

The planosporicin gene cluster from
Planomonospora alba

Emma Jane Sherwood

2011

Thesis submitted to the University of East Anglia for the degree of Doctor of Philosophy

© This copy of the Thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with the author and that neither quotation from the thesis, nor any information derived therefrom may be published without the author's prior, written consent.

Acknowledgements

I must first acknowledge the John Innes Centre as a wonderful place to come as a student fresh out of University with a wish to continue in science but very little knowledge of how this is best achieved.

My three rotation projects during the first year of my PhD gave me a great insight into how different labs work. I give many thanks to Dr Cyril Zipfel who as well as being an excellent supervisor, really helped me adjust to life as a research scientist as I started my first rotation. It was exciting to work in such a vibrant lab with a strong work-hard play-hard philosophy, but I came to realise my primary interest lay in molecular microbiology. The Lomonosoff and Buttner labs gave me a great introduction to life working away from plants in the Chatt building. But most thanks must go to my supervisor Prof Mervyn Bibb who offered me this opportunity at a time when microbiology PhD projects seemed hard to come by and has been unstinting in his expertise, enthusiasm and support ever since.

My lab colleagues have made the time working at John Innes Centre very enjoyable. Particular thanks must go to Maureen Bibb who first introduced me to the tricks of working with *Streptomyces* before I moved on to work with an even more recalcitrant organism. Andy Hesketh first demonstrated the potential of this project and helped ensure a smooth transition as I adopted it as my own. Sean O'Rourke, Lucy Foulston, Jan Claesen and Rob Bell were brilliant companions as we each battled with our own bizarre actinomycete and its lantipeptide gene cluster. Govind Chandra, Gerhard Saalbach, Mike Naldrett, Juan Pablo Gomez-Escribano and Kim Findlay have been very generous with their knowledge; always willing to help with any problems I happened to tackle. The provision of a dedicated media kitchen and very helpful stores meant that I could follow up new ideas with minimal fuss.

Particular thanks go to Sebastian Gehrke whose expertise in chemistry and computational biology has been invaluable to a microbiologist trying to get to grips with other disciplines. The combination of friends I have met from different labs across the site, as well as a number of fellow sports enthusiasts from across the Norwich Research Park, have ensured a healthy work-life balance. Of course I would never have got to where I am today without my parents, never pushy but always there, the Sherwood tenacity and drive to succeed ensured that giving up was never an option.

Abstract

Lantibiotics are ribosomally-synthesised peptide antibiotics that are post-translationally modified to introduce (methyl)lanthionine bridges. The increasing prevalence of antibiotic resistance in bacterial pathogens has renewed focus on the discovery of natural products with antimicrobial properties. The Actinomycetales are renowned for their ability to produce a large variety of antimicrobials, yet of these, very few lantibiotics have been described. Planosporicin is a lantibiotic produced by the actinomycete *Planomonospora alba* that inhibits cell wall biosynthesis in Gram-positive pathogens.

During this work, the structure of planosporicin was refined and reclassified as a type AI lantibiotic with a strong resemblance to those which bind lipid II. The biosynthetic gene cluster responsible for planosporicin biosynthesis was identified by genome scanning and subsequently isolated from a *P. alba* cosmid library. A minimal gene cluster for planosporicin production was defined by heterologous expression in *Nonomuraea* sp. ATCC 39727. Deletion of the gene encoding the prepropeptide abolished planosporicin production in *P. alba*, further confirming the identity of the cluster. Further deletion analyses of likely biosynthetic and processing enzymes identified through bioinformatics revealed they too are essential for planosporicin production.

The planosporicin gene cluster was shown to encode several putative regulatory proteins including an extracytoplasmic function (ECF) σ factor, a likely cognate anti- σ factor and a potential helix-turn-helix DNA binding protein. Further deletion analyses, combined with quantitative RT-PCR, revealed insights into the mechanism of regulation of planosporicin biosynthesis. A model for the pathway-specific regulation of a lantibiotic biosynthetic gene cluster by a master regulator and an ECF σ factor–anti- σ factor complex that responds to sub-inhibitory concentrations of planosporicin is proposed.

Abbreviations

A	adenine
Abu	aminobutyrate
ABC	ATP-binding cassette
Apra	apramycin
ATP	adenosine-triphosphate
AviCys	S-[(Z)-2-aminovinyl]-D-cysteine
BLAST	basic local alignment search tool
bp	base pairs
C	cytosine
Carb	carbenicillin
cDNA	complementary DNA
C terminal	carboxy-terminal
DAB	deoxyactagardine B
Dha	2,3-didehydroalanine
Dhb	(Z)-2,3-didehydrobutyrine
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	deoxynucleoside triphosphate
ESI-MS	electrospray ionisation mass spectroscopy
G	guanine
gDNA	genomic DNA
HPLC	high-pressure liquid chromatography
Hyg	hygromycin B
Kan	kanamycin
kDa	kilodalton
kb	kilo base pairs
Lab	(2S,4S,8R)-labionin
alle	L- <i>allo</i> -isoleucine
LacZ	β -galactosidase
Lan	lanthionine
LC-MS	liquid chromatography mass spectrometry
MALDI-ToF	matrix-assisted laser desorption ionisation time-of-flight

MCS	multiple-cloning site
MeLan	methyl-lanthionine
MIC	minimum inhibitory concentration
min	minute
mRNA	messenger RNA
MRSA	methicillin-resistant <i>Staphylococcus aureus</i>
MS	mass spectrometry
MS/MS	tandem mass spectrometry
Nal	nalidixic acid
N terminal	amino-terminal
NBD	nucleotide-binding domain
NMR	nuclear magnetic resonance
NRPS	non-ribosomal peptide synthetase
nt	nucleotide
OD	optical density
ORF	open reading frame
PCR	polymerase chain reaction
PFGE	pulsed-field gel electrophoresis
PKS	polyketide synthase
qRT-PCR	quantitative reverse transcriptase polymerase chain reaction
RBS	ribosome-binding site
RNA	ribonucleic acid
RNase	ribonuclease
rRNA	ribosomal RNA
RT-PCR	reverse transcriptase polymerase chain reaction
s	second
sp.	species
spec/strep	spectinomycin/streptomycin
T	thymine
Thio	thiostrepton
TM	transmembrane
TMD	transmembrane domain
Vio	viomycin
VRE	vancomycin-resistant enterococci
WT	wild type

Contents

Acknowledgements	2
Abstract	3
Abbreviations	4
Contents	6
Tables and Figures	14
Chapter 1 : Introduction	19
1.1 Actinobacteria	19
1.1.1 Phylogeny	19
1.1.2 Secondary metabolites	22
1.1.3 Regulation of secondary metabolism	24
1.2 Antibiotics	26
1.2.1 The rise in antibiotic resistance	26
1.2.2 The decline in antibiotic discovery	27
1.3 Lantibiotics	30
1.3.1 Classification	30
1.3.2 Lantibiotic producers	32
1.3.3 Lantibiotic mode of action	38
1.3.4 Lantipeptide biosynthesis	46
1.3.4.1 Biosynthetic genes	47
1.3.4.1.1 LanB	47
1.3.4.1.2 LanC	49
1.3.4.1.3 LanM	50
1.3.4.1.4 RamC/LanL	51
1.3.4.1.5 Linaridins	52
1.3.4.2 ABC transporters	52
1.3.4.3 Other modification enzymes	54
1.3.4.3.1 Protease	54
1.3.4.3.2 Decarboxylation	55
1.3.4.3.3 Hydroxylation	56
1.3.4.3.4 D-Ala	57
1.3.4.3.5 Other bridges	57
1.3.4.3.5.1 Lysinoalanine (LysAla)	57
1.3.4.3.5.2 Disulphide	57

1.3.4.3.5.3 Labionin	58
1.3.4.3.6 Single examples	58
1.3.5 Planosporicin	59
1.4 Main Objectives	61
Chapter 2 : Materials and Methods	62
2.1 Bacterial plasmids and strains	62
2.2 Oligonucleotides used in this study	67
2.3 Culture media and antibiotics	74
2.3.1 Antibiotics	74
2.3.2 Agar media	74
2.3.3 Liquid media	78
2.4 Solutions and buffers	82
2.5 General molecular biology methods	83
2.5.1 Plasmid isolation	83
2.5.2 Cosmid isolation	84
2.5.3 Agarose gel electrophoresis	84
2.5.4 Pulsed-field gel electrophoresis	84
2.5.5 DNA extraction from an agarose gel	85
2.5.6 DNA digestion with restriction enzymes	85
2.5.7 Ligation	85
2.5.8 Preparation and transformation of electrocompetent <i>E. coli</i>	86
2.5.9 Transformation of chemically-competent <i>E. coli</i>	86
2.6 General PCR methods and Sanger sequencing	86
2.6.1 General analytical PCR	87
2.6.2 High-fidelity amplification for cloning applications	87
2.6.3 Colony PCR in <i>E. coli</i>	88
2.6.4 Colony PCR in <i>Planomonospora</i>	88
2.6.5 Purification of PCR products	89
2.6.7 Sanger sequencing using Big Dye v3.1	89
2.7 Growth conditions	90
2.7.1 Growth and storage of <i>E. coli</i>	90
2.7.2 Growth and storage of <i>Micrococcus luteus</i>	90
2.7.3 Growth and storage of <i>Streptomyces</i>	90
2.7.4 Growth and storage of <i>Planomonospora</i>	90
2.7.5 Growth and storage of <i>Nonomuraea</i>	91
2.7.6 <i>P. alba</i> growth curve	91

2.8 Microscopy	92
2.8.1 Phase-contrast microscopy	92
2.8.2 Cryo-scanning electron microscopy	92
2.9 Next-generation sequencing and analysis	92
2.10 Isolation of genomic DNA	93
2.10.1 <i>P. alba</i>	93
2.10.1.1 High-molecular weight genomic DNA	93
2.10.1.2 Small-scale genomic DNA extraction	93
2.10.2 <i>Nonomuraea</i>	94
2.10.3 <i>Streptomyces</i>	94
2.11 Cosmid Library Preparation	95
2.11.1 Preparation of insert gDNA	95
2.11.2 Preparation of pSuperCos1 vector and ligation with insert DNA	95
2.11.3 Phage packaging	96
2.11.4 Phage titration	96
2.11.5 Phage transfection of <i>E. coli</i> to construct cosmid library	97
2.11.6 Library picking and transfer to membrane	97
2.11.7 Probe preparation and library hybridisation	97
2.12 Cosmid sequencing and sequence analysis	98
2.13 PCR targeting of cosmids	99
2.13.1 Construction of integrative cosmids	99
2.13.2 Construction of mutant cosmids by gene replacement	100
2.13.3 Construction of minimal gene set	101
2.13.4 FLP-mediated recombination to generate scar mutants	102
2.14 Conjugation methods	103
2.14.1 <i>Streptomyces</i>	103
2.14.2 <i>P. alba</i>	103
2.14.3 <i>Nonomuraea</i>	104
2.15 Generation of mutants by homologous recombination in <i>P. alba</i>	104
2.16 Complementation of mutant phenotypes	105
2.17 Bioassay methods	106
2.17.1 Solid bioassay	106
2.17.2 Liquid bioassay	106
2.18 Mass Spectrometry	107
2.18.1 Matrix-Assisted Laser-Desorption Ionisation Time of Flight (MALDI-ToF)	107
2.18.1.1 Low mass range: 700 - 4000 Da	107
2.18.1.2 High mass range: 4000 - 12000 Da	107

2.18.2 Electrospray ionisation (ESI)	107
2.19 Extraction Methods for planosporicin	108
2.19.1 Methanol extraction from <i>P. alba</i> mycelium	108
2.19.2 Concentration of planosporicin with Diaion HP20 bead matrix	108
2.19.3 Analytical High-Pressure Liquid Chromatography (HPLC)	109
2.20 RNA methods	109
2.20.1 Isolation of RNA from <i>P. alba</i>	109
2.20.2 DNaseI treatment of RNA	110
2.20.3 Confirming RNA quality	111
2.20.3.1 Checking for DNA contamination	111
2.20.3.2 Checking RNA concentration	111
2.20.3.3 Checking RNA quality	111
2.20.4 cDNA synthesis	111
2.20.5 Reverse transcriptase (RT)-PCR	112
2.20.6 Quantitative (q) RT-PCR	113
Chapter 3 : Structure of planosporicin	115
3.1 Introduction	115
3.2 Analysis of <i>P. alba</i> and <i>P. sphaerica</i>	115
3.2.1 Growth characteristics of <i>P. alba</i> and <i>P. sphaerica</i> on agar media	115
3.2.2 Growth characteristics of <i>P. alba</i> in liquid media	116
3.2.3 Scanning electron microscopy of <i>P. alba</i>	119
3.2.4 16S rRNA	121
3.3 Planosporicin production	121
3.3.1 Detection by bioassay	121
3.3.2 Matrix-Assisted Laser Desorption/Ionisation Time-of-Flight analysis	123
3.3.3 Purification of bioactive compounds	126
3.4 Structural analysis of planosporicin	127
3.4.1 Genome scanning	127
3.4.2 Tandem mass spectrometry (MS/MS)	130
3.5 Similarity to other lantibiotics	134
3.5.1 Nisin	134
3.5.2 Microbisporicin	137
3.5.3 Actroracin	138
3.5.4 Clausin	140
3.5.5 Kineosporiacin	141
3.6 Discussion	141

3.7 Summary	143
Chapter 4 : Identification of the planosporicin biosynthetic gene cluster	144
4.1 Introduction	144
4.2 454 sequencing	146
4.3 Cosmid library preparation	148
4.3.1 Genomic DNA	149
4.3.2 Cosmid library construction	150
4.4 Clone identification	151
4.4.1 Probe preparation	151
4.4.2 Hybridisation to cosmid library	153
4.5 Clones picked from the library	153
4.5.1 PCR on <i>P. alba</i>	153
4.5.2 End-sequencing of <i>P. alba</i> cosmids	155
4.5.3 Restriction digest on <i>P. alba</i> cosmids	155
4.6 Cosmid sequencing and annotation	156
4.6.1 Sequencing	156
4.6.2 Annotation	157
4.7 <i>P. sp.</i> DSM 14920 cosmids	159
4.7.1 PCR on <i>P. sp.</i> DSM 14920 cosmids	159
4.7.2 Restriction digest on <i>P. sp.</i> DSM 14920 cosmids	160
4.7.3 End-sequencing of <i>P. sp.</i> DSM 14920 cosmids	162
4.8 Discussion	162
4.9 Summary	163
Chapter 5 : The planosporicin biosynthetic gene cluster	164
5.1 Introduction	164
5.2 Features of the minimal gene set	165
5.3 Biosynthetic enzymes	169
5.3.1 PspB	169
5.3.2 PspC	170
5.3.3 PspTU	171
5.3.4 PspYZ	175
5.3.5 PspEF	178
5.4 Regulatory proteins	181
5.4.1 PspX	181
5.4.2 PspW	182
5.4.3 PspR	185

5.5 Unknown proteins	186
5.5.1 PspJ	186
5.5.2 PspQ	187
5.5.3 PspV	191
5.5.4 ORF-10 and ORF-11	194
5.6 Synteny	195
5.6.1 <i>mib</i> cluster	195
5.6.2 <i>Actinomyces</i> sp. cluster	196
5.6.3 <i>Bacillus</i> sp. clusters	198
5.7 Discussion	200
5.8 Summary	202
Chapter 6 : Heterologous production of planosporicin	203
6.1 Introduction	203
6.2 Expression of the planosporicin gene cluster in <i>Streptomyces</i> species	204
6.2.1 Mobilisation of cosmids	204
6.2.2 Bioassays and MALDI-ToF analysis	207
6.3 Expression of the planosporicin gene cluster in <i>Nonomuraea</i>	209
6.3.1 Mobilisation of cosmids	209
6.3.2 Bioassay and MALDI-ToF analysis	210
6.3.2.1 <i>Nonomuraea</i> M1295	210
6.3.2.2 <i>Nonomuraea</i> M1296	213
6.4 Construction of a minimal gene set for planosporicin production	215
6.4.1 Creation of constructs	215
6.4.2 Bioassay and MALDI-ToF analysis	219
6.5 Heterologous expression of <i>P. sp.</i> DSM 14920 cosmids	223
6.5.1 Creation of constructs	223
6.5.2 Bioassay and MALDI-ToF analysis	223
6.6 Discussion	227
6.7 Summary	229
Chapter 7 : Mutational analysis of planosporicin biosynthesis	230
7.1 Introduction	230
7.2 Methods and tools developed for use with <i>P. alba</i>	231
7.2.1 Conjugation methods	231
7.2.2 Integration sites	233
7.2.3 Resistance cassettes	237
7.3 Targeted deletions within the planosporicin gene cluster of <i>P. alba</i>	244

7.3.1 <i>pspA</i>	245
7.3.1.1 PCR analysis	245
7.3.1.2 Bioassay and MALDI-ToF analysis	248
7.3.2 <i>pspYZ</i>	250
7.3.2.1 PCR analysis	250
7.3.2.2 Bioassay and MALDI-ToF analysis	252
7.3.3 <i>pspTU</i>	253
7.3.3.1 PCR analysis	253
7.3.3.2 Bioassay and MALDI-ToF analysis	254
7.3.4 <i>pspV</i>	258
7.3.4.1 PCR analysis	258
7.3.4.2 Bioassay and MALDI-ToF analysis	259
7.3.5 <i>pspEF</i>	260
7.3.5.1 PCR analysis	260
7.4 Generation of ‘scar’ mutants through FLP-mediated excision of the disruption cassette	261
7.4.1 <i>pspA</i>	261
7.5 Complementations	264
7.5.1 Complementing biosynthetic mutants	264
7.5.1.1 <i>pspA</i>	264
7.5.1.2 <i>pspTU</i>	267
7.5.2 Introduction of a different <i>lanA</i> into WT <i>P. alba</i>	268
7.6 Discussion	269
7.7 Summary	273
Chapter 8 : Mutational analysis of planosporicin regulation	274
8.1 Introduction	274
8.2 Targeted deletions within the planosporicin gene cluster of <i>P. alba</i>	275
8.2.1 <i>pspX</i>	275
8.2.1.1 PCR analysis	275
8.2.1.2 Bioassay and MALDI-ToF analysis	277
8.2.2 <i>pspW</i>	281
8.2.2.1 PCR analysis	281
8.2.2.2 Bioassay and MALDI-ToF analysis	282
8.2.3 <i>pspR</i>	284
8.2.3.1 PCR analysis	284
8.2.3.2 Bioassay and MALDI-ToF analysis	284

8.3 Complementations	285
8.3.1 Complementing regulatory mutants	285
8.3.1.1 <i>pspX</i>	285
8.3.1.2 <i>pspR</i>	286
8.4 Computational analysis	288
8.4.1 Dyad symmetries	288
8.4.1.1 Attenuators	288
8.4.1.2 Binding sites	288
8.4.2 Promoter motifs	289
8.5 Transcriptional analysis	291
8.5.1 Reverse transcriptase (RT) PCR	294
8.5.2 Quantitative (q) PCR	296
8.6 Discussion	301
8.7 Summary	303
Chapter 9 : General discussion	304
9.1 Planosporicin structure and activity	304
9.2 Planosporicin mode of action	305
9.3 The <i>psp</i> gene cluster	307
9.3.1 Introduction of non-proteogenic amino acids	307
9.3.2 ATP-binding cassette transporters	309
9.3.3 Regulatory genes	311
9.3.4 Unknown genes	313
9.4 The model of <i>psp</i> regulation	314
9.5 Future work	316
Chapter 10 : References	319

Tables and Figures

Figure 1.1 : Phylogenetic tree of the class Actinobacteria.	20
Figure 1.2 : A phylogenetic tree based on partial 16S rRNA sequences from representative members of the <i>Streptosporangiaceae</i> family.	21
Figure 1.3 : Structural motifs introduced into lantibiotics by post-translational modification.	31
Figure 1.4 : A schematic of seven representative lantibiotic gene clusters from low-GC Gram-positive bacteria.	33
Figure 1.5 : A schematic of actinomycete lantibiotic biosynthetic gene clusters.	35
Figure 1.6 : Representatives of lantibiotics produced by actinomycetes.	36
Figure 1.7 : The structure of lipid II.	40
Figure 1.8 : Lantibiotics with N-terminal structural similarity to nisin.	41
Figure 1.9 : NMR resolved solution structure of nisin in complex with the lipid II variant 3LII (shortened prenyl tail of 3 isoprene units instead of 11 in full-length lipid II).	42
Figure 1.10 : Structures of lantibiotics containing a conserved methyl-lanthionine ring proposed to bind lipid II (in red).	44
Figure 1.11 : Dehydration and cyclisation of lantibiotic prepropeptides.	48
Table 1.1 : Minimum inhibitory concentrations (in $\mu\text{g/ml}$) for planosporicin, actagardine, mersacidin, nisin and microbisporicin against a range of target pathogens.	60
Table 2.1 : Plasmids and cosmids used and constructed in this study	62
Table 2.2 : <i>E. coli</i> strains used in this study	65
Table 2.3 : <i>Actinomycetes</i> strains used in this study	66
Table 2.4 : <i>Actinomycetes</i> strains constructed in this study	66
Table 2.5 : Primers used in this study	67
Table 2.6 : Concentration of antibiotics used in this study	74
Table 2.7 : Agar media used in this study	74
Table 2.8 : Liquid media used in this study	78
Table 2.9 : Solutions and buffers used in this study	82
Figure 3.1 : Morphology of <i>P. alba</i> and <i>P. sphaerica</i> grown on agar media.	117
Figure 3.2 : Morphology of <i>P. alba</i> in liquid ISP4 media.	118
Figure 3.3 : Growth curve of <i>P. alba</i> measured by OD_{450} and total protein assay.	119
Figure 3.4 : Morphology of <i>P. alba</i> spores.	120
Figure 3.5 : Phylogram based on alignment of partial 16S rRNA sequences from five <i>Planomonospora</i> strains.	121

Figure 3.6 : <i>P. alba</i> solid bioassay. _____	122
Figure 3.7 : Growth curve of <i>P. alba</i> in liquid medium. _____	123
Figure 3.8 : MALDI-ToF mass spectrometry analysis. _____	125
Table 3.1 : The predicted masses of planosporicin adducts. _____	126
Figure 3.9 : Concentrated culture supernatant of <i>P. alba</i> . _____	128
Figure 3.10 : Structure of planosporicin. _____	130
Figure 3.11 : Product ion scan. _____	132
Table 3.2 : The major fragments from MS/MS of the 2191.8054 Da peptide planosporicin. _____	133
Figure 3.12 : Revised structure of planosporicin based on peptide fragments from MS/MS. _____	133
Figure 3.13 : An alignment of the planosporicin prepropeptide (PspA) with eight others. _____	135
Figure 3.14 : A comparison of lantibiotics with structural similarity to planosporicin. ____	136
Figure 3.15 : A comparison of lantibiotics with structural similarity to planosporicin. ____	139
Figure 4.1 : A schematic depicting three genes proposed to be part of the planosporicin biosynthetic gene cluster annotated using the generic designation ' <i>lan</i> '. _____	147
Figure 4.2 : Histogram of the GC content for the assembled contigs from 454 sequencing of <i>P. alba</i> genomic DNA. _____	149
Figure 4.3 : High molecular weight <i>P. alba</i> gDNA analysed by PFGE. _____	150
Figure 4.4 : Schematic of cosmid library construction. _____	152
Figure 4.5 : Hybridisation of the <i>P. alba</i> cosmid library with three different probes. ____	154
Figure 4.6 : Restriction digest patterns of different <i>P. alba</i> cosmids. _____	156
Figure 4.7 : Cosmid B4-1 viewed in Artemis (Sanger). _____	158
Figure 4.8 : PCR analysis to determine which <i>P. DSM 14920</i> cosmids were likely to contain the entire planosporicin gene cluster. _____	159
Figure 4.9 : PCR analysis to check the similarity of <i>P. DSM 14920</i> cosmid 4B8 to <i>P. alba</i> cosmid B4-1, both encoding the planosporicin gene cluster. _____	160
Figure 4.10 : Restriction digest patterns of eight <i>P. alba</i> cosmids and four <i>P. sp. DSM</i> 14920 cosmids with two different enzymes. _____	161
Figure 5.1 : The putative planosporicin gene cluster with each gene shown to scale. ____	165
Table 5.1 : Putative proteins of the planosporicin biosynthetic gene cluster. _____	166
Figure 5.2 : Sequence alignment of PspC with five highly similar LanC proteins. ____	172
Figure 5.3 : Features of the ABC transporter PspTU. _____	173
Figure 5.4 : Features of the ABC transporter PspYZ. _____	176
Figure 5.5 : A schematic showing the syntenous arrangement of the proteins bearing significant similarity to those encoded by <i>pspX</i> , <i>pspW</i> , <i>pspJ</i> , <i>pspY</i> and <i>pspZ</i> . ____	177

Figure 5.6 : Features of the ABC transporter PspEF.	178
Figure 5.7 : Sequence alignment of the three ATP-binding proteins in the planosporicin cluster (PspF, PspZ, PspT) with ten highly similar ATP-binding proteins.	180
Figure 5.8 : A schematic showing the transmembrane helices predicted to occur in PspW and eight PspW-like proteins.	184
Figure 5.9 : Features of PspJ-like proteins.	188
Figure 5.10 : An alignment of PspQ with five PspQ-like proteins.	190
Figure 5.11 : The tripartite nature of the signal sequence of lipoproteins.	191
Figure 5.12 : A schematic showing the context of genes encoding proteins with significant similarity to PspV in the genomes listed.	193
Figure 5.13 : A comparison of the <i>psp</i> and <i>mib</i> clusters using the Artemis Comparison Tool (ACT; Sanger).	197
Figure 5.14 : A schematic showing the arrangement of clusters encoding proteins identified to have significant similarity to PspABC in the genomes listed.	199
Figure 6.1 : Construct pIJ12323; pIJ12321 (cosmid B4-1) with the backbone targeted with the <i>oriT-attP-int-aac(3)IV</i> cassette from pIJ10702.	206
Table 6.1 : Strains made in this Chapter through integration of various plasmids into the Φ C31 phage attachment site.	207
Figure 6.2 : PCR confirmation of the integration of <i>Planomonospora</i> cosmids pIJ12323, pIJ12324, pIJ12325 and pIJ12326 in <i>S. coelicolor</i> M1146 and <i>S. lividans</i> TK24.	208
Figure 6.3 : PCR confirmation of the integration of <i>Planomonospora</i> cosmids pIJ12323 and pIJ12324 in <i>Nonomuraea</i> ATCC 39727.	210
Figure 6.4 : Heterologous expression of the planosporicin gene cluster integrated into the <i>Nonomuraea</i> ATCC 39727 chromosome.	212
Figure 6.5 : Heterologous expression of the planosporicin gene cluster integrated into the <i>Nonomuraea</i> ATCC 39727 chromosome.	213
Figure 6.6 : Heterologous production of planosporicin in <i>Nonomuraea</i> ATCC 39727.	214
Figure 6.7 : Strategy for the construction of a minimal <i>psp</i> gene set.	217
Figure 6.8 : Schematic illustrating the reduced <i>psp</i> gene clusters.	218
Figure 6.9 : PCR confirmation of the integration of <i>Planomonospora</i> cosmids pIJ12327, pIJ12328 and pIJ12329 in <i>Nonomuraea</i> ATCC 39727.	219
Figure 6.10 : Heterologous production of planosporicin from reduced gene sets integrated into the <i>Nonomuraea</i> ATCC 39727 chromosome.	220
Figure 6.11 : Heterologous expression of the planosporicin gene cluster integrated into the <i>Nonomuraea</i> ATCC 39727 chromosome.	221

Figure 6.12 : Heterologous production of planosporicin from reduced gene clusters derived from pIJ12321 integrated into the <i>Nonomuraea</i> ATCC 39727 chromosome.	222
Figure 6.13 : PCR confirmation of the integration of <i>Planomonospora</i> cosmids pIJ12325 and pIJ12326 in <i>Nonomuraea</i> ATCC 39727.	224
Figure 6.14 : Heterologous expression of the planosporicin gene cluster integrated into the <i>Nonomuraea</i> ATCC 39727 chromosome.	225
Figure 6.15 : Heterologous production of planosporicin from <i>P. sp.</i> DSM 14920 cosmids integrated into the <i>Nonomuraea</i> ATCC 39727 chromosome.	226
Figure 7.1 : Conjugation with <i>E. coli</i> ET12567 pUZ8002 harbouring pMS82.	234
Table 7.1 : A list of the site-specific integrative vectors conjugated into <i>P. alba</i> .	235
Figure 7.2 : Colony PCR analysis to confirm the integration of pMS82 into the genome of <i>P. alba</i> .	235
Figure 7.3 : PCR analysis to confirm the integration of pIJ10706 into the <i>P. alba</i> genome.	236
Figure 7.4 : An alignment of a section of the <i>S. coelicolor</i> Φ C31 insertion site with a region of contig00142 from the <i>P. alba</i> 454 contig database.	237
Figure 7.5 : Schematic of four new constructs based on pIJ10706.	239
Figure 7.6 : Growth of <i>P. alba</i> strains with antibiotic selection.	241
Figure 7.7 : Growth of <i>P. alba</i> strains with antibiotic selection.	242
Figure 7.8 : Schematic of reduced constructs used to make a targeted deletion of <i>pspA</i> .	247
Figure 7.9 : PCR confirmation of the deletion of <i>pspA</i> from the <i>P. alba</i> chromosome.	248
Figure 7.10 : Abolition of planosporicin production due to the targeted deletion of <i>pspA</i> in <i>P. alba</i> .	249
Figure 7.11 : Double crossover mutants isolated after growth in liquid culture.	251
Figure 7.12 : PCR confirmation of four <i>P. alba</i> deletion mutants.	252
Figure 7.13 : Lack of production of antibiotic compounds from <i>P. alba</i> deletion mutants.	254
Figure 7.14 : Lack of production of peptides by four <i>P. alba</i> deletion mutants.	255
Figure 7.15 : Production of antibiotic compounds by <i>P. alba</i> M1305 in comparison to WT <i>P. alba</i> when grown on agar medium.	256
Figure 7.16 : Attempted detection of antibiotic compounds either secreted into the medium or accumulated in the cytoplasm.	257
Figure 7.17 : Schematic illustrating the construction of a reduced construct to use to generate a <i>pspV</i> deletion mutant in <i>P. alba</i> .	259

Figure 7.18 : Lack of production of antibiotic compounds from <i>P. alba</i> deletion mutants complemented in <i>trans</i> with the WT gene. _____	265
Figure 7.19 : Lack of production of peptides from <i>P. alba</i> deletion mutants complemented in <i>trans</i> with the WT gene. _____	266
Figure 8.1 : Generation of <i>P. alba</i> Δ <i>pspX</i> :: <i>hyg</i> exconjugants. _____	276
Figure 8.2 : Lack of production of antibiotic compounds from three <i>P. alba</i> Δ <i>pspX</i> :: <i>hyg</i> exconjugants. _____	278
Figure 8.3 : Production of antibiotic compounds from four <i>P. alba</i> deletion mutants. ____	279
Figure 8.4 : Production of peptides from three <i>P. alba</i> deletion mutants. _____	280
Figure 8.5 : PCR confirmation of three <i>P. alba</i> deletion mutants. _____	282
Figure 8.6 : Production of antibiotic compounds by <i>P. alba</i> M1304 and the WT strain. _	283
Figure 8.7 : Production of antibiotic compounds from <i>P. alba</i> deletion mutants complemented in <i>trans</i> with the WT gene. _____	286
Figure 8.8 : Production of peptides from <i>P. alba</i> deletion mutants complemented in <i>trans</i> with the WT gene. _____	287
Figure 8.9 : The 36 bp hairpin loop located between <i>pspA</i> and <i>pspB</i> . _____	288
Figure 8.10 : A sequence logo of the consensus sequence compiled by MEME. ____	290
Table 8.1 : Promoter motifs identified by MEME. _____	291
Figure 8.11 : Growth and planosporicin production in five <i>P. alba</i> strains. _____	292
Figure 8.12 : Appearance of RNA extracted from <i>P. alba</i> separated by gel electrophoresis. _____	293
Figure 8.13 : RT-PCR on four <i>P. alba</i> strains. _____	295
Figure 8.14 : Level of transcription from 11 genes of the <i>psp</i> gene cluster in four different <i>P. alba</i> strains at two different time points. _____	298
Figure 8.15 : Level of transcription from 7 genes of the <i>psp</i> gene cluster in two different <i>P. alba</i> strains at two different time points. _____	300
Figure 8.16 : Model for the regulation of the <i>psp</i> gene cluster. _____	301
Figure 9.1 : Schematic representation of planosporicin post-translational modifications.	308
Figure 9.2 : A schematic depicting the model of regulation of planosporicin biosynthesis. _____	315

Chapter 1 : Introduction

1.1 Actinobacteria

The phylum Actinobacteria consists of 39 families and 130 genera, making it one of the largest lineages within the domain Bacteria (Figure 1.1) (Ventura *et al.* 2007; Miao and Davies 2010). Actinobacteria are Gram-positive bacteria with a high genomic GC content; ranging from 51 mol% GC in some *Corynebacterium* to more than 70 mol% GC in *Streptomyces* and *Frankia* (Ventura *et al.* 2007). The sequencing of many actinobacterial genomes revealed vast genomic diversity within this phylum. Sequenced genomes to date range from the 2.50 Mb *Micrococcus luteus* circular chromosome predicted to encode 2,403 proteins (Young *et al.* 2010), to the 9.67 Mb *Rhodococcus jostii* linear chromosome predicted to encode 9,145 proteins (McLeod *et al.* 2006).

