothiocyanate*. *Anal. Biochem.*, 1984. **136**(1): p. 65-74.
265. Shafiee, A., H. Motamedi, and T. Chen, *Enzymology of FK-506 biosynthesis*. *Eur. J. Biochem.*, 1994. **225**(2): p. 755-764.
266. Lynch, S.A., S.K. Desai, H.K. Sajja, and J.P. Gallivan, *A High-Throughput Screen for Synthetic Riboswitches Reveals Mechanistic Insights into Their Function*. *Chem. Biol.* , 2007. **14**(2): p. 173-184.