Actinobacteria exhibit a wide variety of morphologies, but many, such as the *Streptomycetaceae* family, have a highly differentiated, branched mycelium. This superficially resembles the mycelium of fungi, which they were originally classified as under the older name actinomycetes (mycet is derived from the Greek *mukēt* meaning fungus), referred to in subsequent Chapters. Unusual developmental features are displayed by many actinobacterial genera, such as formation of sporulating aerial mycelium in many members of the Actinomycetales order. These features are linked to the variety of lifestyles exhibited by different genera. Many *Streptomyces* species are soil inhabitants where they play a crucial role in the decomposition of organic matter to recycle nutrients back into the ecosystem. However, other Actinobacteria are plant commensals (*Leifsonia*), nitrogen-fixing symbionts (*Frankia*), gastrointestinal tract inhabitants (*Bifidobacterium*) and even pathogens, such as some *Mycobacterium* species (Ventura *et al.* 2007).

1.1.1 Phylogeny

Among Actinobacteria, the *Streptomycetaceae* are the dominant family, with non-streptomycetes referred to as 'rare actinomycetes' due to their low isolation frequency. The family *Streptosporangiaceae* was first proposed over 20 years ago (Goodfellow *et al.* 1990). This family initially consisted of the genera *Streptosporangium*, *Microbispora*, *Microtetraspora*, *Planobispora* and *Planomonospora*. Subsequently, further genera have

been added, such as *Nonomuraea* (Zhang *et al.* 1998), to form a family of nine related genera with *Streptosporangium* as the type genus. Together, the family *Streptosporangiaceae* forms a distinct clade on the basis of 16S rRNA gene sequence data (Figure 1.2), but display a range of morphologies. Genera such as *Planomonospora* form spore vesicles containing single spores, whereas other genera form spore chains. *Microbispora* has spores in characteristic longitudinal pairs while *Nonomuraea* displays spore chains on aerial hyphae (Dworkin 1996).

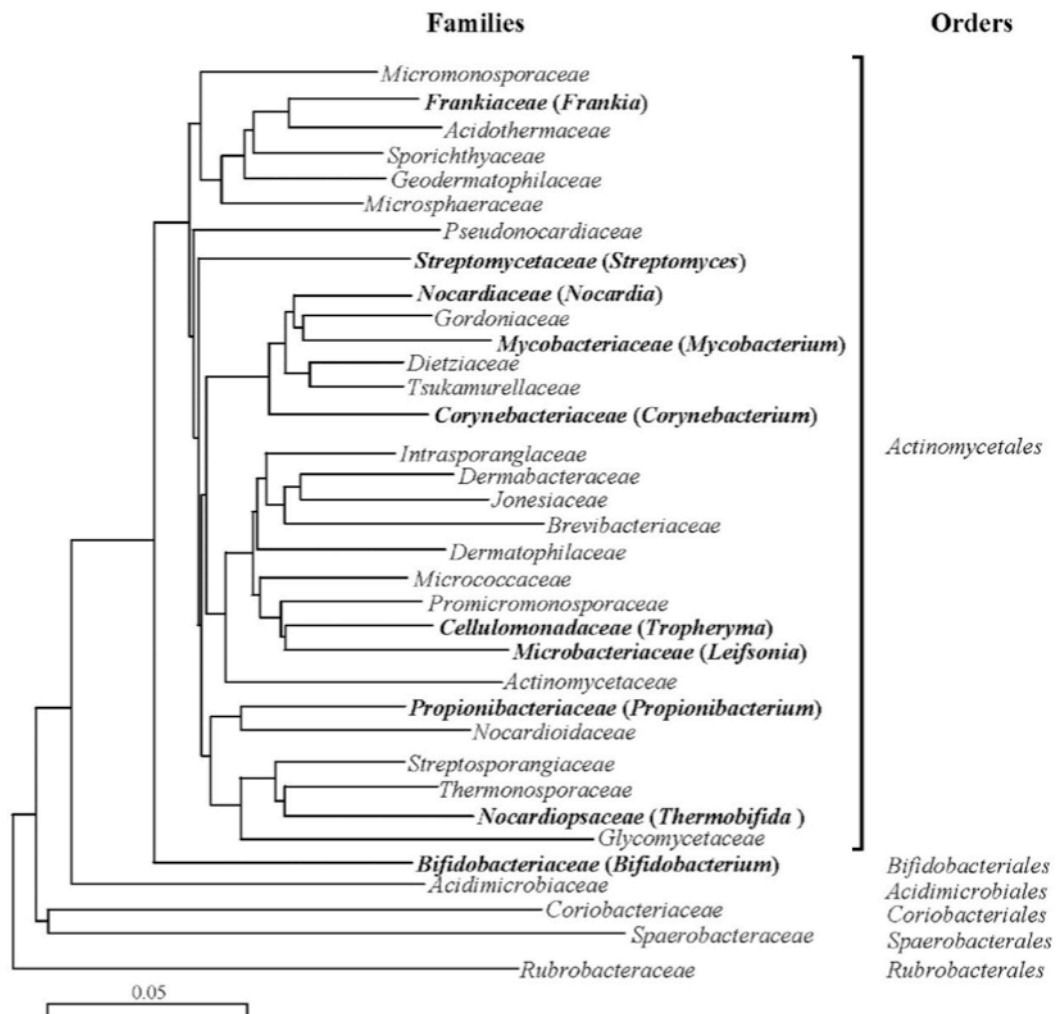


Figure 1.1 : Phylogenetic tree of the class Actinobacteria.

Taxonomy is based on 1,500 nucleotides of 16S rRNA. Families containing members subjected to complete genome sequencing (as of August 2010) are depicted in bold. Scale bar, 5 nucleotides. Adapted from Miao and Davies 2010.

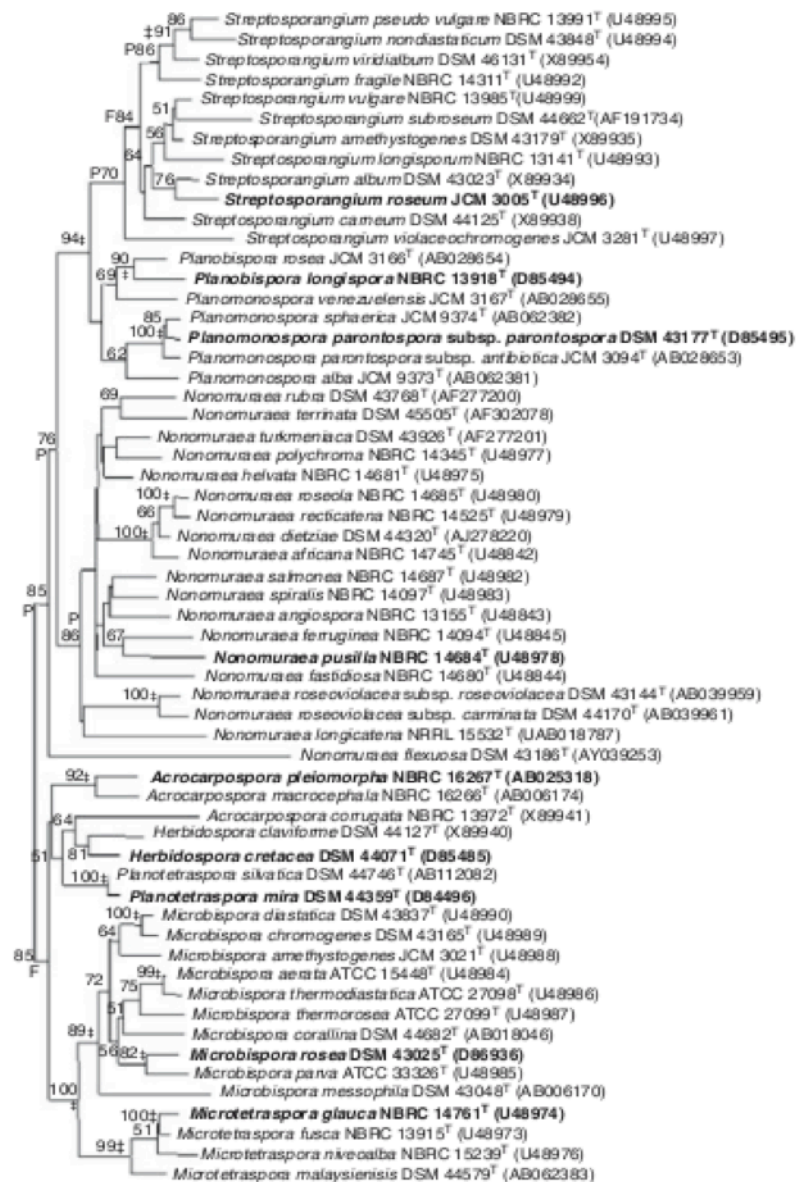


Figure 1.2 : A phylogenetic tree based on partial 16S rRNA sequences from representative members of the *Streptosporangiaceae* family.

Type species are in bold and ^T indicates type strain. The numbers at the nodes indicate the percentage of bootstrap support and an asterisk indicates the branch was also found using other tree-making algorithms. Figure from Dworkin 1996.

The *Planomonospora* genus was initially proposed in 1967 with *Planomonospora parontospora* as the type species (Thiemann *et al.* 1967). The genus now contains five validly described species; *Planomonospora alba* (Mertz 1994), *Planomonospora parontospora* subsp. *antibiotica* (Thiemann *et al.* 1967), *Planomonospora parontospora* subsp. *parontospora* (Thiemann *et al.* 1967), *Planomonospora sphaerica* (Mertz 1994) and *Planomonospora venezuelensis* (Thiemann 1970). *P. venezuelensis* was first identified from soil in Venezuela, whereas *P. alba* was identified from the Sudan,

indicating a wide geographical distribution. A subsequent study isolated 246 *Planomonospora* strains from 137 of the 1200 soil samples examined (Suzuki *et al.* 2001). Strains of the *P. parontospora* group were recovered from 131 of these soil samples, while strains assigned to the *P. venezuelensis* group were isolated from 13 soil samples. It is now clear that *Planomonospora* strains have a worldwide distribution in soils of arid, temperate and tropical regions.

P. alba is an aerobic, Gram-positive, chemo-organotrophic actinomycete that forms branched, non-fragmenting substrate hyphae and sparsely branched aerial hyphae (Mertz 1994). Clavate spore vesicles each contain a single spore, which become motile in water due to peritrichous flagella (Mertz 1994). The cell membrane contains a number of glycerol-based phospholipids such as phosphatidylethanolamine, phosphatidylinositol and phosphatidylglycerol as well as glucosamine-containing phospholipids (Mertz 1994).

1.1.2 Secondary metabolites

Actinobacteria exhibit diverse physiological and metabolic properties in order to adapt to changing environments. Many produce so-called 'secondary metabolites', defined as metabolites that are not essential for growth under laboratory conditions. These include pigments, such as the brightly coloured actinorhodin and prodiginines from *Streptomyces coelicolor* and many antibiotics (Wright and Hopwood 1976a; Rudd and Hopwood 1980). The first actinomycete antibiotic to be discovered was actinomycin, produced by *Actinomyces antibioticus* (Waksman and Woodruff 1941). Subsequently hundreds of secondary metabolites with potent antimicrobial activity have been discovered, primarily from the *Streptomyces* genus. For example, *S. coelicolor* produces both methylenomycin and the calcium-dependent antibiotic (Wright and Hopwood 1976b; Lakey *et al.* 1983). However, other actinomycetes genera also have the capacity to produce antimicrobials of pharmacological and commercial interest.

The term 'secondary metabolites' implies inferiority to primary metabolites, but non-essentiality for growth under controlled lab conditions has no bearing on their important roles in the environment. Moreover, there is increasing evidence that many secondary metabolites may play roles as signalling molecules at low concentrations (Goh *et al.* 2002; Yim *et al.* 2007). Surfactin originally drew attention due to its antimicrobial activity but has recently been shown to stimulate biofilm development in *Bacillus subtilis* through a mechanism that triggers K⁺ leakage resulting in the elevated production of extracellular matrix components (Lopez *et al.* 2009). This may be a common theme in which secondary metabolites act as intra- and inter-specific signals, cues or chemical manipulators at one

concentration, and have a second role as antibiotics, toxins, or ionophores at higher concentrations (Keller and Surette 2006; Shank and Kolter 2009).

The secondary metabolites of the Antibiotic Literature Database (ABL) are mostly produced by the bacterial order Actinomycetales (57 %) with 13 % made up by other bacteria and the remaining 30 % produced by fungi¹. More than 20,000 microbial secondary metabolites possessing some bioactivity *in vitro* have been discovered (Marinelli 2009). Many were initially studied as potential anti-infectives (antibacterials, antifungals, antivirals, and antiparasitics) but eventually developed into leading anticancer drugs, immunosuppressive agents for organ transplantation, or pharmaceuticals targeting metabolic and cardiovascular diseases (Marinelli 2009).

The bioactive molecules produced by actinomycetes belong to a range of chemical classes. Of most relevance to this thesis are the peptide antibiotics. Thiopeptides contain a characteristic macrocyclic core consisting of multiple thiazoles, dehydroamino acids, and a 6-membered nitrogen heterocycle (Velasquez and van der Donk 2011). Recently the thiostrepton biosynthetic gene cluster was identified from whole-genome scanning of *Streptomyces laurentii* demonstrating thiopeptides are ribosomally-synthesised then post-translationally modified (Kelly *et al.* 2009). There are also many non-ribosomally produced peptides which exhibit bioactivity. Non-ribosomal peptides are synthesised by non-ribosomal peptide synthetases (NRPSs) through the sequential condensation of amino acid monomers. Unlike ribosomes, NRPSs can incorporate a range of non-proteogenic amino acids, leading to unusual chemical properties and structures that can be further modified by specific tailoring enzymes to add additional functional groups (Walsh *et al.* 2001; Lautru and Challis 2004). Several of these peptide-based natural products are currently used in the clinic to treat serious Gram-positive infections. For example, the glycopeptide vancomycin is synthesised by an NRPS as a linear heptapeptide which is further modified by tailoring enzymes to glycosylate and introduce oxidative cross-links (van Wageningen *et al.* 1998). The cyclic lipopeptide daptomycin consists of thirteen amino acids, ten of which form a ring, while three form an exocyclic tail with decanoic acid attached to the N-terminus (Miao *et al.* 2005). Daptomycin was the second member of this class to be discovered but its bactericidal activity against vancomycin resistant Gram-positive pathogens ensured that it was the first to be marketed for commercial use (Steenbergen *et al.* 2005). However, this thesis focuses on one particular group of ribosomally synthesised and post-translationally modified peptides, the lantibiotics (introduced in Section 1.3).

¹ www.ricercaperlavita.it/en/patrimonio_fiirv/abl_database.htm (updated until 2006)

1.1.3 Regulation of secondary metabolism

The regulation of secondary metabolism occurs at two levels, it can be pathway-specific and/ or mediated through pleiotropic regulatory genes (Bibb 2005). The use of multiple regulators gives the scope for different signals to influence the production of secondary metabolites. Therefore it is not uncommon to observe conditional phenotypes where the wild type (WT) phenotype can be restored if the growth conditions are altered (Rigali *et al.* 2008).

The highly phosphorylated nucleotide ppGpp (guanosine 5'-diphosphate 3'-diphosphate) mediates the stringent response under conditions of amino acid starvation in bacteria. This response modulates transcription to divert resources away from growth and division in order to adapt until nutrient conditions improve. In actinomycetes these adaptive responses include the induction of antibiotic production, and there is increasing evidence that ppGpp plays a direct role in transcriptional activation (rather than this activation being an indirect consequence of a growth rate reduction). A modified ppGpp synthetase gene, *relA*, was used to induce ppGpp synthesis without a detectable reduction in growth and was shown to elicit the transcription of *actII-ORF4* and the consequent production of actinorhodin in *S. coelicolor* (Hesketh *et al.* 2001). The global role of ppGpp synthesis in *S. coelicolor* was demonstrated through the use of Affymetrix microarrays and revealed an influence on the expression of antibiotic gene clusters, conservons, and morphogenetic proteins (Hesketh *et al.* 2007).

In actinomycetes, the developmental regulation of secondary metabolite production is well-documented and commonly linked to the presence of regulatory metabolites, such as A-factor (the auto-regulatory-factor 2-isocapryloyl-3R-hydroxymethyl- γ -butyrolactone) which induces morphological and chemical differentiation in *Streptomyces griseus*, resulting in sporulation and the production of streptomycin and grinoxazole (Horinouchi and Beppu 1994). A-factor binds the receptor protein ArpA and prompts its dissociation from the *adpA* promoter resulting in transcription of *adpA* (Kato *et al.* 2004). AdpA consists of an N-terminal dimerisation domain and a C-terminal DNA-binding domain, and is the key transcriptional activator of a regulon that includes genes required for secondary metabolism and morphological differentiation (Yamazaki *et al.* 2004). In contrast to other γ -butyrolactone regulatory systems, disruption of *arpA* had no effect on A-factor production, which is instead repressed by AdpA in a two-step regulatory feedback loop (Kato *et al.* 2004). However, A-factor may well be the exception among γ -butyrolactones as a pleiotropic regulator. Other γ -butyrolactones only regulate secondary metabolism and the frequent location of genes for their biosynthesis and receptor proteins within secondary metabolite gene clusters implies that their regulatory role may be pathway-specific. For example, the jadomycin gene cluster of *Streptomyces venezuelae* encodes

JadR2, a γ -butyrolactone receptor homologue which functions with JadR1 in stress-induced activation of jadomycin B biosynthesis (Wang and Vining 2003).

Other regulators include the SARP (*Streptomyces* Antibiotic Regulatory Protein) family and the LAL (Large ATP-binding regulators of the LuxR family) family, responsible for the transcriptional activation of specific secondary metabolic pathways in many actinomycetes. To take a recent example, Aur1PR3 is a transcriptional activator of the SARP family which forms a regulatory cascade along with the TetR family negative regulator Aur1R to control expression of the *aur1* polyketide gene cluster involved in biosynthesis of the angucycline-like antibiotic auricin in *Streptomyces aureofaciens* CCM 3239 (Novakova *et al.* 2011).

However, the regulation of lantibiotic gene clusters is best characterised in low-GC organisms. In several systems, lantibiotics have been shown to induce their own synthesis. The first insight into the mechanism of nisin autoregulation in *Lactococcus lactis* came from the demonstration that the sensor histidine kinase NisK and the response regulator NisR transduced signals resulting in the transcriptional autoregulation of *nisA* which encodes the nisin precursor peptide (Kuipers *et al.* 1995). Subsequently *nisF*, which encodes part of the immunity ABC transporter, was also demonstrated to be nisin inducible (de Ruyter *et al.* 1996). The SpaRK two-component system regulating subtilin production in *B. subtilis* ATCC 6633 is itself under control by two independent regulatory systems. These are autoinduction via subtilin and transcriptional regulation via σ^H (Stein *et al.* 2002b). In contrast, the deletion of *spo0H* (encoding σ^H) had no significant effect on mersacidin production and immunity in *Bacillus* sp. strain HIL Y-85,54728 (Schmitz *et al.* 2006). Instead mersacidin functions as an autoinducer, promoting transcription of *mrsA*, the gene encoding the precursor peptide, resulting in earlier production of mersacidin (Schmitz *et al.* 2006). Two-component regulatory systems are not exclusive to low-GC lantibiotic clusters. The cinnamycin gene cluster encodes the two-component regulatory system CinRK. However this cluster also contains one further regulator, a SARP (CinR1) (Widdick *et al.* 2003). Thus the biosynthesis of lantibiotics is often regulated in a cell density dependent manner in which the lantibiotic acts as a signalling molecule in an autoinduction mechanism.

1.2 Antibiotics

1.2.1 The rise in antibiotic resistance

It is more than sixty years since clinical implementation of the first natural product antibiotic. Yet infectious disease remains a global problem. It is now understood that use (and misuse) of an antibiotic selects for bacteria that are resistant to the drug. Thus clinically significant resistance commonly arises within a couple of decades after the introduction of a new antibiotic. In the case of penicillin, pathogen resistance through the acquisition of a single hydrolytic enzyme occurred less than one year after clinical introduction of the drug in 1945 (Walsh 2003). In contrast, vancomycin resistance did not emerge until 30 years after its first clinical use, due to its initially infrequent use and the fact that resistance required the assembly of a five-gene resistance cassette (Walsh 2003).

The prevalence of hospital-acquired (nosocomial) infections caused by Gram-positive pathogens such as *Staphylococcus aureus*, *Streptococcus pneumoniae*, enterococci and *Clostridium difficile* has increased due to the acquisition of multidrug resistance. *S. aureus* most commonly causes infections of the skin and bloodstream (bacteraemia), as well as pneumonia, which can be severe or fatal. The number of cases of methicillin-resistant *S. aureus* (MRSA) bacteraemia in the UK increased from 2 % in 1990 to 43 % in 2002 (Johnson *et al.* 2005). In recent years an improvement in cleaning regimes has decreased the number of deaths attributed to MRSA², however there is no indication of a decrease in virulence, and a more lethal strain may subsequently develop. Recently, new strains of MRSA have emerged that cause infections in the community. In 2009 it was estimated that approximately 30 % of the UK population were colonised with *S. aureus*, and 1-3 % of the total population colonised with MRSA³. Community-associated (CA)-MRSA strains are genetically and phenotypically distinct from Hospital-acquired (HA)-MRSA. For example, many CA-MRSA strains are producers of the lantibiotic Bsa (bacteriocin of *S. aureus*), presumably conferring a competitive ecological advantage (Daly *et al.* 2010). HA-MRSA is resistant to fluoroquinolones, macrolides, lincosamides, streptogramins, aminoglycosides, rifampin and tetracyclines whereas CA-MRSA is susceptible to many non- β -lactam antibiotics (Van Bambeke *et al.* 2008). Treatment of MRSA with vancomycin led to the acquisition of the VanA determinant from vancomycin-resistant enterococci (VRE) to

² MRSA deaths continue to decrease (August 2011) Office for National Statistics
<http://www.statistics.gov.uk/cci/nugget.asp?id=1067>

³ MRSA in primary care (January 2009) Clinical Knowledge Summaries
www.cks.nhs.uk/mrsa_in_primary_care

generate vancomycin-resistant *S. aureus* (VRSA) (Van Bambeke *et al.* 2008). VRE are a cause of surgical infections and infections of the urinary tract (Appelbaum and Jacobs 2005). However, the worldwide spread of *Enterococcus faecium* (complex-17) is characterised primarily through their resistance to ampicillin rather than vancomycin (Willems *et al.* 2005). Multidrug-resistant *S. pneumonia* continues to cause acute respiratory diseases and meningitis in many parts of the world, although the use of protein conjugated vaccines has reduced the number of new infections (Rice 2009). Epidemiological studies of isolates of one multidrug-resistant lineage of *S. pneumoniae* showed how genomic plasticity permitted the evolution of resistance to fluoroquinolones, rifampicin, and macrolides on multiple occasions over just 40 years (Croucher *et al.* 2011). This insight into the evolution of such pathogens may aid the design of future control strategies and vaccines (Woodford and Livermore 2009).

1.2.2 The decline in antibiotic discovery

Although there are a number of new antibacterial agents active against multi-drug resistant pathogens, only one, daptomycin, has a novel mechanism of action (Steenbergen *et al.* 2005). Daptomycin is proposed to undergo a primary structural change due to an interaction with calcium which facilitates a secondary structural change due to an interaction with the cell membrane, leading to membrane penetration (Jung *et al.* 2004). The insertion of daptomycin into the membrane creates a pore through which potassium ions are released, resulting in depolarisation and cell death (Muthaiyan *et al.* 2008; Van Bambeke *et al.* 2008). Other antibacterials recently approved for clinical use include linezolid (the first oxazolidinone in clinical use), telithromycin (a ketolide derived from clarithromycin), tigecycline (a broad-spectrum intravenous tetracycline) and ceftobiprole (a fifth generation cephalosporin). Other agents in preclinical development include the new glycopeptides oritavancin and dalbavancin, and the diaminopyrimidine iclaprim (Appelbaum and Jacobs 2005). So far, there is no widespread resistance to these newer antimicrobials, but occasional reports of therapeutic failure due to resistance (e.g. linezolid, daptomycin, telithromycin, or newer fluoroquinolones) means the development of new antimicrobials remains a high priority (Woodford and Livermore 2009). As well as glycopeptides, there are several other peptide antibiotics with potential as antibacterial agents. Although host-derived antimicrobial peptides (AMPs) such as pexiganan (a magainin), omiganan (an indolicidin), and iseganan (a protegrin) suffer from weak activity, nonspecific cytotoxicity, and apparent susceptibility to proteolysis, bacterial peptides such as the lantibiotics show more promise and are discussed in detail in Section 1.3 (Vaara 2009).

The need for further antimicrobial agents can be addressed through two strategies. Novel natural products can be discovered either by screening extracts in a whole-cell assay, or potentially through genome scanning. Alternatively, vast libraries of chemicals can be screened *in vitro* to identify novel scaffolds. Compounds discovered through either route are frequently subject to subsequent modification by medicinal chemists to ensure appropriate levels of stability, toxicity, distribution and elimination.

A change in emphasis between natural products and chemical libraries as the key source of new antimicrobial agents has occurred several times over the past seventy years. The development of penicillin as a medicine by Florey and Chain in the 1940's resulted in the first mass-produced antibiotic. This heralded the start of the 'golden age' of antibiotic discovery, lasting over fifty years (Newman *et al.* 2000). The main strategy during this time relied on the detection of bioactive natural products and their purification. The purification and structural elucidation of these compounds often posed technical challenges, adding to the financial cost of new natural product discovery. As technology improved, purification became less challenging, but despite the addition of multiple screening steps, dereplication (the rediscovery of known antibiotics) became a problem.

Medicinal chemistry improved existing activities through the synthetic tailoring of known molecular scaffolds, but in the long-term new lead compounds are needed (Fischbach and Walsh 2009). Combinatorial chemistry enabled the synthesis of new structural classes within the chemical space to form vast libraries of compounds that were then subjected to high-throughput screening (HTS) against drug targets to identify structures to take forward as lead compounds. Although there has been a lot of investment in HTS, only one *de novo* combinatorial compound was approved as a drug during the 25 years from 1981 to 2006 (Newman and Cragg 2007). This was the biaryl urea sorafenib, a multikinase inhibitor for the treatment of advanced renal cancer (Newman and Cragg 2007).

The main issue is that lead compounds identified through *in vitro* screens against specific targets need subsequent modification to improve permeability. Whereas whole-cell screening simultaneously tests for the ability to enter the pathogen as well as activity, leading more quickly to biologically active compounds (Baltz 2007). What has also become clear is that the choice of target to screen against is critical (Payne *et al.* 2007). The search for novel targets by a genomic approach has been unsuccessful and instead agents against under-exploited old targets may provide the answer.

More than two-thirds of clinically used antibiotics are natural products or their semisynthetic derivatives (Newman and Cragg 2007). The success of natural products can be no surprise when millions of years of evolution have selected for structures with biological activity. Natural products achieve structural complexity and bioactivity that are

unlikely to be achieved by combinatorial chemistry. The failure of high-throughput target-based screening of chemical libraries to discover new antibiotics has prompted a return to the discovery of new natural products with antimicrobial properties. There are now two additional methods of accessing novel natural products; metagenomics and genome scanning (Davies 2011). However the proof of principle of metagenomics as a validated method of discovery has not yet been established (Baltz 2007). Thus the latest strategy is a return to the search for natural products through genome mining in combination with the search for novel biodiversity (Zerikly and Challis 2009).

Genome mining is possible due to the fact that biosynthetic genes are clustered. This was first discovered in *S. coelicolor* where all the biosynthetic genes for the polyketide antibiotic actinorhodin were found to be clustered in a chromosomal region of approximately 26 kb (Malpartida and Hopwood 1986). The sequencing of *S. coelicolor* revealed many more biosynthetic gene clusters where the product has not yet been identified, referred to as cryptic gene clusters (Bentley *et al.* 2002). The *Streptomyces* genus is predicted to be capable of producing hundreds of thousands of antibiotics, of which <5 % have been discovered so far (Watve *et al.* 2001). The identification of cryptic gene clusters combined with genetic manipulation can be used in principle to stimulate the production of novel natural products from known species through heterologous expression, gene-targeting and ribosome engineering (Chiang *et al.* 2011; Winter *et al.* 2011). For example, genetic manipulation enabled the expression of the 150 kb type I modular polyketide synthase (PKS) cryptic gene cluster from *Streptomyces ambofaciens* ATCC 23877. The constitutive expression of a LAL regulator encoded within the cluster triggered expression of the biosynthetic genes, which lead to the identification of the stambomycins (Laureti *et al.* 2011).

The actinomycetes have proven themselves to be an important source of antibiotics in the past, and the use of new techniques such as high-throughput fermentation, the isolation of little-studied marine actinomycetes and the mining of genomes for cryptic pathways may facilitate the discovery of further antimicrobials (Baltz 2008; Nett *et al.* 2009). It has been noted that of the actinomycetes, the fragmenting forms are not such prolific producers, so focus has fallen on those whose morphology is characterised by a stable differentiated mycelium (F. Marinelli, personal communication). In particular, there is increasing emphasis on the isolation and screening of rare actinomycetes from under-explored ecological niches (Tiwari and Gupta 2011). The main problem with screening for novel anti-infectives is dereplication. In order to circumvent this, there has also been a drive to re-evaluate 'neglected' molecules through screening for other activities.

The use of novel biosynthetic pathways for natural products in combination with chemistry creates the opportunity for combinatorial biosynthesis as a source of further novel

compounds (Baltz 2007). There are other combinations of microbial genetics and synthetic chemistry that can yield new compounds. For example, mutasynthesis is where a mutant is fed a synthetic analogue that can be utilised by the biosynthetic enzymes to obtain a new product, while chemoenzymatic synthesis is where the enzyme that will synthesise the analogue is expressed in the deletion mutant to obtain the new product.

Over the past 15 years, pharmaceutical industry research into natural products has declined (Li and Vederas 2009). Although natural product anti-infective research still exists, generally in smaller companies, and is supported by government funded research, it is clear that the 'golden age' of discovery is over. As antibiotic treatments last for at most a couple of weeks, their profit-making potential is limited. Consequently, many companies have shut down their antimicrobial discovery and development groups in favour of investment in potential treatments for chronic/terminal illnesses.

1.3 Lantibiotics

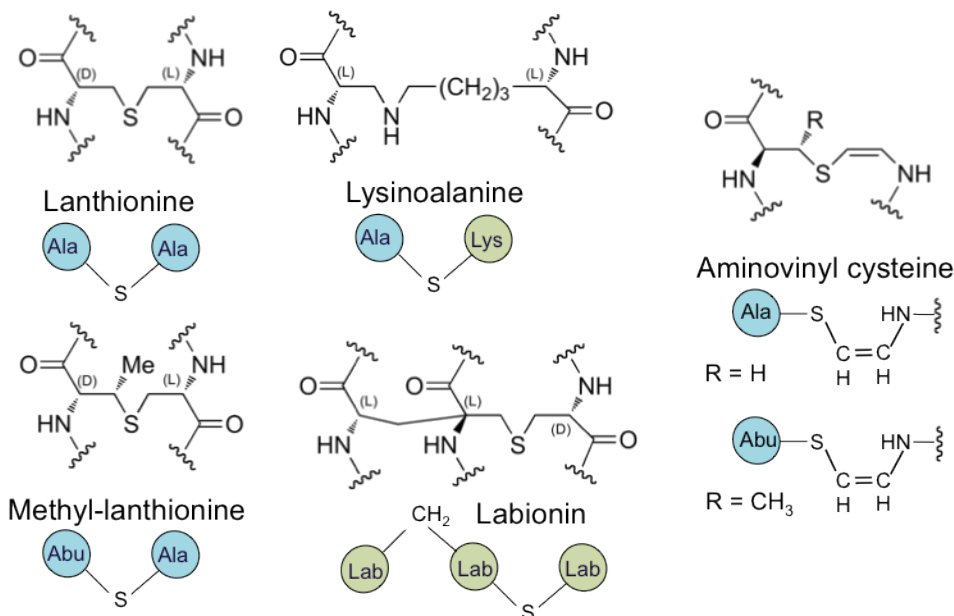
Lantibiotics are a distinct class of antimicrobial compounds, characterised by the presence of unusual post-translational modifications. They are typified by the presence of the ring-forming thioether amino acids lanthionine (Lan) and/or 3-methyl lanthionine (MeLan) (Sahl *et al.* 1995). These non-proteogenic amino acids constrain the molecule into a defined structural conformation that serves the dual purpose of ensuring functional biological activity and resisting proteolytic degradation (Sahl and Bierbaum 1998). As more lantibiotics are characterised, the range of observed post-translational modifications has expanded, but can be classified into those linking two residues within the peptide to form a bridge and those modifications on the α -carbon of a single amino acid (Figure 1.3). By definition, lantibiotics are lanthionine-containing peptides with antibiotic activity. However, some lanthionine-containing peptides have been discovered that do not appear to have antibiotic activity. Thus the term lantipeptides was recently coined to describe such compounds (Goto *et al.* 2010).

1.3.1 Classification

The bacteriocins were originally defined half a century ago to describe colicin-like antimicrobials produced by bacteria (Jacob *et al.* 1953). A classification scheme for the bacteriocins of lactic-acid bacteria was first put forward in 1993 but has subsequently been revised and simplified to form a scheme which incorporates bacteriocins regardless of the Gram-status of the producing organism (Klaenhammer 1993; Cotter *et al.* 2005a).

In both schemes the lantibiotics are class I bacteriocins, distinguished from other bacterially produced antimicrobial peptides on the basis of their extensive post-translational modifications (Klaenhammer 1993; Cotter *et al.* 2005a).

Bridge motifs



Unusual alpha-modified amino acids

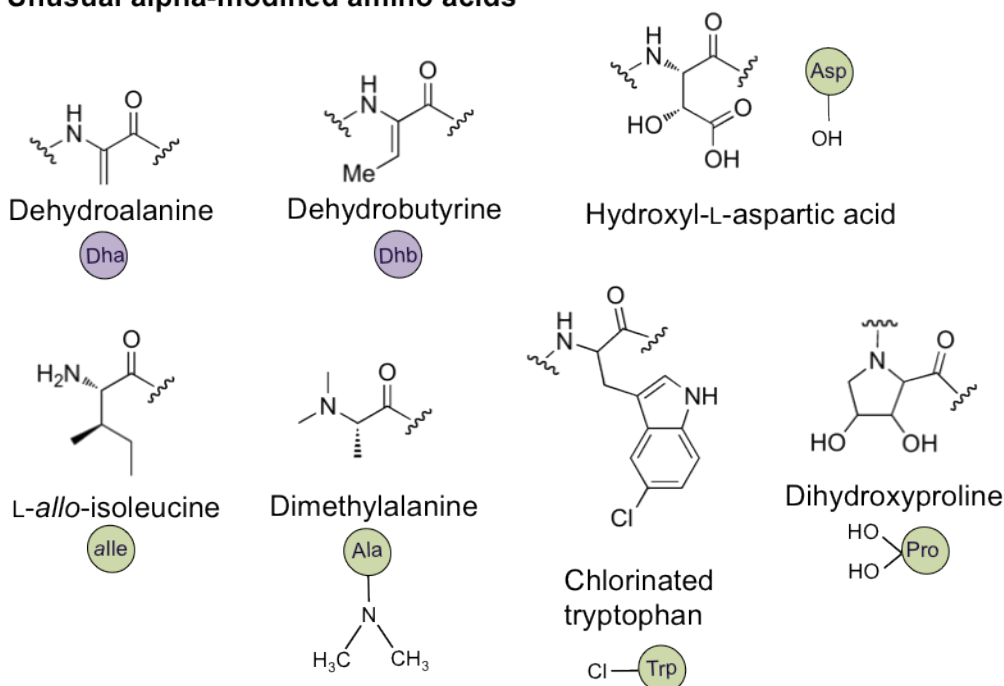


Figure 1.3 : Structural motifs introduced into lantibiotics by post-translational modification. Also shown is the schematic notation used in other figures, which includes the abbreviation for amino butyric acid (Abu).

Various lantibiotic subdivisions have been implemented over the years. Initially lantibiotics were assigned into two subclasses on the basis of structure: type A lantibiotics have an elongated shape, while type B lantibiotics are more globular in conformation (Jung and Sahl 1991). An alternative interpretation of structural features divides class I bacteriocins into seven subgroups. Each subgroup assumes the name of the most renowned member; nisin A, epidermin, Pep5, lactacin 481, mersacidin, cinnamycin and lactacin 3147 (Xie and van der Donk 2004). Other classification schemes have been proposed which are based on primary rather than tertiary structure. One such scheme focuses on the composition of the leader sequence present at the N-terminus of the precursor peptide (prepropeptide) that includes the future mature peptide sequence (propeptide) at its C-terminus. Lantibiotics modified by individual modification enzymes to produce (Me)Lan bridges typically have a conserved FNLD motif in the leader peptide, whereas lantibiotics modified by bifunctional modification enzymes typically have a GG or GA cleavage site and contain multiple Asp and Glu residues in the leader peptide (Chatterjee *et al.* 2005). However the steady discovery of lantibiotics with novel structures and functions demands a classification scheme with inherent flexibility. One such scheme primarily emphasises the amino acid sequence of enzymes and substrates involved in the post-translational modification of lantibiotics. In this scheme, class I lantibiotics such as nisin use distinct enzymes for the individual reactions involved in (Me)Lan formation, while class II, such as mersacidin, use dual-function enzymes. Class III lantipeptides are characterised by overt antimicrobial activity (Willey and van der Donk 2007). Another unifying feature of class III is the use of a dual-function enzyme with homology to the C-terminus of class II enzymes but with a Ser/Thr kinase domain (Goto *et al.* 2010). These classification systems are commonly used in conjunction with each other. Type AI lantibiotics are those which are elongated and use individual modification enzymes to produce (Me)Lan bridges, whilst type AII and type B use bifunctional modification enzymes and are elongated or globular respectively.

1.3.2 Lantibiotic producers

The first lantibiotic was discovered over eighty years ago (Rogers 1928). This was nisin, produced by the low-GC organism *L. lactis* and subsequently used as a preservative in the food industry (Chatterjee *et al.* 2005). Since then more than 60 lantibiotic peptides have been described (Field *et al.* 2010). Many are produced by other low-GC Gram-positive bacteria from which the biosynthetic gene cluster has been identified (Figure 1.4). For many, the structure and mode of action of the lantibiotic has been investigated (Section 1.3.3) and the function of genes encoded within the biosynthetic gene cluster studied (Section 1.3.4).

A smaller number of lantibiotic biosynthetic gene clusters have been sequenced in high-GC organisms (Figure 1.5). The cinnamycin gene cluster from *Streptomyces cinnamoneus cinnamoneus* DSM 40005 provided the first insight into the differences between gene clusters from high-GC as opposed to low-GC clusters (Widdick *et al.* 2003). Cinnamycin is a type B lantibiotic which is closely related to duramycin, duramycin B, duramycin C, and ancovenin, all of which are produced by actinomycetes as 19 residue propeptides which are post-translationally modified to introduce Lan residues in similar positions (Figure 1.6) (Widdick *et al.* 2003).

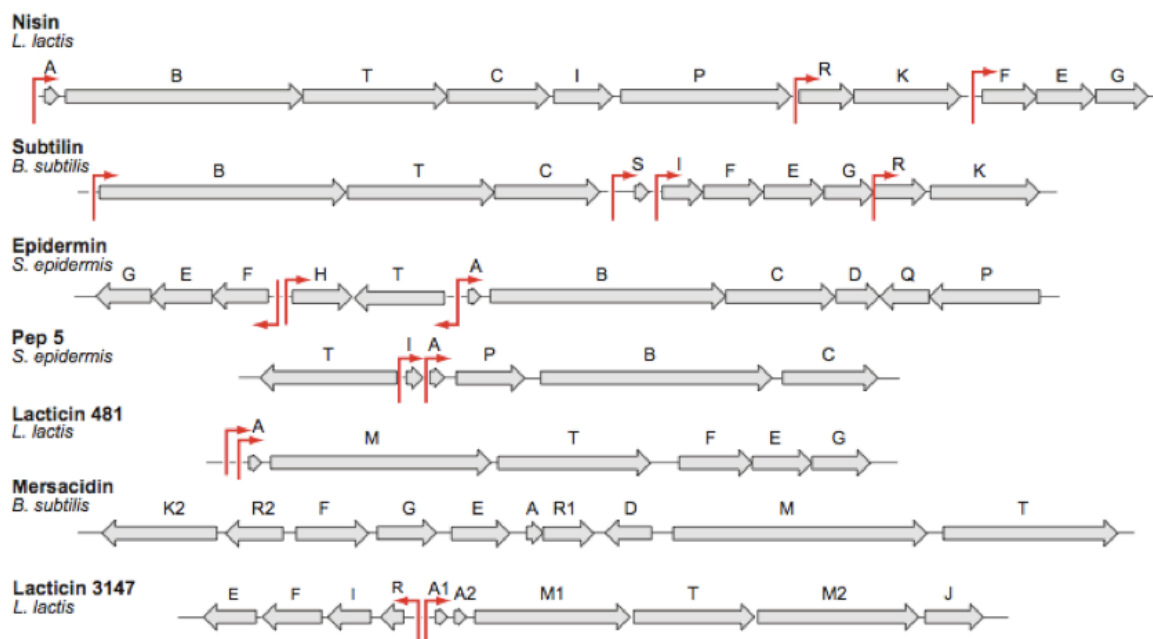


Figure 1.4 : A schematic of seven representative lantibiotic gene clusters from low-GC Gram-positive bacteria.

The gene clusters for type AI lantibiotics nisin, subtilin, epidermin and Pep5, the type AII lantibiotic lactacin 481, the type B lantibiotic mersacidin and the two-component lantibiotic lactacin 3147 are depicted. Several genes are represented in all the displayed gene clusters; i.e. the prepropeptide (A, S), the dehydratase and cyclase (B, C, M) and the ABC transporter (T). Promoter position and direction (where known) are indicated with red arrows. Figure adapted from Willey and van der Donk *et al.* 2007.

Actagardine was initially named gardimycin (Somma *et al.* 1977). This lantibiotic is produced as a 19 residue propeptide from *Actinoplanes garbadinensis* and post-translationally modified to introduce one Lan and three MeLan bridges (Figure 1.6) (Zimmermann *et al.* 1995). Deoxyactagardine B (DAB) is a 19 residue type B lantibiotic structurally similar to actagardine (Figure 1.6). The *gar* gene cluster from *A. garbadinensis* ATCC 31049, and the *lig* gene cluster from *Actinoplanes liguriae* NCIMB41362 have been

cloned and sequenced (Figure 1.5) (Boakes *et al.* 2009; Boakes *et al.* 2010). Despite the presence of a GarO-like mono-oxygenase, DAB lacks the sulfoxide bond seen in actagardine. The only other difference between DAB and actagardine are two amino acids at positions 15 and 16 (Figure 1.6).

The biosynthetic gene clusters of cinnamycin, actagardine and DAB share common features, identifying key differences between high-GC as opposed to low-GC clusters (Widdick *et al.* 2003; Boakes *et al.* 2009; Boakes *et al.* 2010). All lack the typical LanT-type lantibiotic exporter, instead encoding a two-component ABC transporter (Section 1.3.4.2). Also missing is a LanP-type protease to cleave the leader peptide; instead the protease is likely encoded elsewhere in the genome (Section 1.3.4.3.1).

Michiganin A has been purified and characterised from the tomato pathogen *Clavibacter michiganensis* subsp. *michiganensis* and resembles actagardine. Michiganin A is a type B lantibiotic produced as a 21 residue propeptide into which three MeLan bridges are introduced (Figure 1.6) (Holtsmark *et al.* 2006).

A screen of extracts from uncommon actinomycetes by Vicuron Pharmaceuticals identified five potentially novel lantibiotics. Two of these lantibiotics were identified from members of the *Streptosporangiaceae*; planosporicin from *Planomonospora* sp. DSM 14920 and microbisporicin from *Microbispora* sp. ATCC PTA-5024 (Castiglione *et al.* 2007; Castiglione *et al.* 2008). Planosporicin is discussed in detail in Section 1.3.5. The gene cluster for microbisporicin biosynthesis was subsequently cloned and characterised from *Microbispora corallina* (Figure 1.5) (Foulston and Bibb 2010). The cluster revealed a typical actinomycete lantibiotic exporter (MibTU) but also several enzymes not previously observed in any lantibiotic gene clusters. In particular, a tryptophan halogenase (MibH) is proposed to be responsible for the chlorination of tryptophan observed in microbisporicin, and the cluster appears to be regulated through a combination of an extracytoplasmic function sigma factor (MibX) and anti-sigma factor (MibW) as well as a helix-turn-helix DNA binding protein (MibR) (Foulston and Bibb 2011).

SapB is a lantipeptide encoded by *ramS* in the *ram* operon which is regulated by RamR (Figure 1.5). RamS is subject to post-translational modification to introduce four dehydroalanine (Dha) residues and two Lan bridges to create SapB (Figure 1.6). This lantipeptide does not possess antimicrobial activity but is instead a morphogenetic peptide important for aerial mycelium formation in *S. coelicolor* (Kodani *et al.* 2004). Similarly, *Streptomyces tendae* produces SapT, a lantipeptide with three MeLan bridges and one Lan bridge (Figure 1.6). SapB and SapT are both amphiphilic peptides which serve as biosurfactants to facilitate the emergence of newly formed aerial hyphae (Kodani *et al.* 2005).

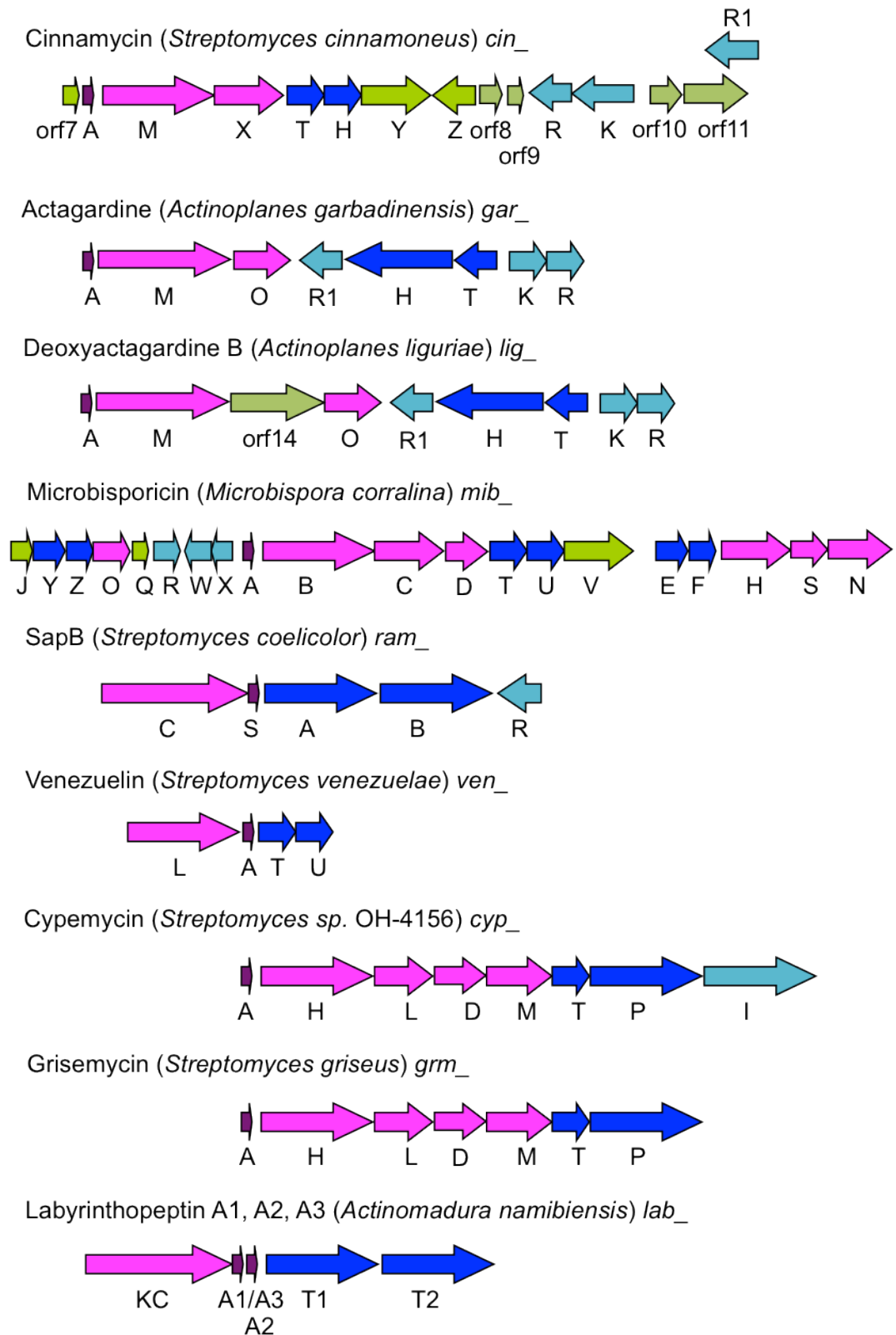


Figure 1.5 : A schematic of actinomycete lantibiotic biosynthetic gene clusters.

Shown are genes encoding the prepeptide (purple), modification enzymes (pink), two-component ABC export systems (dark blue), putative regulators (light blue) and genes of unknown function (green) as described in the text.

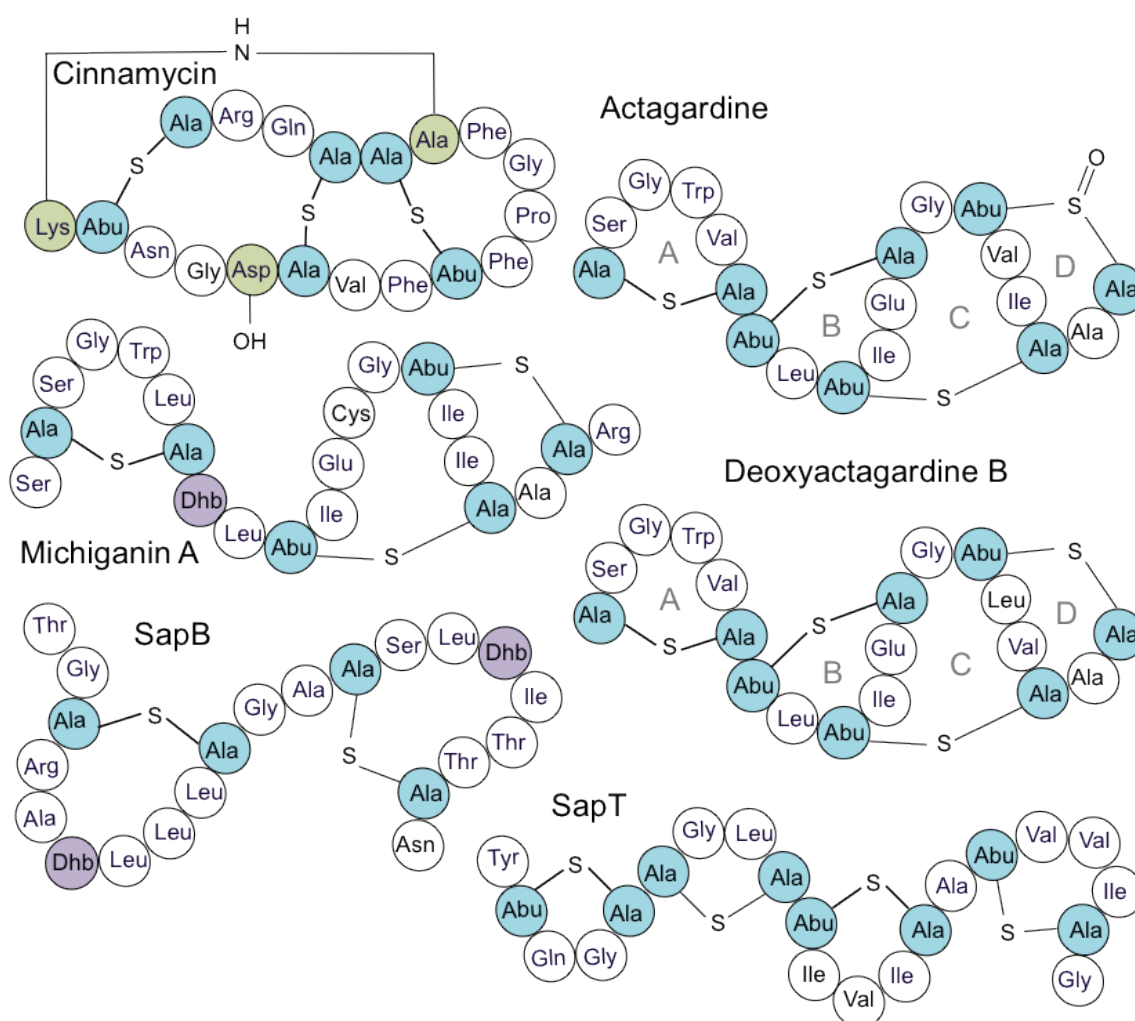


Figure 1.6 : Representatives of lantibiotics produced by actinomycetes.

Cinnamycin (Kaletta *et al.* 1991), actagardine (Zimmermann and Jung 1997), Deoxyactagardine B (Boakes *et al.* 2010), michiganin A (predicted) (Holtsmark *et al.* 2006), sapB (Kodani *et al.* 2004), sapT (Kodani *et al.* 2005), cypemycin (Claesen and Bibb 2010), grisemycin (Claesen and Bibb 2011) and labyrinthopeptins A1, A2 and A3 (Meindl *et al.* 2010). Residues are; unmodified (white), dehydrated (purple), dehydrated and cyclised residues (blue), other modifications (green). The same shorthand notation is used as defined in Figure 1.3. Abbreviations refer to; amino butyric acid (Abu), dehydroalanine (Dha), lanthionine (Ala-S-Ala), dehydrobutyrine (Dhb), methyl lanthionine (Abu-S-Ala), L-*allo*-isoleucine (*alle*) and labionin (Lab). (Me)Lan rings of lantibiotics with structural similarity to mersacidin are labelled in grey.

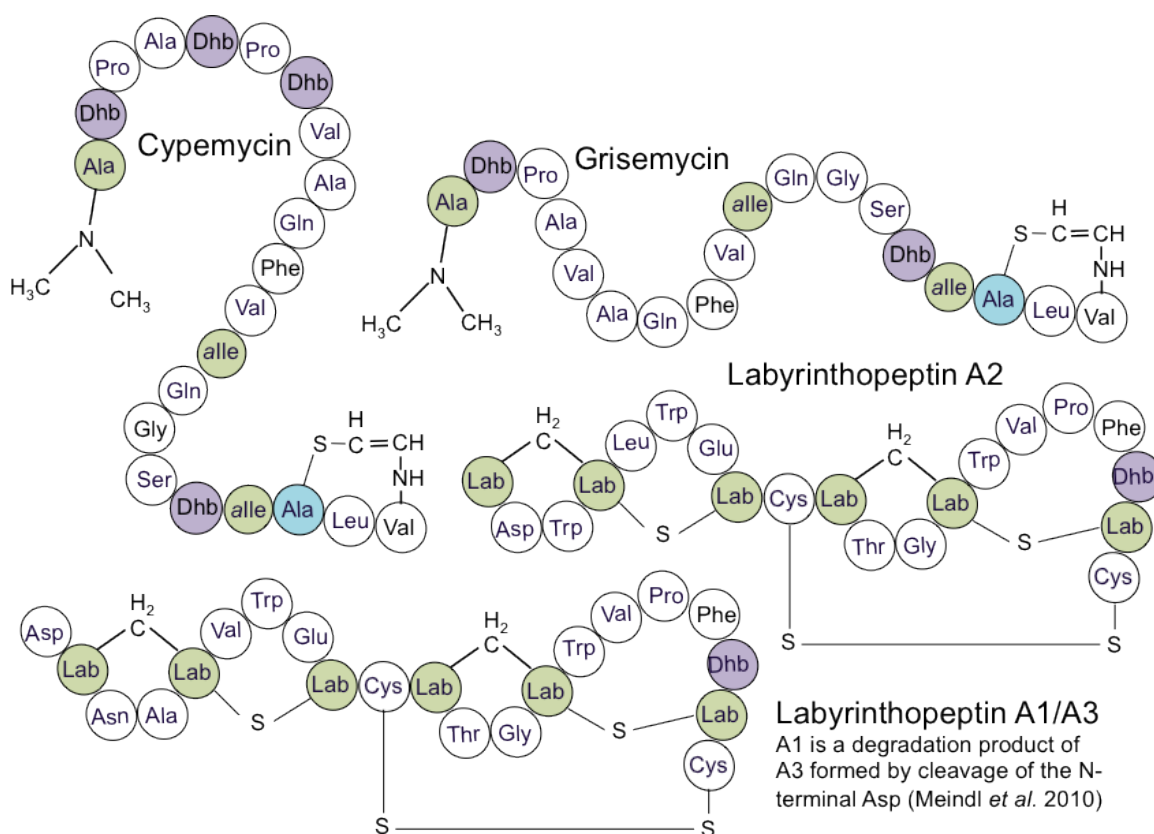


Figure 1.6 : Representatives of lantibiotics produced by actinomycetes (continued).

Genome scanning identified a gene cluster from *S. venezuelae* that encoded a putative prepeptide (VenA) and an unusual lanthionine synthetase (VenL) (Figure 1.5) (Goto *et al.* 2010). Homologues to this lanthionine synthetase, generically termed LanL, were identified in other gene clusters and are discussed further in Section 1.3.4.1.4. *In vitro* enzyme activity of VenL was used to produce venezuelin through the introduction of one Dha and three dehydrobutyrine (Dhb) residues; a Cys residue located in the nascent peptide chain was added to one of the Dhb residues to form a Lan bridge (Goto *et al.* 2010). Venezuelin has a globular structure, but its biological activity and function remain to be determined (Goto *et al.* 2010).

Cypemycin is a peptide antibiotic produced by *Streptomyces* sp. OH-4156 (Komiyama *et al.* 1993). Although it displays some antimicrobial activity against *M. luteus*, it also possesses significant cytotoxic activity against P388 leukemia cells. The cloning and characterisation of the cypemycin gene cluster (Figure 1.5), led to the identification of a new class of modified peptide natural products, the linaridins (Claesen and Bibb 2010). These peptides are ribosomally synthesised and post-translationally modified to introduce a number of lantibiotic-like structural features, notably the dehydration of serine and threonine residues. However, the characteristic (Me)Lan bridges of lantibiotics are not

observed (Figure 1.6). Indeed the enzyme(s) responsible for dehydration are completely different to those of lantibiotic clusters (Section 1.4.3.1.5). Other members of this class have subsequently been identified and characterised, such as grisemycin produced by *Streptomyces griseus* IFO 13350 (Figure 1.5 and 1.6) (Claesen and Bibb 2011).

The labyrinthopeptins were identified recently as a product of the actinomycete *Actinomadura namibiensis* DSM 6313 (Meindl *et al.* 2010). These ribosomally synthesised peptides are post-translationally modified to introduce the amino acid labionin, described in Section 1.3.4.3.5.3 (Figure 1.3). The sequencing of the gene cluster revealed a protein kinase-cyclase (LabKC) that phosphorylates serines then dehydrates these phosphoserines to Dha residues which are subsequently cyclised to form the labionin (Figure 1.5) (Meindl *et al.* 2010). Although labyrinthopeptins A1, A2 and A3 are structurally very different to other type III lantipeptides (Figure 1.6), the presence of a kinase-cyclase modification enzyme assigns it as a member of the type III class.

The thiopeptides are a class of ribosomally synthesised, heterocycle-containing peptides produced by a range of phyla, including some actinomycetes. One striking example is goadsporin produced by *Streptomyces* sp. TP-A0584 (Onaka *et al.* 2005). This polypeptide antibiotic contains thiazole and oxazole rings introduced through enzymes which appear either similar to lanthionine synthetases or enzymes involved in microcin biosynthesis. The microcins are a class of ribosomal peptides, some of which are extensively post-translationally modified to create a range of structures distal to other enterobacteria. GodE bears resemblance to McbC involved in Microcin B17 formation, while GodF and GodE are more similar to LanB-type dehydratases of lantibiotic gene clusters (Onaka *et al.* 2005).

1.3.3 Lantibiotic mode of action

The two main methods through which antibiotics target bacteria are the inhibition of either cell wall biosynthesis or protein synthesis. Both these targets involve molecules which are not found in eukaryotes so are also used widely in clinical treatments of bacterial infections. The most commonly used antibacterials are the β -lactam antibiotics such as the penicillins, aminopenicillins and cephalosporins. Other chemical classes of secondary metabolites with antibiotic properties are the aminocoumarins, tetracyclines, macrolides, aminoglycosides, lincosamides and the polyene antibiotics as well as peptide, glycopeptide and lipopeptide antibiotics. There are also a number of synthetic antibacterials which include the fluoroquinolones, oxalidinones, sulfonamides and nitroimidazoles.

Although the range of organisms inhibited by different lantibiotics varies, all lantibiotics are primarily active against Gram-positive species. However, if the outer membrane of Gram-negatives is disrupted, then some activity can be observed (Stevens *et al.* 1991). Many lantibiotics act through the inhibition of cell wall biosynthesis. Lipid II is a membrane-bound monomeric cell wall precursor that delivers cell wall monomers to the membrane for polymerisation into the cell wall (Figure 1.7). Inhibition of cell wall biosynthesis effectively inhibits cell division and increases likelihood of cell lysis due to internal osmotic pressure (Bauer and Dicks 2005; Breukink and de Kruijff 2006). Lipid II is a complex biosynthetic intermediate biosynthesised from UDP-GlcNAc in eight successive enzyme-catalysed steps (van Heijenoort 2001). Therefore, targeting lipid II poses a more difficult problem to pathogens than targeting a single enzyme whose active site could be modified more easily. The effectiveness of targeting lipid II as an antibacterial strategy is highlighted by the fact that it is targeted by several different classes of antibiotic (Figure 1.7). While other antimicrobial classes often target protein synthesis, lantibiotics are mainly cell wall synthesis inhibitors or pore formers. The formation of pores in the cytoplasmic membrane dissipates the proton motive force (PMF). The PMF is a combination of the chemical potential energy (the difference in concentration of the protons on each side of the membrane) and the electrical potential energy (a consequence of the charge separation across the membrane). These gradients drive ATP synthesis, so the collapse of the PMF means the cell cannot power energy-requiring reactions and consequently dies.

Nisin is the most extensively characterised of all lantibiotics and its mode of action is well understood. To date, six naturally occurring variants have been described, four from different strains of *L. lactis* and two from *Streptococcus* species. The efficacy of nisin and the absence of resistance despite over 40 years of use in the food industry as a preservative is thought to be due to a dual mechanism of cell killing. Nisin Z has been shown through confocal fluorescence microscopy to sequester lipid II from sites of cell wall biosynthesis such as the cell division septum (Hasper *et al.* 2006). The addition of nisin to *Bacillus* sp. and susceptible *L. lactis* was observed to perturb the normal distribution of lipid II, preventing its accumulation at specific sites. Thus the peptidoglycan units carried by lipid II are not available, uncoupling cell wall synthesis from membrane synthesis, and causing bacterial lysis (Hasper *et al.* 2006). The N-terminus of nisin consists of two thioether crosslinks referred to as the A and B rings (Figure 1.8). Nuclear magnetic resonance (NMR) studies have revealed these rings wrap around the pyrophosphate group of lipid II, allowing the formation of five intermolecular hydrogen bonds (Figure 1.9). These rings form a cage encapsulating the pyrophosphate moiety of lipid II (Hsu *et al.* 2004). The binding of lipid II by the N-terminus is a key step in the second mode of action mediated through the C-terminus. The recruitment of nisin

monomers to the membrane facilitates oligomerisation to form trans-membrane pores, leading to osmotic shock and cell lysis (Hasper *et al.* 2004). The region between the third and fourth (Me)Lan rings in nisin acts as a hinge and is essential for pore formation. Site-specific mutations to the nisin hinge region prevent pore formation without affecting the bacteriocidal activity of nisin (Hasper *et al.* 2004). Thus the hinge allows insertion of nisin molecules into the membrane where four nisin molecules form a complex with eight lipid II molecules to form a stable structure resistant to disruption by mild detergents (Hasper *et al.* 2004).

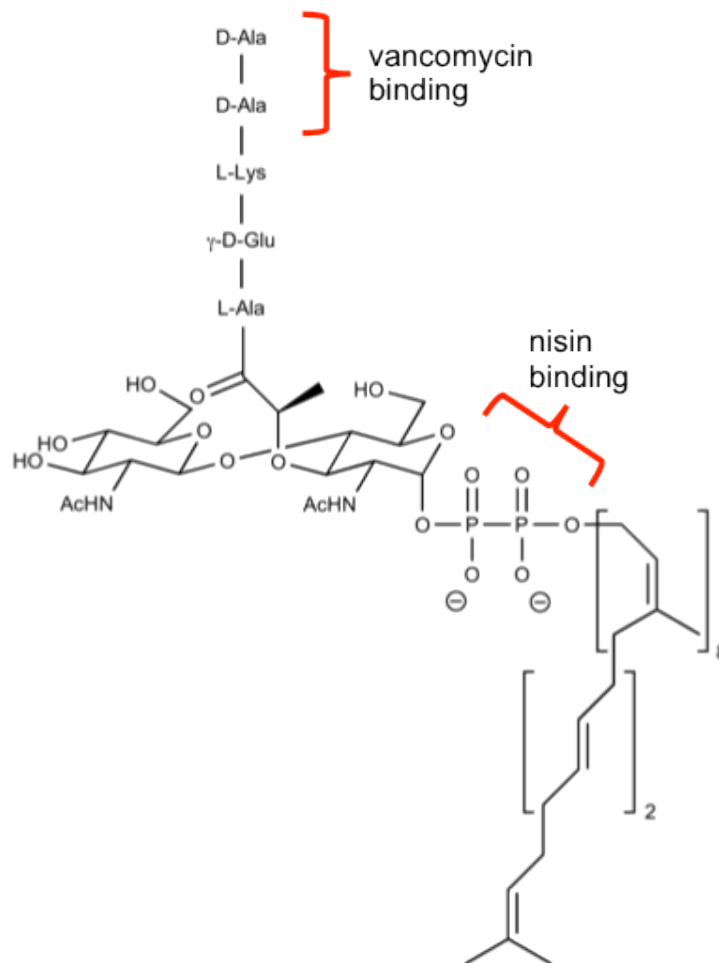


Figure 1.7 : The structure of lipid II.

The *N*-acetylglucosamine- β -1,4-*N*-acetylmuramic acid disaccharide (GlcNAc–MurNAc) is attached by a pyrophosphate to a membrane anchor of 11 isoprenoid units (undecaprenylpyrophosphate). The muramic acid bears a pentapeptide that contains a Lys residue for later cross-linking. The binding site of the glycopeptide vancomycin and the lantibiotic nisin are indicated in red.

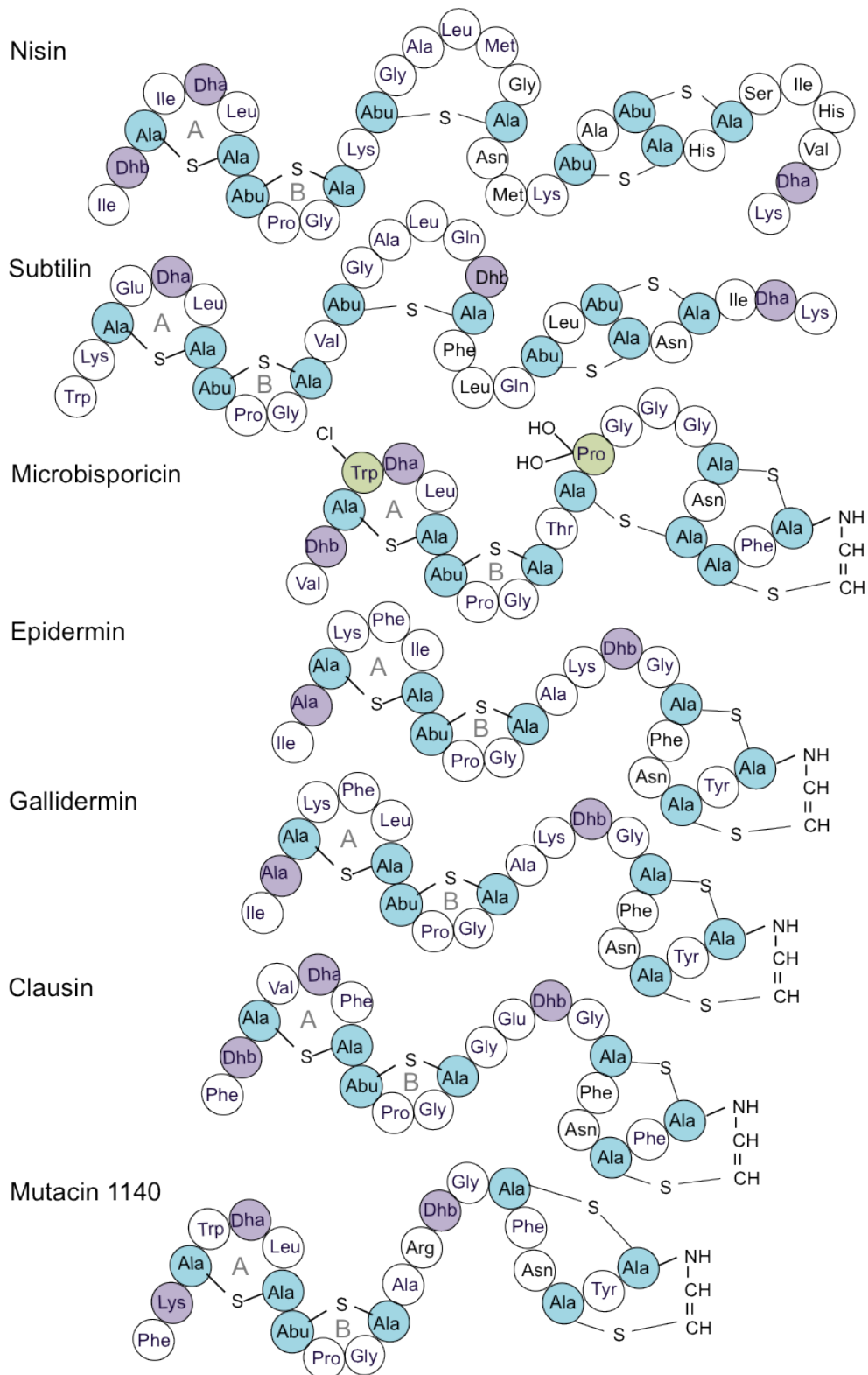


Figure 1.8 : Lantibiotics with N-terminal structural similarity to nisin.

Nisin, subtilin (Gross *et al.* 1973), microbisporicin (Castiglione *et al.* 2008), epidermin (Allgaier *et al.* 1985), gallidermin (Kellner *et al.* 1988), clausin (Bressollier *et al.* 2007) and mutacin 1140 (Smith *et al.* 2000). Residues are; unmodified (white), dehydrated (purple), dehydrated and cyclised residues (blue). The same shorthand notation is used as defined in Figure 1.3. The N-terminal A and B rings proposed to bind lipid II are labelled in grey.

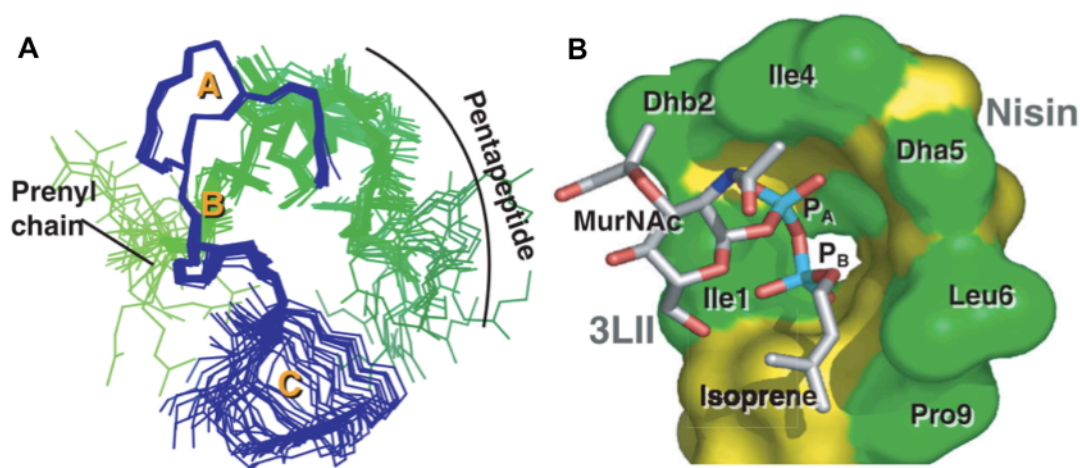


Figure 1.9 : NMR resolved solution structure of nisin in complex with the lipid II variant 3LII (shortened prenyl tail of 3 isoprene units instead of 11 in full-length lipid II).

A; Ensemble of the 20 lowest-energy NMR structures. The backbone of nisin (residues 1-19) is blue with the (Me)Lan rings A-C labelled, 3LII is green. **B**; The N-terminus of nisin (residues 1-12) encage the pyrophosphate moiety of 3LII. Nisin is shown as a space-filled model with side chains in green and labelled. 3LII is shown as a skeletal model with C, N, O and P in white, blue, red and cyan respectively. Figures from Hsu *et al.* 2004.

Crucially, it is the backbone of the nisin peptide chain which forms five hydrogen bonds to the pyrophosphate, with no involvement of the specific side-chains of each amino acid residue. Consequently, other lantibiotics with nisin-like A and B rings at the N-terminus are bacteriocidal due to their ability to bind lipid II (Bonelli *et al.* 2006; Hasper *et al.* 2006). Epidermin and gallidermin share a high degree of homology with the A and B rings of nisin and have been shown to bind lipid II (Figure 1.8). One key difference is at position 4 which is a lysine in epidermin and gallidermin, compared to isoleucine in nisin. The change in charge of the side-chain, from neutral to positive, may increase affinity to the pyrophosphate of lipid II, and it is interesting to note that gallidermin and epidermin have 10-20 times the bacteriocidal activity of nisin, perhaps due a higher efficacy of binding increasing the efficiency of cell wall inhibition (Bonelli *et al.* 2006; Rink *et al.* 2007b). Alternatively, there is some evidence to suggest that the C-termini of epidermin and gallidermin are able to interact with the prenyl chain of lipid II as a second mechanism of inhibition which could account for its increased binding efficacy (Hans-Jorg Sahl, personal communication). However there is also some evidence that like nisin, epidermin and gallidermin may also form pores, although the shorter length (22 amino acids, compared to 34 in nisin) means that the efficiency of lipid II-mediated pore formation by gallidermin appears to vary with the thickness of model membranes (Bonelli *et al.* 2006). Assays with intact cells, indicated that the pore-forming ability of gallidermin did not significantly

contribute to killing (Bonelli *et al.* 2006). Thus the high potency of epidermin and gallidermin is likely due to a higher affinity for lipid II.

Other lantibiotics sharing similarity to nisin at the N-terminus include clausin and mutacin 1140 (Figure 1.8). Clausin has been shown to interact with lipid II (Wiedemann *et al.* 2006a). Similarly, mutacin 1140 (also known as mutacin III) has demonstrated inhibition of cell wall synthesis (Smith *et al.* 2008).

The structure-activity relationships of most lantibiotics are poorly defined due to the lack of molecular detail elucidated from antibiotic-target interactions. The existence of unrelated compounds possessing similar mechanisms of action often means that conclusions drawn from structure-activity studies on one lantibiotic cannot readily be applied to another. For example mersacidin is a type B lantibiotic with no structural similarity to nisin, yet it inhibits cell wall biosynthesis through an interaction with lipid II. Mersacidin acts by inhibiting the transglycosylation reaction of bacterial cell wall biosynthesis via interaction with the disaccharide-pyrophosphate region of lipid II (Brotz *et al.* 1998). This interaction relies on the highly flexible Ala₁₂-Abu₁₃ region and the flexible glycine rich region (Figure 1.10) (Hsu *et al.* 2003). The removal of Glu₁₇ (located in the C-ring) inactivated mersacidin, implying that it may be involved in the Lipid II interaction (Szekat *et al.* 2003). A more extensive mutagenesis analysis revealed that the A-ring of mersacidin could not be extended and that the B-ring was particularly amenable to substitutions and insertions (but not deletions) (Appleyard *et al.* 2009).

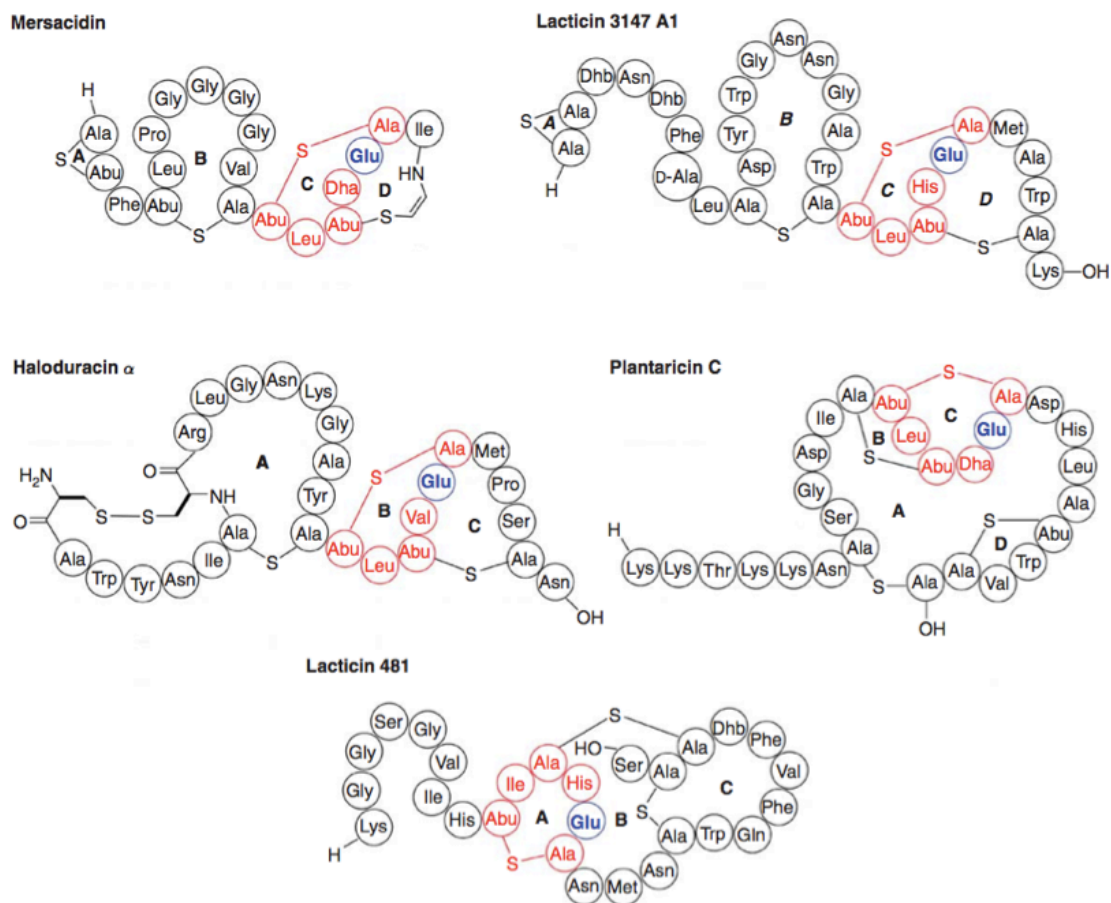


Figure 1.10 : Structures of lantibiotics containing a conserved methyl-lanthionine ring proposed to bind lipid II (in red).

The glutamate shown to be essential for mersacidin and haloduracin antimicrobial activity is illustrated in blue. The proposed structure of plantaricin C is shown (the stereochemistry of the lanthionine between residues 7 and 27 is not known). (Me)Lan rings are labelled A, B, C and D. Figure from (Cooper *et al.* 2010).

Actagardine and DAB are structurally similar to mersacidin and are proposed to bind lipid II in a similar manner (Zimmermann and Jung 1997). Specifically, five out of six residues in the C-ring of mersacidin and the B-ring of actagardine and DAB are conserved, forming a 'CTLTXEC' motif (Figure 1.6 and 1.10). However, the residue that differs (Dha at position 16 of mersacidin, compared to Ile at position 10 of actagardine and DAB) is essential for mersacidin activity. Site-directed mutagenesis to produce S16I mersacidin yielded only small amounts of the product with markedly reduced antimicrobial activity (Szekat *et al.* 2003). The broad activity spectrum of these lantibiotics differs, indicating other structural features play a role in bioactivity.

Other lipid II binding molecules of the mersacidin class include the α -peptide of the two-component lantibiotics lacticin 3147 and haloduracin produced by *L. lactis* subsp. *lactis* DPC 3147 and *Bacillus halodurans* respectively (Figure 1.10). Like actagardine, these only possess similarity with the C-ring of mersacidin (the Dha in position 16 of mersacidin is His in lacticin 3147 A1 and Val in haloduracin α). Opening the equivalent ring in lacticin 3147 abolished bioactivity (Field *et al.* 2007; Cooper *et al.* 2008). However it may only be the sequence conservation within the ring which is essential for lipid II binding; when haloduracin was synthesised without the Lan bridge, antibacterial activity was maintained (Cooper *et al.* 2008). There is overall similarity between the C-rings of mersacidin, lacticin 3147 A1 and plantaricin C, and the B-ring in haloduracin α (Figure 1.10). These rings have structural similarity to the A-ring of lacticin 481 (Figure 1.10) and the C-ring of lichenicidin α (Begley *et al.* 2009). However there is reduced sequence identity as the Leu and Dha at positions 14 and 16 of mersacidin are Ile and His in lacticin 481, and Val and Lys in lichenicidin α . The structural similarity combined with the observed lipid II-dependent activities of several of these lantibiotics implies that this ring similarity may correspond to antimicrobial activity for all members. For example, complete alanine scanning mutagenesis of each of the 30 amino acids of LtnA1 (and the 29 amino acids of LtnA2) revealed nearly all residues in the C-ring of LtnA1 were essential for bioactivity (Cotter *et al.* 2006).

In summary the mode of action has only been demonstrated for a small number of lantibiotics, but this can be extended to others on the basis of structural similarity. It is likely that members of the nisin-like subgroup (subtilin, ericin A and S) and the epidermin-like subgroup (gallidermin, mutacin 1140 and clausin) bind lipid II through the A and B rings. A second mechanism of cell death through pore formation is combined with lipid II binding in several nisin-like lantibiotics as well as other types of lantibiotics. In lacticin 3147, the two components interact synergistically. Lacticin 3147 A1 binds lipid II, however this peptide alone has only marginal antibiotic activity. Crucially, lacticin 3147 A2 is able to bind the LtnA1–lipid II complex, prompting strong inhibition of cell wall biosynthesis and the formation of pores (Wiedemann *et al.* 2006b). However, pore formation may also occur independently as a mechanism of action in other lantibiotics such as plantaricin C, a type AII lantibiotic classified in the lacticin 481 subgroup (Figure 1.10). This globular lantibiotic has a positively charged, unbridged N-terminus proposed to insert into the membrane due to electrostatic interactions with anionic phospholipids (Gonzalez *et al.* 1994; Turner *et al.* 1999). The type B lantibiotics are split into two groups. The mersacidin group bind lipid II in a different manner to the type AI lantibiotics. The structural motif responsible for binding is also observed in several of the two-component lantibiotics; lacticin 3147, lichenicidin and haloduracin. The cinnamycin group possesses a different

mode of action entirely. Cinnamycin binds to phosphatidylethanolamine in the membrane and thus disrupts membrane morphology and stability (Makino *et al.* 2003).

1.3.4 Lantipeptide biosynthesis

The generic locus symbol *lan* is used to refer to lantibiotic biosynthetic genes until a more specific name for each lantibiotic is assigned (e.g. in this work *psp* is used to refer to planosporicin biosynthetic genes) (Patton and van der Donk 2005). Lantibiotics are encoded as a prepropeptide by a structural gene (*lanA*) whose product is post-translationally modified then exported from the cell. The genes encoding these proteins are found clustered together either on the chromosome (e.g. subtilin) or on a plasmid (e.g. nisin) (Chatterjee *et al.* 2005). The prepropeptide consists of the future mature peptide sequence (propeptide) at the C-terminus and an N-terminal leader peptide. Numerous experiments on different lantibiotic systems suggest three different roles for the leader peptide.

There is much evidence for the leader peptide acting as a scaffold for the binding of lantibiotic modification enzymes. The expression of a chimera consisting of the subtilin leader fused to an N-terminal nisin - C-terminal subtilin structural region in a strain of *B. subtilis* 168 (encoding all of the cellular machinery for subtilin biosynthesis except for the prepropeptide), resulted in production of the corresponding mature lantibiotic (Chakicherla and Hansen 1995). Site-directed mutagenesis on different residues of the leader peptides of nisin, Pep5 and mutacin II demonstrate weak dependence of enzymatic processing on point mutations, indicating the recognition of the leader is through tertiary structural elements, not conserved residues (van der Meer *et al.* 1994; Neis *et al.* 1997; Chen *et al.* 2001). The *in vitro* reconstitution of lactacin 481 biosynthesis demonstrated that neither the length of the leader or any one specific residue is essential for correct modification. Instead it is the general properties of each amino acid which matter, for example the hydrophobic residue at position 7 is important for the maturation process (Patton *et al.* 2008).

There is also evidence for a role of the leader peptide in signalling for export. As described above, the subtilin leader fused to a nisin-subtilin structural region is modified and exported from *B. subtilis* 168 (Chakicherla and Hansen 1995). *L. lactis* expressing the *nisT* transporter was capable of excreting both unmodified and partially or fully post-translationally modified forms of prenisin and fusions of the leader peptide with non-lantibiotic peptides (Kuipers *et al.* 2004). The lantibiotic transporter is often capable of recognising similar lantibiotic leader peptides, however cleavage of the leader peptide to release the mature lantibiotic does not always occur. Expression of the subtilin leader

peptide DNA sequence fused to the sequence encoding pronisin Z in *L. lactis* resulted in production of the unmodified leader peptide of subtilin linked to a fully matured nisin Z structural region (Kuipers *et al.* 1993b). Leader peptides often contain the motif for cleavage by the protease. The purification of the N-terminal protease domain of LctT enabled an assessment of the *in vitro* production of lacticin 481 from variants of the prepropeptide (Furgerson Ihnken *et al.* 2008). Site-directed mutagenesis of the LctA leader peptide showed cleavage was slow if Glu₈ was mutated and abolished if the GG motif was mutated, whereas there was no defect if mutations were made in the propeptide or nonconserved residues in the leader peptide (Furgerson Ihnken *et al.* 2008). However, the observation that NisB and NisC have a role in the rate of nisin production could undermine the leader peptide role as a recognition motif for export (van Sappareoa *et al.* 2008). It is suggested that a channeling mechanism of prenisin transfer occurs between NisB, NisC and NisT so that once prenisin has found this cluster of enzymes, the leader peptide may be less important as transfer between enzymes occurs by diffusion.

A third role for the leader peptide is in host immunity. In many systems the modified lantibiotic shows no activity as long as the leader peptide remains attached. This protective role of the leader has been shown for several lantibiotics, including nisin (van der Meer *et al.* 1993; van der Meer *et al.* 1994), subtilin (Stein and Entian 2002; Corvey *et al.* 2003), lacticin 481 (Xie *et al.* 2004) and mutacin II (Chen *et al.* 2001).

1.3.4.1 Biosynthetic genes

1.3.4.1.1 LanB

lanB genes are characteristic of class I lantibiotic clusters. They encode a ~120 kDa protein responsible for the dehydration of serine and threonine amino acids to form 2,3-didehydroalanine (Dha) and (Z)-2,3-didehydrobutyrine (Dhb), respectively (Figure 1.11). The first evidence for a role in the formation of these characteristic non-proteogenic amino acids came through *in vivo* work. Targeted mutations in producing strains indicated a role for LanB in the dehydration of serine and threonine residues in Pep5 and nisin (Meyer *et al.* 1995; Koponen *et al.* 2002). Further work demonstrated that LanB is necessary for dehydration in a non-producing *L. lactis* strain. Expression of the *nisABTC* genes efficiently produced fully post-translationally modified prenisin, while expression of just *nisBT* could produce dehydrated prenisin without thioether rings and a dehydrated form of a non-lantibiotic peptide (Kuipers *et al.* 2004). Interestingly, the dehydrations of Ser/Thr yielding Dha/Dhb groups in thiopeptides are mediated by enzymes resembling LanB-type lantibiotic dehydratases (Li and Kelly 2010).

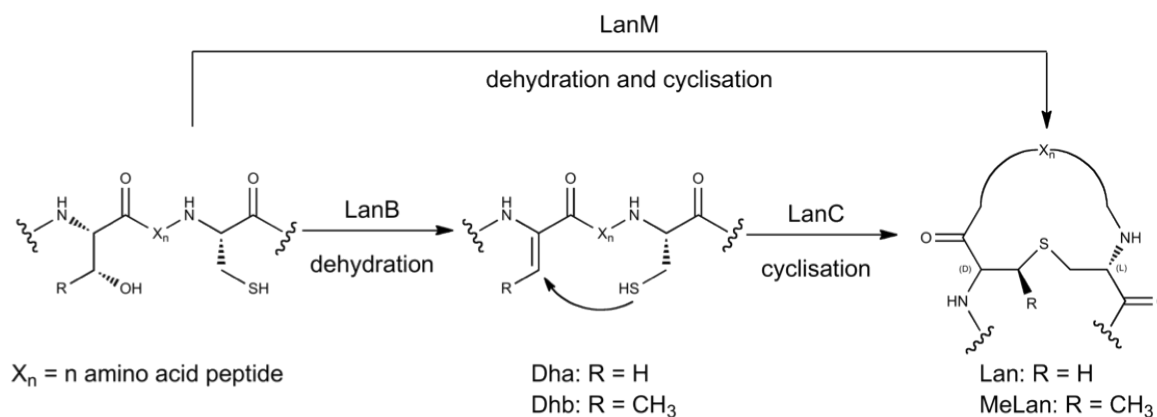


Figure 1.11 : Dehydration and cyclisation of lantibiotic prepeptides.

The dehydration of Ser/Thr to form Dha/Dhb is catalysed by LanB or LanM. The intramolecular addition of Cys thiols onto Dha/Dhb to form Lan/MeLan is catalysed by LanC or LanM.

Subsequently, focus has moved to the expression and purification of an active LanB in order to study the enzymatic reaction mechanism. Although this has so far proved elusive, a substantial amount of information has been accrued along the way. LanB enzymes are overall hydrophilic proteins, but contain some hydrophobic domains which imply it may be membrane associated. This is supported by the observed co-sedimentation of NisB and SpaB with membrane vesicles (Engelke *et al.* 1992). Yeast-two-hybrid and immunoprecipitation studies have suggested that LanBCT enzymes are only active in a multimeric lantibiotic synthetase complex. For example in nisin biosynthesis, yeast-two-hybrid revealed an interaction between NisB and NisC as well as an interaction between NisC and the NisT, and implied that at least two molecules of NisC and two molecules of NisT are part of the complex (Siegers *et al.* 1996). Likewise a complex of at least two each of SpaT, SpaB, and SpaC were demonstrated to be associated with the substrate subtilin through yeast-two-hybrid and coimmunoprecipitation experiments (Kiesau *et al.* 1997). Yet *in vivo* data indicates a lack of mutual dependence for activity, as deletions of *lanC* do not prevent the dehydration by LanB, and deletions of *lanB* do not prevent cyclisation by LanC (Meyer *et al.* 1995; Koponen *et al.* 2002).

In the absence of active purified LanB, heterologous expression systems have been established to demonstrate the substrate promiscuity of LanB through the expression of chimeras and non-lantibiotic peptides. The production of non-naturally occurring peptides containing a series of Dhb residues from *L. lactis* was achieved through expressing both *nisBT* and a plasmid encoding a specific leader peptide fusion construct (Rink *et al.* 2007c). Novel hexapeptides fused to the nisin leader and expressed in a *L. lactis* strain

containing the nisin modifying and export enzymes resulted in the production of correctly modified fusion peptides (Rink *et al.* 2005). A *L. lactis* strain containing *nisBTC* was found to effectively dehydrate and secrete a wide range of non-lantibiotic peptides, many of which also demonstrated ring formation (Kluskens *et al.* 2005). The non-lantibiotic peptides included a number of animal hormones such as vasopressin, angiotensin and erythropoietin. Post-translational modifications by NisB and NisC still occur when the nisin leader is preceded by a 27 amino acid Sec signal peptide, enabling exploitation of the Sec pathway of *L. lactis* for the secretion of dehydrated variants of therapeutic peptides (Kuipers *et al.* 2006).

1.3.4.1.2 LanC

lanC genes encode 400-450 amino acid enzymes responsible for cyclisation reactions yielding Lan and MeLan bridges. Dha and Dhb are subject to nucleophilic attack by the thiol group from a cysteine residue located elsewhere in the peptide chain. This intramolecular cyclisation reaction forms characteristic thioether crosslinks. Those formed from Dha are Lan bridges, while those formed from Dhb are MeLan bridges (Figure 1.11).

Section 1.3.4.1.1 referred to studies in which LanBCT enzymes were found only to be active in a multimeric lanthionine synthetase complex (Siegers *et al.* 1996; Kiesau *et al.* 1997). Evidence for the role of NisC was initially provided through *in vivo* work. His-tagged nisin precursors expressed in a *nisC* mutant were dehydrated but did not contain (Me)Lan bridges, implying NisC is required for nisin maturation (Koponen *et al.* 2002). Subsequently, purified NisC was able to perform all five cyclisation reactions on a dehydrated NisA substrate, confirming the activity of NisC *in vitro*. (Li *et al.* 2006). Further *in vivo* work has demonstrated the relaxed substrate specificity of LanC enzymes. NisC is able to cyclise a variety of peptides unrelated to nisin, as long as they are fused to the C-terminus of the NisA leader peptide (Rink *et al.* 2007a).

Overall, sequence identity between LanC enzymes is low, only 20-30 %, but there are a few strictly conserved histidine and cysteine residues. These were proposed to act as ligands to coordinate zinc binding which likely has a role in the activation of Dha and Dhb residues (Okeley *et al.* 2003). This has been verified through the X-ray-structure of NisC (Li *et al.* 2006). Subsequent mutagenesis identified residues essential for catalysis in NisC and SpaC (Helfrich *et al.* 2007; Li and van der Donk 2007).

In a regioselective reaction, bond formation occurs in one direction. In type A lantibiotics, cyclisation occurs in the C-to-N terminal direction to form an endocyclic enolate, whereas some substrates of type B lantibiotics undergo cyclisation in the opposite direction to form

an exocyclic enolate (Zhou and van der Donk 2002). There is evidence the leader peptide controls the regioselectivity of (Me)Lan formation (Section 1.3.4.1.3). In the cyclisation reaction, the thiol functional group of the cysteine residue undergoes a 1,4-nucleophilic attack on the Michael acceptor system of the Dha/Dhb. It has been observed that the cyclisation of short peptide analogues of type A lantibiotic ring systems occurs spontaneously to form the naturally occurring Lan or MeLan diastereomer (Zhou and van der Donk 2002). However in type B lantibiotics, some cyclisations occur in the N-to-C terminal direction and there is evidence to suggest that in the absence of LanC, a mixture of stereoisomers is created (Zhou and van der Donk 2002). Likewise there is an indication of a similar tendency towards the natural stereoisomer in nonenzymatic cyclisations of precursors. Although a mixture of stereoisomers is created, the major product has the stereochemistry of the natural lathionine. However these results are only clear-cut when the substrate only has the capacity to form one ring. In more complex substrates chemoselectivity is apparent, as cysteine residues will react preferentially with Dha rather than Dhb to form Lans over MeLans. (Zhou and van der Donk 2002). Therefore the main challenge for LanC enzymes is not the activation of the reaction or the control of stereochemistry and regioselectivity of this cyclisation but is instead to overcome the chemoselectivity that would otherwise result in Lan formation over MeLan formation.

1.3.4.1.3 LanM

As an alternative to the individual LanB and LanC genes, type AII lantibiotic gene clusters encode a single bifunctional LanM responsible for both the dehydration and cyclisation reactions (Figure 1.11). While the N-termini of LanM enzymes share no similarity to LanB enzymes, the C-termini share ~20 % identity to LanC enzymes (Okeley *et al.* 2003). Subsequent to the successful *in vitro* reconstitution of LanC enzymes, three LanM enzymes were purified for use in assays. LctM, HalM1 and HalM2, responsible for the modification of lacticin 481, haloduracin α and haloduracin β , respectively, all demonstrated dehydratase and cyclase activity *in vitro* (Xie *et al.* 2004; McClerren *et al.* 2006). Studies on LctM and HalM2 suggested that substrate binding via an N-terminal leader results in dehydration and cyclisation with a strong preference for N to C directionality, providing evidence on the role of the leader peptide in controlling the regioselectivity of (Me)Lan formation (Levengood *et al.* 2007).

The two functions performed by LanM enzymes are likely achieved through different active sites. The detection of completely dehydrated substrates with no or little cyclisation supports the hypothesis that all the dehydration reactions are catalysed before ring cyclisation begins (Miller *et al.* 2006). Conserved residues in the cyclase domain on LanM

include three zinc ligands and the equivalent of His₂₁₂ in NisC, the residue proposed to be involved in the activation of Dha/Dhb (Li and van der Donk 2007). In accordance with this, site-directed mutagenesis of these residues in LctM decreased or abolished cyclisation without affecting the dehydration of lactacin 481 (Paul *et al.* 2007). These results imply separate dehydratase and cyclase domains perform the two reactions independently.

In general, the biosynthetic gene clusters of two component lantibiotics encode an individual LanM for each peptide. Exceptions occur when the two substrates are highly similar, such as with cytolysin (Gilmore *et al.* 1994). This would tend to imply substrate specificity, but as with LanC enzymes, LanM have remarkable versatility in cyclisation of a variety of peptides. Purified LctM was able to process nonlantibiotics and even non-proteinogenic residues such as homocysteine and β -homocysteine to generate novel thioether linkages (Chatterjee *et al.* 2006). This promiscuity has subsequently been exploited in the preparation of analogues of both type A1 (e.g. subtilin (Liu and Hansen 1992)) and type B (e.g. mersacidin (Szekat *et al.* 2003)) lantibiotics. Whilst most studies have focused on model antibiotic-producing bacteria, the smaller genomes of cyanobacteria have made use of the promiscuous nature of LanM enzymes. The planktonic marine bacterium *Prochlorococcus* MIT9313 encodes a single LanM enzyme able to recognise up to 29 different linear ribosomally synthesised peptides as substrates to generate an array of polycyclic, conformationally constrained products with highly diverse ring topologies (Li *et al.* 2010).

1.3.4.1.4 RamC/LanL

RamC is the cyclase responsible for the formation of Lan bridges in the lantipeptide SapB introduced in Section 1.3.2 (Kodani *et al.* 2004). The structure of this cyclase is unlike that of the LanB-like or LanM-like enzymes. The N-terminus of RamC exhibits similarity to catalytic domains of Ser/Thr kinases while the C-terminus has some sequence similarity to the C-terminal domain of LanM (Hudson and Nodwell 2004). The kinase domain is thought to be involved in phosphorylation of the Ser residues to facilitate dehydration, with the C-terminal domain catalysing subsequent Lan formation (despite lacking the characteristic zinc-binding and catalytic residues of LanC-like domains).

More recently it has become apparent that RamC is part of a larger group of lantipeptide synthetases, sometimes referred to as class III (whereas LanB/C are class I and LanM are class II). This class includes enzymes with the generic name LanL, of which VenL was the first to be characterised (Goto *et al.* 2010). Individual expression of each VenL domain revealed that Ser and Thr residues are first phosphorylated by the kinase domain, followed by an elimination catalysed by the lyase domain, which results in Dha and Dhb residues

(Goto *et al.* 2010). Subsequent addition of the Cys thiols onto these dehydrated residues is performed by the cyclase domain, which contains a LanC-like zinc-binding site and active-site residues, yielding (Me)Lan bridges (Goto *et al.* 2010). Significantly, the expression of the VenL N-terminus is the first example of an *in vitro* reconstituted peptide dehydratase, as previously only bifunctional lantipeptide synthetases had been reconstituted *in vitro*.

1.3.4.1.5 Linaridins

Recently, a new class of post-translationally modified peptides have been identified in which modifications previously described in lantibiotics are carried out by unusual enzymes or via an alternative modification pathway to those described in lantibiotic biosynthesis. These clusters contain *cypL* homologs and are proposed to be involved in the biosynthesis of non-cyclised peptides containing dehydrated amino acids, referred to as linaridins (Claesen and Bibb 2010). The type member of the linaridin family is cypemycin (described in Section 1.3.2). The *cyp* cluster does not contain a conventional LanB enzyme, instead CypH and CypL are essential for cypemycin biosynthesis. It is proposed that CypH and/or CypL are responsible for dehydration of the Thr residues of cypemycin to Dhb, however neither CypH nor CypL activity could be reconstituted *in vitro* (Claesen and Bibb 2010).

1.3.4.2 ABC transporters

ATP-binding cassette (ABC) transporters transduce the free energy of ATP hydrolysis to power the translocation of substrates across cell membranes. The ABC superfamily consists of transporters with two distinct domains, the transmembrane domain (TMD) and the nucleotide-binding domain (NBD). The TMD structure is variable but always contains several alpha helices enabling the protein to be embedded in the membrane bilayer where it recognises the substrate and undergoes a conformational change resulting in transport across the membrane. In contrast the NBD is highly conserved in sequence and is located in the cytoplasm where it forms the site for ATP binding.

The TMD and NBD may be present as two domains of one protein or be encoded on separate peptides. The structural architecture of each ABC transporter consists minimally of two TMDs and two NBDs, both of which may be present as homodimers or heterodimers. Four individual polypeptide chains consisting of two TMD and two NBD subunits, may combine to form a full transporter. More commonly, the NBD and TMD domains are fused together into a single protein. Many exporters consist of a homodimer

of two monomers containing an N-terminal TMD and a C-terminal NBD. In an extension of this, all four domains may be fused into a single polypeptide chain arranged as TMD-NBD-TMD-NBD.

ABC transporters encoded in lantibiotic gene clusters demonstrate a wide variety of architectures. The type of transporter observed in each cluster depends both on the type of lantibiotic exported and the phylogeny of the producing organism. Low-GC producers of type AI lantibiotics typically encode a single LanT with significant homology to the hemolysin B ABC transporter. The *E. coli* hemolysin exporter HlyB is a single polypeptide encoding two TMD and two NBD domains. NisT consists of an N-terminal TMD containing six transmembrane helices and a C-terminal NBD, and dimerises to form the functional transporter (Qiao and Saris 1996).

The ABC transporters of type All lantibiotic clusters from low-GC organisms are similar in architecture to type AI clusters but many bear an additional domain at the N-terminus. This domain has peptidase activity to cleave the leader peptide as the lantibiotic is exported from the cell. This peptidase recognises a double-glycine cleavage site in combination with other features of the substrate. This confers the substrate specificity which prevents the heterologous production of lactacin 481 through the LcnC transporter from an *L. lactis* strain expressing LctA and LctM genes (Uguen and Uguen 2002).

In contrast, high-GC producers of type AI and type B lantibiotics encode the NBD and TMD domains on two different peptides. For example the cinnamycin, actagardine and microbisporicin exporters are CinTH, GarTH and MibTU respectively (Widdick *et al.* 2003; Boakes *et al.* 2009; Foulston and Bibb 2010). It is presumed that two copies of each protein come together to form a tetramer consisting of two homodimers.

A second ABC transporter is present in many lantibiotic gene clusters to provide a mechanism of producer immunity. Many of these proteins are homologous to members of the type B ABC transporters of the HisP family. The first to be identified was the NisFEG transporter of *L. lactis* 6F3. Deletion mutants of *nisF*, *nisE* and *nisG* continued to produce nisin but were more sensitive to nisin compared to the WT strain (Siegers and Entian 1995). The functional transporter has the architecture NisF₂EG with NisE and NisG forming a heterodimer of TMDs and NisF forming a homodimer of two NBDs.

Several proposals have been suggested to link the mechanism of lantibiotic action to the mechanism of the ABC transporter. It has been proposed that LanFEG proteins may scavenge lantibiotics that have bound lipid II and/or inserted into the membrane and secrete them back out of the cell (Chatterjee *et al.* 2005). This is supported by the observation that lantibiotic concentrations are higher in the supernatant from cells that express the LanFEG transporter, indicating a reasonable efficacy of releasing the

lantibiotic from the cell surface. Evidence for this comes from heterologous expression of *nisFEG* in a nisin-sensitive strain of *B. subtilis* which conferred immunity to nisin. This was assessed through quantitative *in vivo* peptide release assays which revealed that NisFEG reduced the number of nisin molecules associated with the membrane, presumably through transport back into the extracellular space (Stein *et al.* 2003). Similarly, HPLC analysis of the epidermin-producing strain *Staphylococcus epidermidis* Tu3298 indicated that the EpiFEG transporter functions to release epidermin from the cytoplasmic membrane (Otto *et al.* 1998).

1.3.4.3 Other modification enzymes

1.3.4.3.1 Protease

As mentioned in Section 1.3.4.2, class II LanT transporters have an additional N-terminal peptidase domain and are proposed to cleave the leader peptide to release the mature lantibiotic concomitantly with export out of the cytoplasm. The protease domain of LctT has been experimentally validated through purification and use *in vitro* to cleave the LctA prepropeptide at the double-glycine site to release mature lactacin 481 (Ihnken *et al.* 2008).

In contrast, class I lantibiotic gene clusters do not use a bifunctional transporter. Instead, many encode a separate subtilisin-type serine protease, which may be located intracellularly, extracellularly, or anchored to the cell wall surface. NisP, the protease responsible for cleavage of the nisin leader, is an example of a protease located on the external face of the cell membrane (Schneewind *et al.* 1995). Evidence for this is provided through the observation that supernatant and membrane-free extracts of *L. lactis* 9800 were unable to remove the leader from the NisA precursor, whereas whole *L. lactis* 9800 cells or cell extracts from *E. coli* overexpressing NisP demonstrated protease activity (van der Meer *et al.* 1993). Thus NisT exports the post-translationally modified nisin prepropeptide, prior to cleavage of the leader peptide. Further work on this system has revealed that the arginine at position -1 and the alanine at position -4 in the NisA prepropeptide are necessary for cleavage by NisP (van der Meer *et al.* 1994). The presence of (Me)Lan bridges appear to be an additional requirement, as unmodified or dehydrated precursor peptides were not recognised by NisP *in vivo* (Kuipers *et al.* 2004). EpiP is proposed to be an extracellular protein initially translated as a pre-pro-enzyme that is subsequently cleaved to reveal the active protease. Like NisP, the arginine at position -1 of the prepropeptide is essential for cleavage of the leader, but unlike NisP, there appears to be no requirement for the presence of (Me)Lan bridges (Geissler *et al.* 1996). The proteases for Pep5, epilancin K7 and lactocin S production (PepP, ElkP and LasP

respectively) are all proposed to cleave the leader peptide whilst the prepropeptide is still in the cytoplasm. This processing step is linked to the ability of the transporter to export the lantibiotic across the membrane (Meyer *et al.* 1995).

Although all lantibiotics are ribosomally synthesised with a leader peptide, many biosynthetic gene clusters appear not to encode a protease. The best characterised example is the subtilin cluster from *B. subtilis* ATCC 6633. Initial evidence implying the role of an extracellular protease came from the observation that culture supernatant from a *Bacillus* strain unable to produce subtilin was able to cleave the leader peptide when incubated with subtilin containing (Me)Lan bridges (Stein and Entian 2002). Further analysis revealed three of the five extracellular serine proteases encoded by *B. subtilis* were able to cleave the leader peptide to release the mature lantibiotic (Corvey *et al.* 2003). The heterologous expression of numerous subtilin-nisin chimeras established that the subtilin leader fused to the nisin Z structural region was not cleaved by NisP in *L. lactis* (Kuipers *et al.* 1993b). The reciprocal experiment used *B. subtilis* ATCC 6633 to heterologously express the subtilin leader peptide fused to the N-terminus of nisin (1-11) in turn fused to the C-terminus of subtilin (12-32) to produce a bioactive nisin-subtilin chimera with the leader peptide removed (Chakicherla and Hansen 1995). This implied that proteolytic processing by LanP enzymes has more substrate specificity than the leader peptide cleavage performed by proteases encoded elsewhere in the genome.

1.3.4.3.2 Decarboxylation

The carboxy-terminus of many lantibiotics is post-translationally modified to introduce a S-[(Z)-2-aminovinyl]-D-cysteine (AviCys) modification (Figure 1.3). This modification has been observed in type A1 lantibiotics (e.g. epidermin, gallidermin, microbisporicin, clausin and mutacin 1140 in Figure 1.8), type B lantibiotics (e.g. mersacidin in Figure 1.10) and other lantipeptides (e.g. cypemycin and grisemycin in Figure 1.6). In the majority of these lantibiotics the C-terminal cysteine is decarboxylated to yield an enethiol intermediate which subsequently forms an AviCys bridge via cyclisation with a Dha or Dhb in the nascent peptide chain. The decarboxylation step is catalysed through a LanD protein, a member of the homooligomeric flavin-containing cysteine decarboxylase (HFCD) superfamily (Chatterjee *et al.* 2005). The best characterised LanD is EpiD from the epidermin biosynthetic gene cluster in *S. epidermidis*. The enzymatic activity has been reconstituted *in vitro* with the purified flavoenzyme able to decarboxylate EpiA, and *in vivo* through the coexpression of EpiA with EpiD in *E. coli* (Kupke *et al.* 1994; Kupke and Gotz 1997). Crystal structures have revealed the cofactors required for functionality. These are flavin mononucleotide (FMN) in EpiD and flavin adenine dinucleotide (FAD) in MrsD

(Blaesse *et al.* 2000; Blaesse *et al.* 2003). Available evidence so far suggests that AviCys formation is essential for lantibiotic production, the lack of this C-terminal modification may well render lantibiotics vulnerable to carboxypeptidase degradation in the external environment. Although epidermin can be synthesised without the AviCys modification, bioactivity is reduced through either a direct reduction in efficacy or as a consequence of reduced stability (Augustin *et al.* 1992). In contrast, deletion of *mibD* from *M. corallina* abolished production of all microbisporicin-related peptides (Foulston and Bibb 2010). Similarly the generation of mersacidin variants with alterations at residues adjacent to the decarboxylated cysteine residues severely reduced production, implying an essential role of cysteine decarboxylation in mersacidin biosynthesis (Appleyard *et al.* 2009).

As well as LanD enzymes, the HFCD family also encompasses (*R*)-4'-phospho-*N*-pantothenoylcysteine (PPC) decarboxylases (Kupke *et al.* 2000). CypD bears the most significant alignment to PPC decarboxylases instead of LanD enzymes, yet it catalyses the decarboxylation of cypemycin. Evidence for this is supplied by the *in vivo* deletion of *cypD* from *S. sp.* OH-4156, resulting in the production of non-decarboxylated cypemycin, and the *in vitro* reconstitution of decarboxylase activity with the purified enzyme and the CypA prepropeptide (Claesen and Bibb 2010). This suggests that CypD is functionally analogous to other LanD enzymes, and that both are derived from a common HFCD ancestor to perform the decarboxylation of C-terminal cysteine. Unlike other AviCys residues, the sequencing of the *cypA* prepropeptide revealed that it is formed between two cysteine residues; Cys₁₉ and Cys₂₂. The deletion of *cypD* abolished only the decarboxylation of Cys₂₂, and left a Dha residue at position 19. It was proposed that the decarboxylation by CypD follows the canonical mechanism, but the dethiolation of Cys₁₉ to Dha is performed by the same enzymes that catalyse the dehydration of threonine in cypemycin, CypH and/or CypL (Claesen and Bibb 2010). Interestingly, linaridins from gene clusters only encoding a homologue to the 5' end of *cypH* are devoid of cysteines, which would be consistent with a role for the C-terminal domain of CypH in the dethiolation of cysteine (Claesen and Bibb 2010).

1.3.4.3.3 Hydroxylation

Hydroxylation most commonly occurs on the beta-carbon of L-Asp (Figure 1.3). Although found in a number of mammalian proteins, this modification has so far only been observed in the cinnamycin/duramycin group of lantibiotics (Fredenhagen *et al.* 1990; Kaletta *et al.* 1991). The aspartate at position 15 of CinA is hydroxylated to form *erythro*-3-hydroxy-L-aspartic acid. The enzyme CinX encoded in the *cin* cluster is the alpha-ketoglutarate/iron(II)-dependent hydroxylase responsible for this modification (Okesli *et al.*

2011). A much rarer occurrence in ribosomally synthesised peptides is proline hydroxylation. Yet in microbisporicin, both the 3,4-dihydroxyproline and 4-hydroxyproline have been identified and it has been proposed that they are introduced through the action of the cytochrome p450, MibO (Foulston and Bibb 2010).

1.3.4.3.4 D-Ala

The ribosome only constructs peptides from L-amino acids, but D-amino acids can be subsequently introduced by post-translational modification. *Lactobacillus sake* L45 produces Lactocin S, which contains three D-Ala residues, each derived from the post-translational modification of L-Ser (Skaugen and Nes 1994). The two-component lantibiotic lacticin 3147 contains one D-Ala in the A1 peptide (Figure 1.10), and two in the A2 peptide (Ryan *et al.* 1999; Martin *et al.* 2004). Subsequent work on the lacticin 3147 gene cluster from *L. lactis* subsp. *lactis* DPC 3147 indicated that LtnJ is a dehydrogenase capable of dehydrogenating Dha residues which are in turn derived from dehydrated L-Ser (Cotter *et al.* 2005b).

1.3.4.3.5 Other bridges

1.3.4.3.5.1 Lysinoalanine (LysAla)

Lysinoalanine bridges are found naturally in body organs and more are generated in food proteins through the cooking process (Sternberg *et al.* 1975). This non-proteogenic amino acid is present in cinnamycin and the duramycins (Figure 1.3 and 1.6) (Fredenhagen *et al.* 1990; Kaletta *et al.* 1991). In cinnamycin this bridge is critical in mediating the interaction with phosphatidylethanolamine and is formed through a reaction between Lys₁₉ and Ser₆, using Dha₆ as an intermediate. Heterologous expression in *E. coli* indicated that Cinorf7 is critical for the formation of this cross-link (Okesli *et al.* 2011).

1.3.4.3.5.2 Disulphide

As disulfide bonds play an important role in the folding and stability of many secreted proteins, it is not surprising that they are present in some lantibiotics. This covalent bond is formed between the thiol groups of two cysteine residues. It is very unusual for a lantibiotic to contain a free cysteine residue, and to date these have only been found in the β -peptide of plantaricin W (Holo *et al.* 2001) and michiganin A (although the structural similarity of michiganin A to actagardine implies this cysteine may be involved in MeLan

formation with Dhb₈) (Holtsmark *et al.* 2006). Several other lantibiotics have prepropeptides which contain more cysteine residues than there are serine and threonine residues combined, implying the excess cysteines may be involved in disulphide bridge formation. Such is the case with sublancin 168 in which the tertiary structure is stabilised through two MeLan and two disulphide bridges (Paik *et al.* 1998). Other lantibiotics containing a disulphide bridge include the α -peptide of plantaricin W (Holo *et al.* 2001), bovicin HJ50 (Xiao *et al.* 2004), the α -peptide of haloduracin (McClerren *et al.* 2006) and thermophilin 1277 (Kabuki *et al.* 2009). Although the labyrinthopeptins are not technically lantibiotics, as they do not contain a Lan residue, they do contain dehydrated residues and labyrinthopeptin A2 also contains a disulphide bond (Meindl *et al.* 2010).

1.3.4.3.5.3 Labionin

As well as containing a disulphide bond, the labyrinthopeptins acquire their name due to a novel type of bridge structure, (2S,4S,8R)-labionin, a carbacyclic triamino acid residue created through the post-translational modification of two serines and a cysteine (Figure 1.3) (Meindl *et al.* 2010). The enzyme responsible for this modification was identified as LabKC, a kinase-cyclase whose activity has been reconstituted *in vitro* (Muller *et al.* 2010). In contrast to other *in vitro* synthetases of lantibiotics, LabKC used guanosine triphosphate (GTP) instead of ATP for the phosphorylation and dehydration of serine residues (Muller *et al.* 2010).

The presence of a kinase-cyclase domain classifies this enzyme into the class III modification enzymes alongside LanL (Goto *et al.* 2010) and RamC (Hudson and Nodwell 2004; Kodani *et al.* 2004) described in Section 1.3.4.1.4. Class III enzymes contain three distinct catalytic domains, an N-terminal phosphoSer/Thr lyase, a middle Ser/Thr kinase and a C-terminal LanC-like cyclase. Like RamC, the LanC-like domain of LabKC lacks the conserved zinc-binding and catalytic residues seen in LanC and LanM cyclases. It is likely this class will be subdivided further to reflect differences in RamC/LanL/LabKC enzymes that ultimately result in different nucleotide requirements and different bridge formations.

1.3.4.3.6 Single examples

For even rarer modifications, so far only one example of their presence in a lantibiotic is known. Microbisporicin contains a chlorinated tryptophan, which is proposed to be introduced by the enzymes MibHS comprising a tryptophan halogenase and flavin reductase respectively (Foulston and Bibb 2010). The C-terminal Lan of actagardine is oxidized to form a sulfoxide residue. This reaction is catalysed by the luciferase-like

monooxygenase GarO, as deletion of this gene from *A. garbadinensis* resulted in the production of deoxy variants of actagardine (Boakes *et al.* 2009). These chlorination and oxidation modifications alter the physico-chemical properties of the molecule and may also have an affect on lantibiotic activity. The *N*-succinylation of subtilin reported during late stages of cell growth does result in reduced bioactivity (Chan *et al.* 1993). Likewise the absence of a *N,N*-dimethylalanine (Me₂-Ala) in cypemycin abolishes any antimicrobial activity. Identification of the cypemycin gene cluster from *Streptomyces* sp. OH-4156 revealed that Me₂-Ala is formed by methylation of the N-terminal Ala residue of the prepropeptide (Claesen and Bibb 2010). Cypemycin also possesses two L-*allo*-isoleucine residues introduced by side-chain isomerisation of L-Ile (Claesen and Bibb 2010).

1.3.5 Planosporicin

Planosporicin represents one of the few lantibiotics to be described that is produced by an actinomycete. A number of experiments imply the mode of action of planosporicin is through inhibition of cell wall biosynthesis. The lantibiotic is active against *Staphylococcus aureus* but inactive to its L-forms due to their lack a functional cell wall (Castiglione *et al.* 2007). Inhibition tests of macromolecular syntheses in *S. aureus* showed that peptidoglycan synthesis is severely inhibited at concentrations over 1 µg/ml, whereas DNA, RNA and protein synthesis are only marginally affected (Castiglione *et al.* 2007). Treatment of *Bacillus megaterium* with planosporicin caused the accumulation of UDP-linked peptidoglycan precursors in the growing bacterial cells, implying planosporicin is a cell wall inhibitor acting at a step subsequent to the synthesis of lipid II (Castiglione *et al.* 2007). Planosporicin activity against VRE implies it targets lipid II through a moiety other than the D-Ala-D-Ala pentapeptide (Castiglione *et al.* 2007).

In general, the minimum inhibitory concentrations (MICs) for planosporicin are comparable to that of mersacidin and superior to that of actagardine (Table 1.1). In particular streptococci, anaerobic clostridia and some Gram-negative bacteria, such as *Moraxella catarrhalis*, are especially sensitive to planosporicin, portraying a wide spectrum of antibacterial activity. *M. catarrhalis* and *S. pneumoniae* are commonly recognised as the most common pathogens of the respiratory tract (Enright and McKenzie 1997). Furthermore, planosporicin is active against drug-resistant forms of staphylococci and enterococci such as MRSA and VRE.

Target pathogens	MIC ($\mu\text{g/ml}$)				
	Planosporicin	Actagardine	Mersacidin	Nisin	Microbisporicin
L100 <i>Staphylococcus aureus</i> ATCC6538P	2	32	4	0.5	≤ 0.13
L3751 <i>Staphylococcus aureus</i> L form	>128	>128	64	16	>128
L819 <i>Staphylococcus aureus</i> Smith ATCC 19636	16	32	4	2	≤ 0.13
L1400 <i>Staphylococcus aureus</i> MRSA	16	16	8	2	≤ 0.13
L613 <i>Staphylococcus aureus</i> MRSA	32	16	64	8	≤ 0.13
L3798 <i>Staphylococcus aureus</i> VISA	128	128	128	32	2
L3797 <i>Staphylococcus aureus</i> VISA met ^R	>128	>128	128	8	2
L3798 <i>Staphylococcus epidermidis</i> ATCC 12228	32	128	16	2	≤ 0.13
L1729 <i>Staphylococcus haemolyticus</i> met ^R	>128	>128	8	4	8
L49 <i>Streptococcus pyogenes</i>	0.5	2	n.d.	n.d.	≤ 0.13
L44 <i>Streptococcus pneumoniae</i>	4	32	4	0.25	≤ 0.13
L559 <i>Enterococcus faecalis</i>	16	32	32	4	1
L560 <i>Enterococcus faecalis</i> Van A	64	128	64	4	0.5
LA533 <i>Enterococcus faecalis</i> Van A	128	16	32	4	<u>1</u>
L568 <i>Enterococcus faecium</i>	64	64	64	2	<u>2</u>
L569 <i>Enterococcus faecium</i> Van A	128	128	64	2	1
LB518 <i>Enterococcus faecium</i> Van A	>128	>128	128	1	2
L884 <i>Lactobacillus garviae</i>	4	4	16	n.d.	≤ 0.13
L148 <i>Lactobacillus delbrueckii</i> ATCC 04797	16	>128	>128	>128	4
L3607 <i>Clostridium perfringens</i> ATCC 13124	≤ 0.25	4	8	≤ 0.13	≤ 0.125
L4018 <i>Clostridium difficile</i>	1	4	8	≤ 0.13	≤ 0.125
L970 <i>Haemophilus influenzae</i> ATCC 19418	>128	>128	>128	>128	32
L76 <i>Moraxella catarrhalis</i> ATCC 8176	1	32	2	1	0.25
L1613 <i>Neisseria meningitidis</i> ATCC 13090	>128	>128	>128	8	0.5
L997 <i>Neisseria gonorrhoeae</i>	>128	>128	>128	4	0.25
L47 <i>Escherichia coli</i>	>128	>128	n.d.	>128	>128
L145 <i>Candida albicans</i>	>128	>128	n.d.	>128	>128

Table 1.1 : Minimum inhibitory concentrations (in $\mu\text{g/ml}$) for planosporicin, actagardine, mersacidin, nisin and microbisporicin against a range of target pathogens.

Abbreviations refer to; American Type Culture Collection (ATCC), methicillin-resistant (met^R), methicillin-resistant *S. aureus* (MRSA), vancomycin-resistant (VanA), vancomycin-intermediate-resistant *S. aureus* (VISA). Reproduced from Castiglione *et al.* 2008.

Although nisin displays comparatively higher antimicrobial activity, its low solubility limits its use to that of a food preservative (Gowans *et al.* 1952). In contrast, planosporicin was found to be most stable at pH 2 to 6.5 and no degradation was observed when incubated at temperatures ranging from 15 to 40 °C (Castiglione *et al.* 2007). This high stability, solubility and activity at physiological pH implies that planosporicin has potential for systemic chemotherapy. This is substantiated through *in vivo* experiments where administration of planosporicin to mice infected with *S. pyogenes* indicated high efficacy

and low toxicity (Castiglione *et al.* 2007). Although planosporicin exhibits only modest activity against bacterial pathogens, (MIC ranges from 0.25 to >128 µg/ml) the promising physico-chemical properties imply that if activity can be enhanced it could represent a promising pharmaceutical.

1.4 Main Objectives

The overall aim of this project was to identify, clone and characterise the planosporicin biosynthetic gene cluster through a number of objectives:

Identification of a planosporicin-producing *Planomonospora* strain from a culture collection and confirmation of the structure of the lantibiotic.

Identification of the planosporicin biosynthetic gene cluster through genome scanning of the planosporicin producer and cloning of the entire gene cluster through the construction of a genomic cosmid library.

Heterologous expression of the gene cluster and delineation of the minimal gene set required for planosporicin biosynthesis.

Generation of deletions within the gene cluster to confirm individual gene functions inferred from bioinformatic analysis.

Determine similarities/differences to other lantibiotic gene clusters in particular with respect to the mechanism of regulation of gene expression.

Chapter 2 : Materials and Methods

2.1 Bacterial plasmids and strains

Table 2.1 : Plasmids and cosmids used and constructed in this study

Plasmid	Description	Selection markers	Source/ Reference
General plasmids			
pIJ790	λ -RED (<i>gam</i> , <i>bet</i> , <i>exo</i>), <i>cat</i> , <i>araC</i> , <i>rep101ts</i>	Chl	(Gust <i>et al.</i> 2004)
pUZ8002	<i>tra</i> , <i>neo</i> , RP4	Kan	(Paget <i>et al.</i> 1999)
pSuperCosI	pUC ori cos	Carb, Kan	Stratagene
pIJ10702	<i>attP int</i> (Φ C31) <i>oriT</i> pUCori (also known as pMJCos1)	Carb, Apra	(Yanai <i>et al.</i> 2006)
pIJ773	pBS SK+ containing cassette P1-FRT- <i>oriT-aac(3)IV</i> - FRT-P2	Apra	(Gust <i>et al.</i> 2004)
pIJ778	pBS SK+ containing cassette P1-FRT- <i>oriT-aadA</i> - FRT-P2	Carb, Spec/ Strep	(Gust <i>et al.</i> 2004)
pIJ780	pBS SK+ containing cassette P1-FRT- <i>oriT-vph</i> - FRT-P2	Carb, Vio	(Gust <i>et al.</i> 2004)
pIJ6902	<i>attP int</i> (Φ C31) <i>oriT</i> ter fd to P_{tipA}	Carb, Thio	(Huang <i>et al.</i> 2005)
pIJ10700	pBS SK+ containing cassette P1-FRT- <i>oriT-hyg</i> -FRT- P2	Hyg	(Gust <i>et al.</i> 2004)
pGEM®-T Easy	TA-cloning vector	Carb	Promega
pSET152	<i>attP int</i> (Φ C31) <i>oriT</i> pUCori	Apra	(Bierman <i>et al.</i> 1992)
pIJ10706	pSET <i>hyg</i>	Hyg	S. O'Rourke, John Innes Centre
pSET Ω	<i>attP int</i> (Φ C31) <i>oriT</i> pUCori	Spec/ Strep	J. Nodwell, McMaster University
pRT801	<i>attP int</i> (Φ BT1) <i>oriT</i> pUCori	Apra	(Gregory <i>et al.</i> 2003)

Plasmid	Description	Selection markers	Source/ Reference
pMS82	<i>attP int(ΦBT1) oriT pUCori</i>	Hyg	(Gregory <i>et al.</i> 2003)
Cosmids for heterologous expression			
pIJ12321	SuperCosI B4-1	Carb, Kan	This work
pIJ12322	SuperCosI F13-1	Carb, Kan	This work
pIJ12323	pIJ10702 B4-1	Carb, Apra	This work
pIJ12324	pIJ10702 F13-1	Carb, Apra	This work
pIJ12325	pIJ10702 4B8	Carb, Apra	M. Sosio, NAICONS
pIJ12326	pIJ10702 9A7	Carb, Apra	M. Sosio, NAICONS
pIJ12327	pIJ12323 with genes downstream of <i>pspV</i> removed	Carb, Apra	This work
pIJ12328	pIJ12329 with genes upstream of <i>pspE</i> removed	Carb, Apra	This work
pIJ12329	pIJ12329 with genes upstream of <i>orf-12</i> removed	Carb, Apra	This work
Tools			
pIJ12330	pIJ10706- <i>aac(3)IV</i>	Hyg, Apra	This work
pIJ12331	pIJ10706- <i>aadA</i>	Hyg, Spec/ Strep	This work
pIJ12332	pIJ10706- <i>vph</i>	Hyg, Vio	This work
pIJ12333	pIJ10706- <i>tsr</i>	Hyg, Thio	This work
Gene deletion constructs			
pIJ12334	pIJ12321 Δ <i>pspA::(oriT-aac(3)IV)</i>	Carb, Kan, Hyg	This work
pIJ12335	pIJ12322 Δ <i>pspA::(oriT-aac(3)IV)</i>	Carb, Kan, Apra	This work
pIJ12336	pIJ12321 Δ <i>pspA::(oriT-aadA)</i>	Carb, Kan, Spec/ Strep	This work
pIJ12337	pIJ12321 Δ <i>pspA::(oriT-hyg)</i>	Carb, Kan, Hyg	This work

Plasmid	Description	Selection markers	Source/ Reference
pIJ12338	pGEM-T Easy containing $\Delta pspA::(oriT-hyg)$ with 4 kb flanking sequence	Carb, Hyg	This work
pIJ12339	pGEM-T Easy containing $\Delta pspA::(oriT-hyg)$ with 2 kb flanking sequence	Carb, Hyg	This work
pIJ12340	pIJ12328 $\Delta pspA::(oriT-hyg)$	Carb, Kan, Hyg	This work
pIJ12521	pIJ12328 $\Delta pspX::(oriT-hyg)$	Carb, Kan, Hyg	This work
pIJ12522	pIJ12328 $\Delta pspW::(oriT-hyg)$	Carb, Kan, Hyg	This work
pIJ12523	pIJ12328 $\Delta pspYZ::(oriT-hyg)$	Carb, Kan, Hyg	This work
pIJ12524	pIJ12321 $\Delta pspTU::(oriT-hyg)$	Carb, Kan, Hyg	This work
pIJ12525	pIJ12327 $\Delta pspR::(oriT-hyg)$	Carb, Kan, Hyg	This work
pIJ12526	pIJ12327 $\Delta pspEF::(oriT-hyg)$	Carb, Kan, Hyg	This work
pIJ12527	pIJ12321 $\Delta orf-19$ to $orf-12$, $\Delta pspV::(oriT-hyg)$	Carb, Kan, Hyg	This work
pIJ12528	pGEM-T Easy containing $\Delta pspR::(oriT-hyg)$ with 2 kb flanking sequence	Carb, Hyg	This work
pIJ12529	pGEM-T Easy containing $\Delta pspEF::(oriT-hyg)$ with 2 kb flanking sequence	Carb, Hyg	This work
pIJ12530	pGEM-T Easy containing $\Delta pspV::(oriT-hyg)$ with 2 kb flanking sequence	Carb, Hyg	This work
pIJ12531	pIJ12321 $\Delta pspA::(scar)$ $\Delta bla::(oriT-hyg)$	Carb, Kan	This work

Complementation constructs

pIJ12532	pSET152- P_{pspA} - $pspA$ - $XbaI$	Apra	This work
pIJ12533	pSET152- P_{pspA} - $Spel$ - $XbaI$	Apra	This work
pIJ12534	pSET152- P_{pspA} - $Spel$ - $pspTU$ - $XbaI$	Apra	This work
pIJ12535	pSET152- P_{pspA} - $Spel$ - $pspTUV$ - $XbaI$	Apra	This work

Plasmid	Description	Selection markers	Source/ Reference
pIJ12536	pSET152-P _{pspJ} - <i>SpeI</i> - <i>XbaI</i>	Apra	This work
pIJ12537	pSET152-P _{pspJ} - <i>SpeI</i> - <i>pspYZ</i> - <i>XbaI</i>	Apra	This work
pIJ12538	pSET152-P _{pspJ} - <i>SpeI</i> - <i>pspYZQ</i> - <i>XbaI</i>	Apra	This work
pIJ12539	pSET152-P _{pspX} - <i>pspX</i> - <i>XbaI</i>	Apra	This work
pIJ12540	pSET152-P _{pspR} - <i>pspR</i> - <i>XbaI</i>	Apra	This work
pIJ12138	pIJ10706-P _{mibA} - <i>mibA</i>	Hyg	(Foulston 2010)
pIJ12362	pIJ10706-P _{mibA} - <i>mibABCD</i>	Hyg	(Foulston 2010)

Table 2.2 : *E. coli* strains used in this study

Strain	Genotype	Antibiotic resistance	Reference
DH5 α	<i>recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac</i>	None	(Sambrook and Russell 2001)
ET12567	<i>dam13::Tn9 dcm6 hsdM hsdR recF143 zjj201::Tn10 galK galT22 ara14 lacY1 xyl5 leuB6 thil tonA31 rpl136 hisG4 tsx78 mtli glnV44 F⁻</i>	Chl, Tet	(Macneil <i>et al.</i> 1992)
BW25113	Δ (<i>araD-araB</i>)567 Δ <i>lacZ4787</i> (:: <i>rrnB-4</i>) <i>lacIp-4000</i> (<i>lacI^Q</i>) λ^- <i>rpoS369</i> (Am) <i>rph-1</i> Δ (<i>rhaD-rhaB</i>)568 <i>hsdR514</i>	None	(Datsenko and Wanner 2000)
DH5 α /BT340	As DH5 α with pCP20 (<i>FLP+</i> , λ <i>ci857+</i> , λ P _R Rep ^{ts} , Ap ^R , Cm ^R)	Chl	(Cherepanov and Wackernagel 1995; Datsenko and Wanner 2000)
XL-1 Blue MR	<i>recA1, endA1, gyrA96, thi-1, hsdR17</i> (r _K ⁻ , m _K ⁺), <i>supE44, relA1, lac, [F', proAB, lacI^QZDM15::Tn10</i> (tet ^r)]	Nal	(Sambrook and Russell 2001)

Table 2.3 : *Actinomycetes* strains used in this study

Species	Strain	Source/ Reference
<i>Planomonospora alba</i>	NRRL 18924	(Mertz 1994)/ NRLL culture collection
<i>Planomonospora sphaerica</i>	NRRL 18923	(Mertz 1994)/ NRLL culture collection
<i>Streptomyces coelicolor</i>	M1146	(Gomez-Escribano and Bibb 2011)
<i>Streptomyces lividans</i>	TK24	(Hopwood <i>et al.</i> 1983)
<i>Microbispora corallina</i>	NRRL 30420	(Lee 2003)/ NRRL culture collection
<i>Nonomuraea</i> sp.	ATCC 39727	F. Marinelli, University of Insubria, Italy
<i>Micrococcus luteus</i>	ATCC 4698	Novacta Biosciences

Table 2.4 : *Actinomycetes* strains constructed in this study

Strain	Description	Source/ Reference
<i>S. lividans</i> TK24 used to construct derivatives:		
M1446	pIJ10702 in Φ C31 <i>attB</i>	(Claesen 2010)
M1286	pIJ12323 in Φ C31 <i>attB</i> (B4-1)	This work
M1287	pIJ12324 in Φ C31 <i>attB</i> (F13-1)	This work
M1290	pIJ12325 in Φ C31 <i>attB</i> 4B8)	This work
M1291	pIJ12326 in Φ C31 <i>attB</i> (9A7)	This work
<i>S. coelicolor</i> M1146 used to construct derivatives:		
M1410	pIJ10702 in Φ C31 <i>attB</i>	(Claesen and Bibb 2010)
M1288	pIJ12323 in Φ C31 <i>attB</i> (B4-1)	This work
M1289	pIJ12324 in Φ C31 <i>attB</i> (F13-1)	This work
M1292	pIJ12325 in Φ C31 <i>attB</i> 4B8)	This work
M1293	pIJ12326 in Φ C31 <i>attB</i> (9A7)	This work
<i>Nonomuraea</i> sp. ATCC 39727 used to construct derivatives:		
M1294	pIJ10702 in Φ C31 <i>attB</i>	(Foulston and Bibb 2010)

Strain	Description	Source/ Reference
M1295	pIJ12323 in Φ C31 <i>attB</i> (B4-1)	This work
M1296	pIJ12324 in Φ C31 <i>attB</i> (F13-1)	This work
M1297	pIJ12325 in Φ C31 <i>attB</i> (4B8)	This work
M1298	pIJ12326 in Φ C31 <i>attB</i> (9A7)	This work
M1299	pIJ12327 in Φ C31 <i>attB</i>	This work
M1300	pIJ12328 in Φ C31 <i>attB</i>	This work
M1301	pIJ12329 in Φ C31 <i>attB</i>	This work
<i>P. alba</i> NRLL18924 used to construct derivatives:		
M1302	Δ <i>pspA</i> ::(<i>oriT-hyg</i>)	This work
M1303	Δ <i>pspX</i> ::(<i>oriT-hyg</i>)	This work
M1304	Δ <i>pspW</i> ::(<i>oriT-hyg</i>)	This work
M1305	Δ <i>pspYZ</i> ::(<i>oriT-hyg</i>)	This work
M1306	Δ <i>pspTU</i> ::(<i>oriT-hyg</i>)	This work
M1307	Δ <i>pspV</i> ::(<i>oriT-hyg</i>)	This work
M1308	Δ <i>pspR</i> ::(<i>oriT-hyg</i>)	This work
M1309	Δ <i>pspEF</i> ::(<i>oriT-hyg</i>)	This work

2.2 Oligonucleotides used in this study

Table 2.5 : Primers used in this study

Primer	Sequence	Description
General primers		
pSETF	AAAAGTTCGACAGCGTCTCC	anneal within pSET vectors
pSETR	CTGGCGTAATAGCGAAGAGG	
pSET152F	TCGCCATTCAGGCTGC	amplify across MCS of pGEM-T
pSET152R	CTCATTAGGCACCCCAGG	
pSET152F2	GTTTTCCCAGTCACGACGTT	amplify across MCS of pSET152
pSET152R2	TGTGGAATTGTGAGCGGATA	

Primer	Sequence	Description
pRT801F	TTTGCAAGCAGCAGATTACG	anneal within pRT801
pRT801R	CCAAGGTTGAGAAGCTGACC	
pMS82forB	CTACGGGCCGACCGAGGCG	anneal within pMS82
pMS_Rseq	CTAGTAGTTCCTTCGTCACC	
seq pr. pMS82_rev	CTGATGTCATCAGCGGTGG	
pMS82_for seq pr.	TGTACGCCACCGCCTGGTTCTGCGA	
P1	ATTCCGGGGATCCGTCGACC	universal priming sites
P2	TGTAGGCTGGAGCTGCTTC	
Chapter 3: Partial sequencing of 16S rRNA		
T27	AGAGTTTGATCMTGGCTCAG	(Mazza <i>et al.</i> 2003)
R1492	TACGGYTACCTTGTTACGACTT	
Chapter 4 : 454 contig test primers and primers for <i>psp</i> probe		Anneal within:
1289FlanA	CGGTGAGTAAATCCTCGTC	<i>pspA</i>
1289RlanA	CCGTAGACGTCGAGGAAGC	
Probe3088F	CGGCGAGCGGTGCGGAAGTC	<i>pspB</i>
Probe3088R	CTCGCCGAGCTGGTGCGCCA	
3088F2	AACAGGCCGAGCTTGG	<i>pspB</i>
3088R2	CGCCACGTCCTGAACC	
PMSLanAf	CGGTGAGTAAATCCTCGTC	<i>pspA-pspB</i>
PMSLanBf	CCGTAGACGTCGAGGAAGC	
lanEF_F	GGACAAGGACTTCTGGTACGC	<i>pspE-pspF</i>
lanEF_R	GGGTTTCATCGTCAGATCTCC	
End_F	ACATTTCCCCGAAAAGTGC	cosmid end sequencing
End_R	TTGTCCGTGGAATGAACAATG	
ES1F	GAAGGTCGGCACGTACTCC	left border of <i>psp</i> cluster
ES1R	ACTTCGTCGACGGTGAGG	
ES2F	GTTCTACGTGACGGACTTCG	right border of <i>psp</i> cluster
ES2R	TTCCTTGACTTCGGATGACC	

Primer	Sequence	Description
Chapter 6 : Generation of a minimal gene set		Application:
downdisruptF	CGGTCCCGGCGCCCGCCGGGCTGAACGGA CGCGCACCGGACTAGTATTCCGGGGATCCG TCGACC	Δ orf+7- orf+14 ::aac(3)IV
downdisruptR	ATCCAGTAGGGCCGCGGTGAGCGTCTCGGC GCGGCCGCGACTAGTTGTAGGCTGGAGCTG CTTC	
downconfirm	GAAAGGCGCCCGAAGCGCTG	
updisruptF	TCTACCTGCCCGCCTTCGAACGCGTCGTCA CGCAGGCCTCTAGAATTCGGGGATCCGTC GACC	
updisruptR1	CGCCCCGGTCCTCCCGGGTCCCGGCGGGA GGGCGCCCGGTCTAGATGTAGGCTGGAGCT GCTTC	Δ orf-19-orf- 12 ::aac(3)IV
updisruptR2	GCACTGCCATCCGGCCTGCCAGAGGGTGAA TTCCGCGATTCTAGATGTAGGCTGGAGCTG CTTC	Δ orf-19-orf- 10 ::aac(3)IV
upconfirm	GCGGCGGCCGTGGTGTTCAC	
Dbv5_F	GTGACCCTGACCTTGTACC	anneal within <i>dbv5</i>
Dbv5_R	CTCACTGCTGGTTCGGTATCA	
Chapter 7 : Assessment of integration sites and resistance cassettes		
BamHlclone	ATATGGATCCGCTATAATGACCCCGAAGCA	flanking resistance cassettes
Xbalclone	ATATTCTAGAGAATAGGAACTTCGGAATAGG AAC	
vphR	GTCGTCGACTGGGACGAG	
tsrF	CGCTCTCTGGCAAAGCTC	
pSET(Nhe)F	GCTCGCTGATGATATGCTGA	flanking pSET <i>NheI</i>
pSET(Nhe)R	GTACATCACCGACGAGCAAG	
LF025F	CGCTGACATCCTGGAGAC	anneal within <i>mibA</i>
LF025R	GCACAGCGACCAGCTC	
LF026F	GAGCACGACCAGGGAAAG	anneal within <i>mibD</i>
LF026R	GAACGTGATCGGGCAGTC	
Chapter 7 : Constructing biosynthetic mutants		Application:
PspAdisruptF	CCGCGAACACCTGTTTCGATTATTCGAAGGA GTAATCATGATTCCGGGGATCCGTCCGACC	Δ pspA::hyg

Primer	Sequence	Description
PspAdisruptR	GCGGGCGCCCGGCCCCACGGGCGTGTGCG GGTGCGGTCATGTAGGCTGGAGCTGCTTC	
PspAconfirmF	CACACGCCATGTACGTATCC	
PspAconfirmR	CGCATCCTGATGTCGTAGC	
HindlanA4F	ATATA <u>AAGCTT</u> GTCATGTCGTGGCCTTCC	<i>HindIII</i> -2kb- Δ <i>pspA</i> :: <i>hyg</i> - 2kb- <i>SpeI</i>
SpelanA4R	ATATA <u>CTAGT</u> GGTGTGGGCGAGGTTAGC	
4kb5'confirm	CTCCGACCGGCTTTGAC	
4kb3'confirm	GACTTCCTGGTCTTCCTGTCC	
HindlanA8F	ATATA <u>AAGCTT</u> GCTTTCCAGAGGTAGGTGTAG C	<i>HindIII</i> -4kb- Δ <i>pspA</i> :: <i>hyg</i> - 4kb- <i>SpeI</i>
SpelanA8R	ATATA <u>CTAGT</u> CGTTGATCAGGTCGTAGGC	
8kb5'confirm	GAGATCCGTCCGTGCTGA	
8kb3'confirm	GTCTACGACGCCGTGATGT	
-5 disruptF2	GTCGGCGGCGAGCAGGGCGGCGTAGCCGT CCTCCAGCGAATTCCGGGGATCCGTCGACC	Δ <i>pspYZ</i> :: <i>hyg</i>
-4disruptR	ACCGGCCCCACCACGGCATCGGGGAGGC GTGATGACCGTGTAGGCTGGAGCTGCTTC	
-5 confirmF	AGGACGTTGAGTTCGACAGC	
-4 confirmF	ACCGTGTTCTGGAGATGG	
PspTUdisruptF	GGCCGTCGGGGACGGAGACGCCGGCGGCC TTCGAGATAAATTCGGGGATCCGTCGACC	Δ <i>pspTU</i> :: <i>hyg</i>
PspTUdisruptR	CAGCAGGGCCAGGACCGCCGCCAGCATCG CGCCAGCACTGTAGGCTGGAGCTGCTTC	
PspTUconfirmF	GCCTACGACCTGATCAACG	
PspTUconfirmR	GAAGTAGCGCAGTTCGATCC	
PspTUconfirmR 10700	TGCTTCGGGGTCATTATAGC	
PspVdisruptF	CCTGGCCCTGCTGCGGATCCGGAGCGGCC GATGAGCGCCATTCCGGGGATCCGTCGACC	Δ <i>pspV</i> :: <i>hyg</i>
PspVdisruptR	GGGAGCGACCGTGAACCGCATCCGGTGCG CGTCCGTTTATGTAGGCTGGAGCTGCTTC	
PspVconfirmF	ATGTTCCCATGCTGTTCC	
PspVconfirmR	ACTTCCAGGACCTGATCACC	

Primer	Sequence	Description
pspV_2kbF	ACTTCAACCTCGGCATGG	2kb- Δ pspV
pspV_2kbR	ACGTTACACCACGATCAAGC	::hyg-2kb
PspEdisruptF	GGGCGTGACCACGATCGGCCGAGCAAAGG AGACACGGTGATTCCGGGGATCCGTCGACC	Δ pspEF::hyg
PspFdisruptR	CGGGTACGGGCGGCCGTACGCCGTCCTCT GCGGGGGTCATGTAGGCTGGAGCTGCTT	
PspEconfirmE	TTCCTTCCTGACTGGAGTGG	
PspFconfirmF	GTGCTCAATCTCGCCATAACC	
PspEFconfirmR 10700	GGGCAGGATAGGTGAAGTAGG	
pspE_2kbF	CGGCCTGATCTATCTGAAGG	2kb- Δ pspEF
pspF_2kbR2	GCAGGAACTGTCCCTCTACC	::hyg-2kb
blaF	CCCTGATAAATGCTTCAATAATATTGAAAAAG GAAGAGTA	amplify bla- cassette-bla
blaR	AATCAATCTAAAGTATATATGAGTAAACTTGG TCTGACAG	
Chapter 7 : Complementation of biosynthetic mutants		Amplify:
PpspAF_BamHI	ATGGATCCCACACGCCATGTACGTATCC	<u>BamHI-P_{pspA}</u> - <u>pspA-XbaI</u>
PspAR_XbaI	ATTCTAGAAGTGCTTCCTGTGACGATCC	
PpspAR_SpeI BamHI	ATATGGATCCACTAGTGATTACTCCTTCGAA TAA	<u>BamHI-P_{pspA}</u> - <u>SpeI, XbaI</u>
PpspAF2_BamHI	ATATGGATCCGATCTCCTCGGCCACACC	
pspTUF_SpeI	ATATACTAGTCTGGCATGACCGGGCCGT	<u>SpeI-pspTU-</u> <u>XbaI</u>
pspTUR_XbaI	ATATTCTAGAGAGGACGCCTGTCTGCTC	
pspTUVR_XbaI	ATATTCTAGAGGCCACCTGGAGGAGTTG	<u>pspTUV-XbaI</u>
PpspJF_SpeI BamHI	ATATGGATCCACTAGTGTCTGGCCTTCCCT GGT	<u>BamHI-P_{pspJ}</u> - <u>SpeI, XbaI</u>
PpspJR_BamHI	ATATGGATCCGGCCATGATCTGCTCGAC	
pspYZF_XbaI	ATATTCTAGAGATGTTTCGTTCGGTCATTTCGT	<u>SpeI-pspYZ-</u> <u>XbaI</u>
pspYZR_SpeI	ATATACTAGTGTGATGACCGCGCTGTTG	
pspYZQF_XbaI	ATATTCTAGACCGATTCTTGATCACGTTT	<u>pspYZQ-XbaI</u>
Chapter 8 : Constructing regulatory mutants		Application:
ECFdisruptF	CAACCGCGCAGCGGCCGATCATGTTGCCAG CCGCCTTCGATTCCGGGGATCCGTCGACC	Δ pspX::hyg

Primer	Sequence	Description
ECFdisruptR	GGATACGTACATGGCGTGTGTCCGGACGAG GATTTACTCACCGCTGTAGGCTGGAGCTGC TTC	
ECFconfirmF	CCAGCAACAGGTCTGAAGG	
ECFconfirmR	CGTAAATCACGCAATCATGG	
-2disruptF	TCGGCGTACGGACCCCGCCGGCCGGGG GCTCGCGTCAATTCCGGGGATCCGTGACC	Δ pspW::hyg
-2disruptR	CGCTGCGCGGTTGGAAGGATCGGGACGAC GATGACGGAATGTAGGCTGGAGCTGCTTC	
-2confirmF	GTCATGTCGTGGCCTTCC	
-2confirmR	GTCGAGGAACTCGCCTACC	
-7disruptF	ACCCGCGAGGAGGGCGTACGGGCGGCCG GAGGCCTTCAATTCCGGGGATCCGTGACC	Δ pspR::hyg
-7disruptR	CGGGGGACCGCATTGCGCCGGCACGCTTC GCAGGCTGTCTGTAGGCTGGAGCTGCTTC	
-7confirmF	GGCTGCACAAACACTACGG	
-7confirmR	TGATGATCACCGAACTCACG	
pspR_2kbF	GTAAGCGTCGTGGTCATCG	2kb- Δ pspR ::hyg-2kb
pspR_2kbR	CGCTGTTGTGGATCTTGAGG	
Chapter 8 : Complementing regulatory mutants		Amplify:
PpspXF_BamHI	ATGGATCCGTCGTCTGGGCAGTGTTC	<u>BamHI</u> -P _{pspX} - <u>XbaI</u>
PspXR_XbaI	ATTCTAGAACGAGACGCTGGTGATGG	
pspRF_XbaI	ATATTCTAGAGAAGGTCGGCACGTACTION	<u>BamHI</u> -P _{pspR} - <u>XbaI</u>
PpspRR_BamHI	ATATGGATCCCTCAGCAGGACCATGAAGC	
Chapter 8 : RT-PCR		Anneal within:
pspE_RTF	GTGACGCTCGAGTTCTGG	<i>pspE</i>
pspE_RTR	ACCTTGCGGACGAAGTAGC	
pspF_RTF	GGCTGCACAAACACTACGG	<i>pspF</i>
pspF_RTR	CTTCATGAGGGTCGTCTTGC	
pspR_RTF	GCCAGGTCGTGACAGC	<i>pspR</i>
pspR_RTR	CACGTCCATGACCATCTCC	
pspQ_RTF	CGGAACTTCGGCTACACC	<i>pspQ</i>
pspQ_RTR	GGCCGAGACTCACCTTGC	

Primer	Sequence	Description
pspZ_RTF	GGGAGTTCCTCGACTACATCG	<i>pspZ</i>
pspZ_RTR	GACGGATCTCCCATCAGC	
pspY_RTF2	AGGTACGAGGCCCTGATGG	<i>pspY</i>
pspY_RTR	GGTGCAGAAGGTGTTGACG	
pspJ_RTF	CTGGAGCTGCACCTGTCTG	<i>pspJ</i>
pspJ_RTR	GCCATCTCCAGGAACACG	
pspW_RTF	ACTGCCGCAGGAGACTGG	<i>pspW</i>
pspW_RTR	CCAGCAACAGGTCTGAAGG	
ECF_RTF	GTCGAGGAACTCGCCTACC	<i>pspX</i>
ECF_RTR	GTCCCGATCCTTCCAACC	
pspA_RTF2	CCTGCCGCAGAACACC	<i>pspA</i>
pspA_RTR2	GGGTGCACCACGAGACG	
pspT_RTF	GTCTCGGCCTGGTCTTCC	<i>pspT</i>
pspT_RTR	GAACAGGTCGAGCATCAGG	
pspU_RTF	GTCGTCTTCAGCGTCATCG	<i>pspU</i>
pspU_RTR	GTACGTCATGGGGTTGAGG	
pspV_RTF	TCTTCCCCTACCATCTCAACC	<i>pspV</i>
pspV_RTR	GTTGACCACGGGTTTGTAGG	
hrdB_RTF	CCAAGGGCTACAAGTTCTCG	<i>hrdB</i>
hrdB_RTR	GGCCAGCTTGTTGATGACC	
q16sF	TCACGGAGAGTTTGATCCTGGCTC	<i>rrnD</i>
q16sR	GGTCGCCCTCTCAGGCCGG	
Orf-10_RTF	GGAGGACTCGCCGTACCC	<i>orf-10</i>
Orf-10_RTR	ACGTTGTTGTCGCTGATCC	
Orf+7_RTF	GCTCCAGTCTGGTGGAGGTC	<i>orf+7</i>
Orf+7_RTR	TGATCAGGTCCTGGAAGTCG	

2.3 Culture media and antibiotics

2.3.1 Antibiotics

Table 2.6 : Concentration of antibiotics used in this study

Antibiotic	Concentration in media (µg/ml)
Carbenicillin (Carb)	100
Kanamycin (Kan)	50
Chloramphenicol (Chl)	25
Apramycin (Apra)	50 (<i>E. coli</i> and <i>Streptomyces</i>) 20 (<i>Planomonospora</i>)
Hygromycin B (Hyg)	40 (in low salt media e.g. DNA) 80 (in high salt media e.g. LB)
Nalidixic acid (Nal)	25
Spectinomycin (Sp)	50 (<i>E. coli</i>), 200 (<i>Streptomyces</i> sp.)
Streptomycin (Str)	50 (<i>E. coli</i>), 10 (<i>Streptomyces</i> sp.)
Viomycin (Vio)	30
Thiostrepton (Thio)	50

2.3.2 Agar media

Unless stated otherwise, the media used for the culturing of *S. coelicolor* and *E. coli* were prepared as previously described (Kieser *et al.* 2000; Sambrook and Russell 2001).

Table 2.7 : Agar media used in this study

Medium	Composition	Instructions for preparation
L-Agar	Agar	The ingredients, except agar, were dissolved, in the distilled water and 200 ml aliquots were dispensed into 250 ml Erlenmeyer flasks containing the agar. The flasks were closed and autoclaved.
	Difco bacto tryptone	
	NaCl	
	Glucose	
	dH ₂ O	

Medium	Composition		Instructions for preparation
LB-Agar	Agar	15 g	The ingredients, except agar, were dissolved, in the distilled water and pH altered to 7.5 with NaOH 200 ml aliquots were dispensed into 250 ml Erlenmeyer flasks containing the agar. The flasks were closed and autoclaved.
	Difco bacto tryptone	10 g	
	Yeast extract	5 g	
	NaCl	10 g	
	dH ₂ O	To 1000 ml	
Difco nutrient agar (DNA)	Difco nutrient agar	4.6 g	Difco Nutrient Agar was placed in each 250 ml Erlenmeyer flask and distilled water was added. The flasks were closed and autoclaved.
	dH ₂ O	200 ml	
Soft Nutrient Agar (SNA)	Difco nutrient broth powder	8 g	The ingredients, except agar, were dissolved, in the distilled water and 200 ml aliquots were dispensed into 250 ml Erlenmeyer flasks containing the difco bacto agar. The flasks were closed and autoclaved.
	Difco bacto agar	5 g	
	dH ₂ O	To 1000 ml	
Mannitol soya flour medium (SFM or MS)	Agar	20 g	The mannitol was dissolved in the water and 200 ml aliquots poured into 250 ml Erlenmeyer flasks each containing agar and soya flour. The flasks were closed and autoclaved twice (115 °C/15 min), with gentle shaking between the two runs.
	Mannitol	20 g	
	Soya flour	20 g	
	Tap water	To 1000 ml	
Oatbran Medium (OBM)	Porridge oats	40 g	The ingredients were mixed and dispensed into 250 ml Erlenmeyer flasks. The flasks were closed and autoclaved twice (115 °C/15 min), with gentle shaking between the two runs.
	Lab M agar	20 g	
	Tap water	To 1000 ml	
Trace elements for ISP4	FeSO ₄ .7H ₂ O	0.1 g	The ingredients were dissolved, in the distilled water. 10 ml aliquots were filtered through 0.2 µm and dispensed into universals.
	MnCl ₂ .4H ₂ O	0.1 g	
	ZnSO ₄ .7H ₂ O	0.1 g	
	dH ₂ O	To 100 ml	

Medium	Composition		Instructions for preparation
ISP4	Difco bacto agar	20 g	The ingredients were combined in the distilled water. Solution 1 and solution 2 were mixed together and the pH adjusted to pH 7 - 7.4 200 ml aliquots were dispensed into 250 ml Erlenmeyer flasks. The flasks were closed and autoclaved. Once molten, yeast extract was added to a concentration of 0.1 % and mixed. 200 ml aliquots were dispensed into 250 ml Erlenmeyer flasks containing the difco bacto agar. The flasks were closed and autoclaved.
	Solution 1:		
	Soluble starch	10 g	
	dH ₂ O	500 ml	
	Solution 2:		
	CaCO ₃	2 g	
	K ₂ HPO ₄	1 g	
	MgSO ₄	1 g	
	NaCl	1 g	
	(NH ₄) ₂ SO ₄	2 g	
	Trace elements	1 ml	
	dH ₂ O	500 ml	
AF/MS	Lab M agar	10 g	The ingredients were combined in the distilled water and the pH adjusted to pH 7.3. 200 ml aliquots were dispensed into 250 ml Erlenmeyer flasks. The flasks were closed and autoclaved.
	Dextrose	20 g	
	Yeast extract	2 g	
	Soy flour	8 g	
	CaCO ₃	4 g	
	NaCl	1 g	
		dH ₂ O	
D/Seed	Lab M agar	10 g	The ingredients were combined in the distilled water and the pH adjusted to pH 7. 200 ml aliquots were dispensed into 250 ml Erlenmeyer flasks. The flasks were closed and autoclaved.
	Peptone	5 g	
	Soluble starch	20 g	
	Meat extract	2 g	
	Yeast extract	3 g	
	Soy flour	2 g	
	CaCO ₃	1 g	
		dH ₂ O	

Medium	Composition		Instructions for preparation
M8	Lab M agar	10 g	The ingredients were combined in the distilled water and the pH adjusted to pH 7. 200 ml aliquots were dispensed into 250 ml Erlenmeyer flasks. The flasks were closed and autoclaved.
	Soluble starch	20 g	
	Meat extract	2 g	
	Yeast extract	2 g	
	Glucose	10 g	
	CaCO ₃	3 g	
	Hydrolyzed casein	4 g	
	dH ₂ O	To 1000 ml	
ATCC medium 172	Agar	15 g	The ingredients were combined in the distilled water. 200 ml aliquots were dispensed into 250 ml Erlenmeyer flasks. The flasks were closed and autoclaved (121 °C/15 min).
	Soluble starch	20 g	
	Glucose	10 g	
	Yeast extract	5 g	
	N-Z amine type A	5 g	
	CaCO ₃	1 g	
	dH ₂ O	To 1000 ml	
V0.1 (Marcone <i>et al.</i> 2010c)	Agar	15 g	The ingredients were dissolved, in the distilled water and the pH adjusted to pH 7.2. 200 ml aliquots were dispensed into 250 ml Erlenmeyer flasks. The flasks were closed and autoclaved.
	Soluble starch	2.4 g	
	Dextrose	0.1 g	
	Meat extract	0.3 g	
	Yeast extract	0.5 g	
	Triptose	0.5 g	
	dH ₂ O	To 1000 ml	

2.3.3 Liquid media

Unless stated otherwise, the media used for the culturing of *S. coelicolor* and *E. coli* were prepared as previously described (Kieser *et al.* 2000; Sambrook and Russell 2001).

Table 2.8 : Liquid media used in this study

Medium	Composition		Instructions for preparation
L (Lennox)-Broth	Difco bacto tryptone	10 g	The ingredients were dissolved, in the distilled water and aliquots were dispensed into universals or 250 ml flasks and autoclaved.
	Difco yeast extract	5 g	
	NaCl	5 g	
	Glucose	1 g	
	dH ₂ O	To 1000 ml	
LB (Luria-Bertani)-broth	Difco bacto tryptone	10 g	The ingredients were dissolved, in the distilled water and pH adjusted to 7. Aliquots were dispensed into universals or 250 ml flasks and autoclaved.
	Difco yeast extract	5 g	
	NaCl	10 g	
	dH ₂ O	To 1000 ml	
SOB (minus Mg)	Tryptone	20 g	After dissolving the solutes in water, 10 ml 250 mM KCl was added and the pH was adjusted to pH 7 with 5 N NaOH. The volume was then made up to 1000 ml with deionised water and autoclaved.
	Yeast extract	5 g	
	NaCl	0.5 g	
	dH ₂ O	To 950 ml	
SOC			SOC medium is identical to SOB medium except that after autoclaving, 20 ml of sterile 1 M solution of glucose and 5 ml of sterile 2 M MgCl ₂ were added.
2 X YT medium	Difco bacto tryptone	16 g	The ingredients were dissolved, in the distilled water and 10 ml aliquots were dispensed into universals and autoclaved.
	Difco yeast extract	10 g	
	NaCl	5 g	
	dH ₂ O	To 1000 ml	

Medium	Composition		Instructions for preparation
Tryptone soya broth (TSB)	Tryptone soya broth powder	30 g	Tryptone soya broth powder was placed in each 250 ml Erlenmeyer flask and distilled water was added. The flasks were closed and autoclaved.
	dH ₂ O	To 1000 ml	
Super yeast extract-malt extract (SuperYEME)	Yeast Extract	3 g	The ingredients were combined in the distilled water. 200 ml aliquots were dispensed into 250 ml Erlenmeyer flasks. The flasks were closed and autoclaved.
	Peptone	5 g	
	Malt Extract	3 g	
	Glucose	10 g	
	Sucrose	340 g	
	MgCl ₂ 6H ₂ O	1.1 g	
	Glycine	5 g	
	L-Proline	0.075 g	
	L-Arginine	0.075 g	
	L-Histidine	0.1 g	
	Uracil	0.015 g	
	L-Cystine	0.075 g	
	dH ₂ O	To 1000 ml	
Trace elements for ISP4	FeSO ₄ .7H ₂ O	0.1 g	The ingredients were dissolved, in the distilled water. 10 ml aliquots were filtered through 0.2 µm and dispensed into universals.
	MnCl ₂ .4H ₂ O	0.1 g	
	ZnSO ₄ .7H ₂ O	0.1 g	
	dH ₂ O	To 100 ml	

Medium	Composition	Instructions for preparation	
ISP4	Solution 1:	The ingredients were combined in the distilled water. Solution 1 and solution 2 were mixed together and the pH adjusted to pH 7 to 7.4. 200 ml aliquots were dispensed into 250 ml Erlenmeyer flasks. The flasks were closed and autoclaved. Prior to use, yeast extract was added to a concentration of 0.1 % and mixed.	
	Soluble starch		10 g
	dH ₂ O		500 ml
	Solution 2:		
	CaCO ₃		2 g
	K ₂ HPO ₄		1 g
	MgSO ₄		1 g
	NaCl		1 g
	(NH ₄) ₂ SO ₄		2 g
	Trace element solution		1 ml
	500 ml		
	dH ₂ O		
AF/MS	Dextrose	20 g	The ingredients were combined in the distilled water and the pH adjusted to pH 7.3. 200 ml aliquots were dispensed into 250 ml Erlenmeyer flasks. The flasks were closed and autoclaved.
	Yeast extract	2 g	
	Soy flour	8 g	
	CaCO ₃	4 g	
	NaCl	1 g	
	dH ₂ O	To 1000 ml	
D/seed	Peptone	5 g	The ingredients were combined in the distilled water and the pH adjusted to pH 7. 200 ml aliquots were dispensed into 250 ml Erlenmeyer flasks. The flasks were closed and autoclaved.
	Soluble starch	20 g	
	Meat extract	2 g	
	Yeast extract	3 g	
	Soy flour	2 g	
	CaCO ₃	1 g	
	dH ₂ O	To 1000 ml	

Medium	Composition		Instructions for preparation
M8	Soluble starch	20 g	The ingredients were combined in the distilled water and the pH adjusted to pH 7. 200 ml aliquots were dispensed into 250 ml Erlenmeyer flasks. The flasks were closed and autoclaved.
	Meat extract	2 g	
	Yeast extract	2 g	
	Glucose	10 g	
	CaCO ₃	3 g	
	Hydrolyzed casein	4 g	
	dH ₂ O	To 1000 ml	
V (Marcone <i>et al.</i> 2010c)	Soluble starch	24 g	The ingredients were dissolved in the distilled water and the pH adjusted to pH 7.2. 200 ml aliquots were dispensed into 250 ml Erlenmeyer flasks. The flasks were closed and autoclaved.
	Dextrose	1 g	
	Meat extract	3 g	
	Yeast extract	5 g	
	Tryptose	5 g	
	dH ₂ O	To 1000 ml	
VSP (Marcone <i>et al.</i> 2010c)			VSP is identical to V except for the addition of sucrose to a final concentration of 50 g/L and of L-proline to 0.5 g/L after autoclaving.
Medium 266 (YS) (Technikova-Dobrova <i>et al.</i> 2004)	Soluble start	10 g	The ingredients were dissolved in the distilled water and the pH adjusted to pH 7.2. 200 ml aliquots were dispensed into 250 ml Erlenmeyer flasks. The flasks were closed and autoclaved.
	Yeast extract	2 g	
	dH ₂ O	To 1000 ml	

Medium	Composition		Instructions for preparation
Streptosporangium Medium (SM) (Roes and Meyers 2008)	Soluble starch	10 g	The ingredients were dissolved in the distilled water and the pH adjusted to pH 7.3. 200 ml aliquots were dispensed into 250 ml Erlenmeyer flasks. The flasks were closed and autoclaved.
	Glycerol	10 g	
	Tryptone	2.5 g	
	Bacteriological peptone	5 g	
	Yeast extract	2 g	
	NaCl	1 g	
	CaCO ₃	3 g	
	Tap Water	To 1000 ml	

2.4 Solutions and buffers

Table 2.9 : Solutions and buffers used in this study

Solution/buffer	Composition and instructions for preparation	
Cosmid isolation solution I	Glucose	50 mM
	Tris-HCl (pH 8)	25 mM
	EDTA	10 mM
Cosmid isolation solution II	NaOH	0.2 M
	SDS	1 %
Cosmid isolation solution III	Potassium acetate (pH 5.5)	3 M
	Acetic acid	
Cosmid Library Probing Hybridisation Buffer	SSC	6x
	Denhardt's solution	1x
	SDS	0.5 %
	Denatured calf thymus DNA	50 µg/ml
Cosmid Library Probing Wash Buffer	SSC	0.1x
	SDS	1 %

Solution/buffer	Composition and instructions for preparation	
DNA loading buffer	Bromophenol blue	0.125 % (w/v)
	Xylene- cyanol blue	0.125 % (w/v)
	Glycerol	62.5 % (v/v)
	SDS	0.625 % (w/v)
Elution Buffer	Tris-Cl	10 mM
		pH 8.5
SET buffer	NaCl	75 mM
	Tris-HCl (pH 8)	20 mM
	EDTA (pH 8)	25 mM
GET buffer	Glucose	50 mM
	Tris-HCl (pH 8)	25 mM
	EDTA (pH 8)	10 mM
TE buffer	Tris-HCl (pH 8)	10 mM
	EDTA (pH 8)	1 mM
20x SSC	Sodium citrate	300 mM
	NaCl	3 M
TAE buffer	Tris	40 mM
	Acetic acid	1.142 %
	EDTA	1 mM
TBE buffer	Tris base	89 mM
	Boric acid	89 mM
	EDTA (pH 8.0)	2 mM

2.5 General molecular biology methods

2.5.1 Plasmid isolation

QIAprep[®] miniprep kits (Qiagen) were used according to the manufacturer's instructions. Briefly, a 10 ml LB culture of *E. coli* containing the plasmid of interest was grown overnight then centrifuged at 3000 rpm for 10 min. The cell pellet was lysed under alkaline conditions to denature high molecular weight chromosomal DNA. The lysate was then neutralised and centrifuged in a microcentrifuge at 13000 rpm to remove renatured chromosomal DNA, precipitated protein and cell debris. The supernatant was then applied

to a silica-gel membrane mounted in a microcentrifuge tube where the plasmid DNA is selectively adsorbed. The membrane was washed under high salt and ethanolic buffer conditions to remove RNA, protein and metabolites. DNA was eluted from the column in elution buffer. Plasmid DNA was routinely stored at -20 °C.

2.5.2 Cosmid isolation

Cosmid isolation was carried out by alkaline lysis as described by (Sambrook and Russell 2001). A 9 ml overnight culture of *E. coli* was centrifuged to pellet the cells which were then resuspended by vortexing in 200 µl solution I. 400 µl solution II were added and the tubes inverted twelve times. 300 µl of solution III were then added and mixed by inverting the tubes six times. The tube was then centrifuged at 13000 rpm in a microcentrifuge for 5 min at room temperature. The supernatant was mixed with 400 µl phenol/chloroform, vortexed for two min and then centrifuged at 13000 rpm in a micro centrifuge for 5 min. 750 µl of the upper phase were then transferred to a 1.5 ml tube, 750 µl of ice cold isopropanol were added and the tube placed on ice for at least 10 min. The precipitated DNA was then centrifuged at 13000 rpm in a micro centrifuge for 5 min. The pellet was washed with 500 µl 70 % ethanol and centrifuged at 13000 rpm in a microcentrifuge. The pellet was dried by leaving the tube open for 10 min at room temperature prior to resuspension in 100 µl elution buffer.

2.5.3 Agarose gel electrophoresis

Typically, a 1 % agarose gel was cast with TBE buffer with 0.5 µg/ml ethidium bromide. Size markers were provided by Hyperladder I (Bioline), 1 kb ladder (NEB) or 100 bp ladder (NEB) as indicated in figure legends. 10 µl samples in 1x loading dye were loaded into the appropriate wells. Agarose gel electrophoresis was performed in 1 % TBE buffer at 100 V until completion.

2.5.4 Pulsed-field gel electrophoresis

Pulsed-field gel electrophoresis (PFGE) was carried out using CHEF DR-III pulse field gel apparatus (Bio-Rad) in combination with a cooling module (Bio-Rad) to maintain the buffer at a constant temperature. 0.5 % TBE buffer was circulated through the electrophoresis tank (at 0.75 L/ min) and pre-cooled to 14 °C. A 1 % TBE gel was cast in the standard casting frame supplied (14 x 13 cm). 5-10 µl of MidRange PFG Marker I (NEB) was

introduced into one well of the cast gel, then sealed with molten agarose. The gel was immersed in the buffer in the electrophoresis tank and the samples loaded into the appropriate wells. Electrophoresis was carried out for 18 hours at 14 °C, 6 V/cm with an initial switch time of 1 second and a final switch time of 25 seconds.

2.5.5 DNA extraction from an agarose gel

DNA fragments separated in agarose gels were excised from the gel using a clean scalpel and purified using the QIAquick[®] gel extraction kit (Qiagen). Briefly, the agarose gel slice containing the DNA fragment of interest was dissolved in a neutral pH, high salt buffer provided with the kit and applied to a silica-gel membrane mounted in a microcentrifuge tube. DNA is adsorbed while unwanted primers and impurities were washed away, leaving the DNA fragment to be eluted in 30 µl elution buffer.

2.5.6 DNA digestion with restriction enzymes

Restriction enzyme digestion of cosmids, plasmids or genomic DNA was carried out according to the enzyme manufacturer's instructions. In the case of double digests, an appropriate buffer was selected by consulting the manufacturer's literature (Roche or NEB). The reaction volume was usually 15 µl for diagnostic digests and 50-100 µl for preparative digests. Unless otherwise instructed, digests were typically carried out for 1.5 hours at 37 °C.

2.5.7 Ligation

Ligation of DNA fragments was carried out using T4 DNA ligase and buffer (Invitrogen) following manufacturer's instructions. Typically 50 ng of restriction-digested, gel-purified vector were used in a 1:3 molar ratio with the restriction-digested, gel-purified insert fragment. For blunt-end ligations, vector ends were typically treated with Calf Intestinal Alkaline Phosphatase (NEB) at 37 °C for 30 min. Ligations were typically carried out overnight at 4 °C. 3 µl of the ligation reaction were used to transform chemically competent *E. coli* DH5α.

2.5.8 Preparation and transformation of electrocompetent *E. coli*

The desired *E. coli* strain was grown overnight at 37 or 30 °C in 10 ml LB broth containing the appropriate selection. 50 ml of an appropriate growth medium (typically L broth or SOB) were inoculated 1 in 100 with the overnight culture and grown at 37 or 30 °C with shaking at 250 rpm to an OD₆₀₀ of ~0.4. The cells were recovered by centrifugation at 4000 rpm for 5 min at 4 °C in a Sorvall GS3 rotor. The medium was decanted and the pellet resuspended by gentle mixing in 50 ml ice-cold 10 % glycerol. Centrifugation was repeated as above and the pellet resuspended in 25 ml ice-cold 10 % glycerol, and the process repeated once more with 10 ml ice-cold 10 % glycerol. Finally centrifugation was repeated as above and the pellet resuspended in ~1 ml ice-cold 10 % glycerol. 100 µl of this cell suspension were typically mixed with ~100 ng of cosmid DNA. Electroporation was carried out in a 0.2 cm ice-cold electroporation cuvette using a BioRad GenePulser II set to: 200 Ω, 25 µF and 2.5 kV. The time constant was typically 4.5–4.9 ms. 700 µl ice cold LB medium was immediately added to the shocked cells. The cells were incubated with shaking at 250 rpm for 1 hour at 37 or 30 °C. Typically 200 µl and 600 µl of the transformation were spread onto LB agar containing the appropriate selection (or DNA agar in the case of hygromycin selection). Plates were incubated overnight at 37 or 30 °C.

2.5.9 Transformation of chemically-competent *E. coli*

Chemically-competent *E. coli* DH5α (Invitrogen) were transformed with DNA according to the manufacturer's instructions. Briefly, a 50 µl aliquot of cells was defrosted on ice for 5 min. Approximately 100 ng of DNA were added to the cells which were incubated on ice for 30 min before a 20-40 second heat shock at 42 °C, then rested on ice for 2 min. 900 µl of pre-warmed SOC were added and cells were incubated with shaking at 250 rpm for 1 hour at 37 °C. Typically 200 and 700 µl of the transformation were spread on LB or DNA containing selection.

2.6 General PCR methods and Sanger sequencing

A BioRad DNA Engine Thermal Cycler was typically used to perform Polymerase chain reaction (PCR). PCR tubes were strips of 8 thin-walled 0.2 ml tubes with domed-caps (Thermo Scientific). Most primers were designed with an annealing temperature of 55 to 60 °C, consequently the annealing temperature used for most applications was 56 °C. The conditions typically used for different applications are listed below. The use of 4.8 % DMSO in PCR reactions facilitates the amplification of templates with a high GC content.

2.6.1 General analytical PCR

	Amount/Concentration	Volume (μ l)	Final Concentration
Template DNA	1-10 ng	1	
Forward Primer	10 μ M	1	0.4 μ M
Reverse Primer	10 μ M	1	0.4 μ M
dNTPs	20 mM (5 mM each)	1	400 μ M (100 μ M each)
Taq Buffer (Roche)	10 x	5	1x
DMSO	40 %	6	4.8 %
Taq (Roche)	5 U/ μ l	0.5	2.5 U
dH ₂ O		34.5	
Total		50	

95 °C 2 min
 95 °C 45 s
 55-60 °C 30 s
 72 °C 1-3 min
 72 °C 7 min

} x 27 cycles

2.6.2 High-fidelity amplification for cloning applications

	Amount/ Concentration	Volume (μ l)	Final Concentration
Template DNA	100 ng/ μ l	1	
Forward Primer	10 μ M	5	1 μ M
Reverse Primer	10 μ M	5	1 μ M
dNTPs	20 mM (5 mM each)	1	400 μ M (50 μ M each)
Expand HiFi Buffer 2 (Roche)	10 x	2.5	1 x
DMSO	40 %	6	4.8 %
Expand HiFi Taq (Roche)	3.5 U/ μ l	0.71	2.48 U
dH ₂ O		28.79	
Total		50	

94 °C 2 min
 94 °C 45 s
 55-60 °C 45 s
 72 °C 1-3 min
 72 °C 7 min

} x 27 cycles

2.6.3 Colony PCR in *E. coli*

The PCR reaction mix was as 2.6.1 but with an extra 1 µl dH₂O. Colonies of *E. coli* were picked from agar plates using a sterile Gilson pipette tip (Starlabs) and were smeared onto the wall of a PCR tube before adding the appropriate PCR mix. The tubes were placed in the Thermo cycler using the conditions:

95 °C 5 min
 95 °C 45 s
 55-60 °C 45 s
 72 °C 1-3 min
 72 °C 5 min

} x 30 cycles

2.6.4 Colony PCR in *Planomonospora*

Single colonies of *P. alba* were patched on to plates of D/seed and incubated for three days at 30 °C. 50 µl 100 % DMSO were aliquoted into 1.5 ml microcentrifuge tubes for *Planomonospora* mycelium to be scraped into using a sterile toothpick. The tube was shaken vigorously for at least one hour and then briefly centrifuged to pellet cell debris. 2.5 µl of the supernatant were used as the template in the reaction as shown below. Control DNA samples were diluted 1 µl into 1.5 µl DMSO and were used in the same way.

	Amount/Concentration	Volume (µl)	Final Concentration
Forward Primer	10 µM	1	0.4 µM
Reverse Primer	10 µM	1	0.4 µM
dNTPs	20 mM (5 mM each)	1	400 µM (100 µM each)
Taq Buffer (Roche)	10 x	5	1 x
DMSO + template	100 %	2.5	5 %
Taq (Roche)	5 U/µl	0.5	2.5 U
dH ₂ O		39	
Total		50	

95 °C 5 min
 95 °C 45 s
 56-60 °C 45 s
 72 °C 1 min
 72 °C 5 min

} x 30 cycles

2.6.5 Purification of PCR products

The QIAquick™ PCR purification kit (Qiagen) was used according to the manufacturer's instructions to remove unincorporated primers, dNTPs and enzymes from PCR products >100 bp. 5 µl of the PCR reaction mixture were used in agarose gel electrophoresis to check the correct product was amplified. The remaining PCR mixture was diluted 5 times in high salt buffer and applied to a silica-gel membrane mounted in a microcentrifuge tube. The DNA fragment was washed then typically eluted in 2 x 25 µl elution buffer heated to 50 °C.

2.6.7 Sanger sequencing using Big Dye v3.1

The ABI BigDye® 3.1 dye-terminator reaction mix (Applied Biosystems) was used to label purified PCR products or vectors. The manufacturer's instructions were consulted to find what amount of DNA was appropriate as a template, based on its length. Big Dye labeling reaction was performed as detailed below and subsequently submitted to The Genome Analysis Centre for ABI Sanger Sequencing. The resulting sequence chromatogram files were analysed using VectorNTI ContigExpress software.

	Amount	Volume (µl)	Final Concentration
Template	varies	varies	varies
Primer	10 µM	3.2	0.32 µM
ABI BigDye® 3.1 dye-terminator reaction mix		1	
BigDye® 3.1 Reaction Buffer	5 x	2	1 x
dH ₂ O		varies	
Total		10	

95 °C 10 min
 96 °C 10 s
 50 °C 5 s
 60 °C 4 min

} x 25 cycles

2.7 Growth conditions

2.7.1 Growth and storage of *E. coli*

E. coli strains were typically grown with the appropriate antibiotic selection either on L agar at 30 or 37 °C or in L broth with shaking at 250 rpm. For long-term storage strains were grown overnight in L broth and stored in a final concentration of 20 % glycerol at -20 °C or -80 °C.

2.7.2 Growth and storage of *Micrococcus luteus*

M. luteus ATCC 4698 was typically grown either on L agar at 30 or 37 °C or in L broth with shaking at 250 rpm. For long-term storage strains were grown overnight in L broth and stored in a final concentration of 20 % glycerol at -20 °C or at -80 °C.

2.7.3 Growth and storage of *Streptomyces*

Streptomyces was grown and manipulated as described in (Kieser *et al.* 2000). Growth media were as stated in Section 2.3.2 and 2.3.3. For manipulation of *Streptomyces* and growth of confluent lawns for long-term storage, SFM agar medium was used with the appropriate antibiotic selection at 30 °C. Spores were collected in approximately 2 ml 20 % glycerol using a sterile cotton pad through which spores were filtered by collecting with a 2 ml syringe and transferring to a 2 ml cryotube. Spore stocks were spun and 1 ml supernatant was removed before resuspension and storage at -20 °C. Spore stocks were titred by making serial dilutions in water and plating on SFM. The resulting colonies were counted from at least three dilutions and averaged. For liquid culture, $\sim 10^8$ *Streptomyces* spores were inoculated into the appropriate media in a 250 ml flask containing a baffle and incubated at 240 rpm at 30 °C.

2.7.4 Growth and storage of *Planomonospora*

P. alba NRLL18924 and *P. sphaerica* NRLL18923 were obtained from the NRRL culture collection. Growth media were as stated in Section 2.3.2 and 2.3.3. For manipulation of *Planomonospora* and growth of confluent lawns for long-term storage, ISP4 agar medium was used with the appropriate antibiotic selection at 30 °C. Spores were resuspended in 10 ml H₂O then filtered through non-absorbant cotton-wool in a glass collection tube as described in (Kieser *et al.* 2000). Spores were spun down and resuspended in 1 ml 20 %

glycerol then transferred to a 2 ml cryotube for storage at -20 °C. Spore stocks were titred by making serial dilutions in water and plating on ISP4. The resulting colonies were counted from at least three dilutions and averaged. For liquid culture, $\sim 10^7$ *Planomonospora* spores were inoculated into ISP4 medium in a 25-100 ml flask and incubated with shaking at 240 rpm at 30 °C. The use of 3-6 glass beads (2 mm diameter) instead of a baffle, prevents foaming and associated anaerobic growth.

2.7.5 Growth and storage of *Nonomuraea*

Nonomuraea ATCC 39727 was a gift from Prof Flavia Marinelli (University of Insubria, Italy). *Nonomuraea* ATCC 39727 was grown on V0.1 agar medium at 30 °C. For liquid culture, *Nonomuraea* ATCC 39727 was typically inoculated 1 in 10 into VSP medium in a 25-100 ml flask containing 3-6 glass beads (2 mm diameter) and incubated with shaking at 240 rpm at 30 °C. For long-term storage (master cell bank) culture broth containing mycelium was aliquoted into 2 ml cryotubes and stored at -80 °C.

2.7.6 *P. alba* growth curve

Cultures to assess the growth rate of *P. alba* were set up in the following way. Approximately 10^7 *P. alba* spores or 1 ml of broth containing mycelium were used to inoculate 10ml ISP4 in a 25 ml flask containing three glass beads. The culture was grown for 24-48 hours at 30 °C with shaking at 240 rpm. 100 µl of the resulting mycelium was diluted in 1 ml ISP4 and the optical density at OD₄₅₀ assessed. Three 100 ml flasks containing 25 ml ISP4 and 10 glass beads were inoculated to a starting OD₄₅₀ of 0.1-0.2. The flasks were incubated at 30 °C with shaking at 240 rpm. At appropriate time intervals samples were removed. The growth rate of *P. alba* was assessed in liquid ISP4 by two methods. First, optical density at OD₄₅₀ (using a 1 in 10 dilution when OD₄₅₀ > 1) with ISP4 as a blank. Alternatively a total protein assay was used. 1 ml of culture was sampled at the appropriate time intervals. The DC protein assay (Bio-Rad) was used according to the manufacturer's instructions. Bovine serum albumin (BSA) standards of 0.1, 0.2, 0.4, 0.6, 0.8, 1.0, 1.2 and 1.4 ng/ml were prepared in cuvettes in 0.5 ml volumes. The standard curve generated from the standards was used to calculate the protein concentrations in the experimental samples.

2.8 Microscopy

2.8.1 Phase-contrast microscopy

Samples were placed on 76x26 mm glass slides (VWR international) and covered with a glass cover slip (18x18 mm; VWR international). Slides were typically observed at 400 times magnification with a Photomicroscope II in phase-contrast mode (Zeiss).

2.8.2 Cryo-scanning electron microscopy

High-resolution SEM was carried out by Kim Findlay at the John Innes Centre using a Zeiss Supra 55 VP FEG SEM with a Gatan Alto 2500 cryo system at an accelerating voltage of 3 kV.

2.9 Next-generation sequencing and analysis

Genomic DNA isolated from *P. alba* was sequenced using 454 sequencing technology (The Genome Analysis Centre, Norwich Research Park). One quarter of a run of 454 sequencing yielded 72 Mb of sequence data. Sequence reads were assembled using Newbler assembly software. Summary statistics for the data assembly were:

Total number of contigs : 1017
Number of contigs at least 500 nt long: 944

Minimum length of all contigs : 110
Maximum length of all contigs : 141100
Mean length of all contigs : 7186
Median length of all contigs : 3186

Minimum length of contigs at least 500 nt long : 506
Maximum length of contigs at least 500 nt long : 141100
Mean length of contigs at least 500 nt long : 7718
Median length of contigs at least 500 nt long : 3592

Minimum GC fraction in all contigs : 0.569
Maximum GC fraction in all contigs : 0.901
Mean GC fraction in all contigs : 0.739
Median GC fraction in all contigs : 0.742

Minimum GC fraction in contigs at least 500 nt long : 0.569
Maximum GC fraction in contigs at least 500 nt long : 0.833
Mean GC fraction in contigs at least 500 nt long : 0.739
Median GC fraction in contigs at least 500 nt long : 0.742

Total number of nucleotides in all contigs : 7308565
Total number of nucleotides in contigs at least 500 nt long : 7285878

The assembled contig sequences were used to construct a database searchable by BLAST by Dr Govind Chandra using the format from the BLAST suite of programs version 2.2.18 (National Center for Biotechnology Information; (Altschul *et al.* 1990)).

2.10 Isolation of genomic DNA

2.10.1 *P. alba*

2.10.1.1 High-molecular weight genomic DNA

A seed culture was inoculated with *P. alba* spores or mycelium and grown for seven days at 30 °C 250 rpm. 15 ml ISP4 containing the appropriate selection was inoculated 1 in 10 with the seed culture. After three days at 30 °C 250 rpm, 10 ml were harvested by centrifugation and washed twice with 10 ml 10.3 % sucrose to remove media traces. The pellet was resuspended in 5 ml SET buffer and the pH adjusted to pH 8 with 1 N NaOH. 5 mg of lysozyme (Sigma) and 0.25 mg ribonuclease A (Sigma) was added, mixed by inversion, then incubated at 37 °C for 30 min. 3 mg proteinase K (Roche) was added and the incubation continued at 37 °C a further 15 min. Next, SDS was added to a final concentration of 1 %, along with 2 ml of 5 M NaCl and mixed by inversion. The DNA was extracted with 1 ml phenol-chloroform-isoamylalcohol pH 8 using a slow rotary mixer. The phenol and aqueous phases were separated by centrifugation at 4000 rpm for 10 min. The aqueous phase was removed to a fresh 15 ml tube and the DNA extraction repeated with 1 ml chloroform. This was mixed slowly by a rotary mixer, then spun at 4000 rpm for 10 min. The DNA was precipitated from the aqueous phase with ice-cold isopropanol. DNA was spooled from tubes using a sealed glass Pasteur pipette and washed in 70 % ethanol. The DNA was air-dried and resuspended in approximately 200 µl elution buffer at 50 °C. gDNA was stored at 4 °C.

2.10.1.2 Small-scale genomic DNA extraction

P. alba mycelium was grown for 48-72 hours in liquid culture. 10 ml were harvested by centrifugation and washed with 10 ml 10.3 % sucrose to remove media traces. The pellet was resuspended in 2.5 ml GET buffer and the pH adjusted to pH 8 with 1 N NaOH. 1.5 mg of lysozyme and 100 µg RNase A were added, mixed by inversion, then incubated at 37 °C for 30 min. 3 mg proteinase K were added and the incubation continued at 37 °C a further 15 min. Next, SDS was added to a final concentration of 1 %. The DNA was extracted with 1 ml phenol-chloroform-isoamylalcohol pH8 with vortexing for 30 seconds to mix thoroughly, then centrifugation at 6000 rpm for 15 min at 4 °C. The aqueous phase

was removed to a fresh 15 ml tube, and if cloudy, DNA extraction repeated with 1 ml chloroform. This was vortexed for 30 seconds, then centrifuged at 6000 rpm for 15 min at 4 °C. The DNA was precipitated from the aqueous phase with ice-cold isopropanol. DNA was spun down at 7000 rpm for 15 min at 4 °C and the pellet washed in 70 % ethanol. The DNA was air-dried and resuspended in elution buffer. gDNA was stored at 4°C (concentrated samples) and at a concentration of 10 ng/µl at -20 °C.

2.10.2 *Nonomuraea*

Nonomuraea mycelium was harvested from 10 ml of culture by centrifugation after 48 hours growth in liquid culture and washed with 10 ml 10.3 % sucrose to remove media traces. The pellet was resuspended in 5 ml SET buffer. The pH of the solution was adjusted to pH 8 with 1 N NaOH. 50 mg lysozyme were added and the tube incubated at 37 °C for 24 hours, adding fresh lysozyme at intervals (approximately 50 mg every 8-12 hours). After 16 hours incubation, 25 µg RNase A were added and the incubation continued at 37 °C. After 20 hours, 3 mg proteinase K were added and the incubation continued at 37 °C. After 24 hours incubation, SDS was added to a final concentration of 1 % and the solution mixed by inversion followed by an overnight incubation at 55 °C. Finally, 2 ml 5 M NaCl were added and the DNA was extracted with 4 ml phenol-chloroform-isoamylalcohol (pH 8) with vortexing to mix thoroughly. After centrifugation at 4000 rpm for 10 min, the aqueous phase was removed to a fresh 15 ml tube and 4 ml chloroform added. Again the tube was vortexed and centrifuged as above. The aqueous phase was moved to a fresh tube and the DNA was precipitated with ice-cold isopropanol. DNA was spun down at 7000 rpm for 15 min at 4 °C and the pellet washed in 70 % ethanol. The DNA was dried and resuspended in elution buffer. gDNA was stored at 4 °C.

2.10.3 *Streptomyces*

10 ml of 50:50 TSB:superYEME liquid medium in a baffled Universal was inoculated with $\sim 10^8$ *Streptomyces* spores. The culture was incubated at 30 °C for 24-36 hours. The mycelium was harvested and resuspended in 2.5 ml of SET buffer then 1.5 mg lysozyme and 0.25 mg RNase A were added. The mycelium was incubated at 37 °C for 30 min. 0.6 mg of proteinase K (20 mg/ml) was added and incubation continued for a further 15 min. SDS was added to a final concentration of 1 % and mixed by inversion before incubating at 55 °C for 15 min. The DNA was extracted with 1 ml phenol-chloroform-isoamylalcohol pH 8 with vortexing to mix thoroughly and centrifugation at 6000 rpm for 15 min at 4 °C. The aqueous phase was removed to a fresh tube and the DNA was precipitated by adding

1/10 volume 0.3 M sodium acetate (pH 6) and 1 volume ice-cold isopropanol. DNA was pelleted by centrifugation for 15 min at 7000 rpm and washed in 70 % ethanol. The DNA was air-dried and resuspended in 100 µl elution buffer at 50 °C. gDNA was subsequently stored at 4 °C.

2.11 Cosmid Library Preparation

2.11.1 Preparation of insert gDNA

High molecular weight *P. alba* gDNA was partially digested with *Sau3AI* (Roche). 50 µg gDNA were digested with 0.025 units of *Sau3AI* for 30 seconds in a volume of 250 µl. The reaction was immediately quenched with EDTA. The digested DNA was extracted by adding 500 µl phenol-chloroform-isoamylalcohol pH 8 and gently mixing on a rotary mixer. The phenol and aqueous phases were separated by centrifugation at 13000 rpm for 4 min at 4 °C. The aqueous phase was moved to a fresh tube using a wide-bore tip. The DNA extraction repeated with 500 µl chloroform. The tube contents were mixed slowly by a rotary mixer, then spun at 13000 rpm for 4 min at 4 °C. Sodium acetate was added to the supernatant to a final concentration of 0.3 M. The DNA was precipitated with 500 µl ice-cold isopropanol. DNA was pelleted at 13000 rpm for 4 min at 4 °C. The pellet was washed in 70% ethanol by gentle inversion. The DNA was air-dried and resuspended in 25 µl elution buffer heated to 50 °C. The partially digested gDNA was dephosphorylated with calf intestinal alkaline phosphatase (CamBio) in 50 µl total reaction at 37 °C for 1 hour. The reaction was immediately quenched with EDTA then the partially digested dephosphorylated DNA was extracted with phenol-chloroform-isoamylalcohol pH 8 the chloroform and precipitated with isopropanol as described above. The DNA was air-dried and resuspended in 25 µl elution buffer heated to 50 °C. The partially digested and dephosphorylated DNA was checked on PFGE. The partial *Sau3AI* digestion reduced the gDNA from 100–200 kb to 30-50 kb in size for ligation into the SuperCosI vector.

2.11.2 Preparation of pSuperCos1 vector and ligation with insert DNA

10 µg SuperCosI vector (Stratagene) were linearised with 9 U/µg *XbaI* (Roche) in a total volume of 100 µl at 37 °C for 1 hour. Complete digestion was indicated by a linear SuperCosI band at 7.9 kb when 1 µl of the digestion was checked on a 1 % TBE agarose gel. The vector DNA was purified from the remaining reaction with the Qiagen PCR purification kit. The DNA was eluted twice in 25 µl elution buffer warmed to 50 °C. The purified *XbaI* digested DNA was treated with 30 U calf intestinal alkaline phosphatase

(Cambio), in 100 μ l total reaction for 1 hour at 37 °C. The DNA was purified with the Qiagen PCR purification kit and eluted twice in 25 μ l elution buffer warmed to 50 °C. The DNA was digested with 10 U/ μ g *Bam*HI (NEB) in 100 μ l total volume at 37 °C for 1 hour. The DNA was purified with the Qiagen PCR purification kit and eluted with twice in 25 μ l elution buffer warmed to 50 °C. The generation of fragments of the expected sizes of 1.3 kb and 6.6 kb was checked by running 2 μ l of the treated DNA and alongside the 7.9 kb *Xba*I linearised DNA on a 1 % TBE agarose gel by electrophoresis. Ligation was carried out as follows with the reaction incubated at 4 °C overnight

	Sample	Negative Control
Insert DNA <i>Sau</i> 3AI digested	2.5 μ g	2.5 μ g
pSuperCosI <i>Bam</i> HI digested	1 μ g	1 μ g
10 x T4 ligase buffer (Promega)	2 μ l	2 μ l
T4 DNA ligase 3U/ μ l (Promega)	1 μ l	-
H ₂ O	to 20 μ l	to 20 μ l

2.11.3 Phage packaging

4 μ l of each ligation reaction were packaged into λ -phage using the GigaPack III Gold Packaging Extract (Stratagene) according to the manufacturer's instructions. Briefly, packaging extracts were removed from -80 °C storage to dry ice. Extracts were rapidly thawed and 4 μ l of each ligation reaction immediately added. The extract was gently mixed and incubated at room temperature for 2 hours. 500 μ l SM buffer were added, then 20 μ l chloroform were also added. Packaged phage were stored at 4 °C and used within 1 month.

2.11.4 Phage titration

Escherichia coli XL-I Blue MR (Stratagene) was streaked from a glycerol stock at -20 °C on to LB agar and was grown at 37 °C overnight. A single colony from this plate was used to inoculate 10 ml LB broth and grown overnight at 37 °C. 1 ml of this culture was used to inoculate 50 ml LB broth containing 10 mM MgSO₄ and 0.2 % maltose. This culture was grown at 37 °C for 3 hours. Cells were recovered by centrifugation and diluted to an OD₆₀₀ of 0.5 in 10 mM MgSO₄. Packaged phage were diluted in SM buffer to a totally volume of 25 μ l. Transfection of *E. coli* cells with 1 μ l of 1/5, 1/10, 1/50 and 1/100 dilutions of phage was carried out by mixing 25 μ l cells with 25 μ l each dilution and incubating cells at room temperature for 30 min. 200 μ l LB broth were added to each tube and incubated at 37 °C

for 1 hour with gentle shaking every 15 min. The entire transfection reaction was plated out on LB agar containing 100 µg/ml carbenicillin and incubated overnight at 37 °C. The resulting colonies were counted to give the phage titre per µl of phage extract used. The resulting clones were picked into 5 ml LB broth and grown at 37 °C overnight. The packaged extract was estimated to contain 20 colony forming units per µl of undiluted extract.

2.11.5 Phage transfection of *E. coli* to construct cosmid library

This was carried out as a scaled-up version of the phage titration. It was calculated that in order to obtain the 3072 colonies needed to make the library, 225 µl of the packaged extract were transfected to construct the final cosmid library. 500 µl *E. coli* XL-I Blue MR cells at OD₆₀₀ of 0.5 in 10 mM MgSO₄ were added to 225 µl of the pooled packaged phage extracts made up to 500 µl with SM buffer. 5 ml LB broth were added and incubated at 37 °C for 1 hour with gentle mixing every 15 min. 500 µl of the transfection mix were plated out on each of 6 Genetix plates (240 x 240 x 20 mm) containing 250 ml LB agar with 100 µg/ml carbenicillin to give an estimated plating density of 750 colonies per plate. Plates were incubated inverted overnight at 37 °C.

2.11.6 Library picking and transfer to membrane

3072 colonies were picked using Q-Bot (Genetix) into 8 x 384 well archive plates containing freezing broth (LB and glycerol) and 100 µg/ml carbenicillin. Two copies of the archive plates were replicated and all copies of the library were stored at -80 °C. The entire library was spotted two-fold onto nylon membrane in a double off-set pattern. The clones on the membrane were grown on LB agar overnight and baked on to the membrane at 80 °C. The membrane was stored at -20 °C.

2.11.7 Probe preparation and library hybridisation

Three probes were amplified from *P. alba* gDNA by PCR using three pairs of primers. 1289FlanA and 1289RlanA primers amplify 230 bp of *pspA*. 3088F and 3088R amplify 547 bp of *pspB*. PMSLanAf and PMSLanBf amplify 1701 bp spanning *pspA* and 5' *pspB*. The PCR products were separated by electrophoresis on a 1 % TBE agarose gel, purified by PCR purification and ligated into pGEM-T vector (Promega) by TA cloning according to the manufacturer's instructions. 2 µl of the ligation were used to transform DH5α cells and

plated onto LB containing 100 µg/ml carbenicillin, 0.5 mM IPTG and 80 µg/ml X-Gal. One white colony of each insert cloned in pGEM-T was grown in liquid LB broth containing 100 µg/ml carbenicillin. The vector was mini-prepped from the cell pellet using a Qiagen kit. A restriction digest with *EcoRI* revealed the correct sized insert was present in each vector. The insert was amplified out of the vector and purified by gel-extraction with a Qiagen kit and the concentration measured using Nanodrop (Thermo Scientific). Each amplified probe was sequenced and found not to contain any mutations. 25 ng of the purified DNA fragment were diluted to a final volume of 50 µl in TE. The DNA was denatured at 95-100 °C for 5 min and snap cooled on ice for 5 min before centrifuging briefly. The DNA was added to a tube containing Rediprime II random prime labelling reaction mix (Amersham), to which was then added 5 µl ³²P-αdCTP and mixed by pipetting. The tube was incubated at 37 °C for 30 min and the reaction stopped by adding 5 µl 0.2 M EDTA. The tube was heated to 100 °C for 5 min then snap cooled on ice for 5 min before centrifuging briefly.

Three copies of the membrane containing the cosmid library clones were prepared by soaking for 1 hour at 42 °C in a pre-hybridisation solution of: 5 x SSC, 0.5 % SDS, 1 mM EDTA (pH 8). The solution was replaced with fresh pre-hybridisation solution and soaked for a further 2 hours. The bacterial debris were scraped off using a damp paper towel and the membranes were rinsed twice in 2 x SSC.

The membranes were each placed into a large hybridisation tube containing 50 ml of hybridisation buffer warmed to 60 °C and were pre-hybridised for 5 hours at 65 °C in a Techne rotisserie oven. The pre-hybridisation solution was replaced with 10 ml of fresh hybridisation buffer. To each of the three tubes, 30 µl of one labelled probe were added. The filter was hybridised at 60 °C in a Techne rotisserie oven for a total of 20.5 hours. The hybridisation solution was removed and the filter washed five times in 30 ml of pre-warmed wash buffer at 60 °C until the counts per minute were reduced to ~20. The filter was wrapped in clingfilm and exposed to a phosphor plate overnight. The plate was visualised using a phosphorimager (Fuji). Alignment of the membrane with the grid pattern of the 386 well plates along with the double off-set pattern of the spots allowed the location of the positive clones to be identified. Positive clones were selected from the library and grown on LB agar containing 100 µg/ml carbenicillin.

2.12 Cosmid sequencing and sequence analysis

Cosmid end-sequencing was carried out as described in 2.6.7 using the primers End_F and End_R to obtain approximately 800 base pairs of sequence each end. Cosmids B4-1 and F13-1 did not appear to contain lantibiotic biosynthetic genes on the ends of the cosmid so were sequenced using Sanger sequencing by the Cambridge University DNA

Sequencing Service. The complete cosmid sequences were annotated using Artemis (Sanger Centre, Cambridge). Open-reading frames were called on the basis of the GC content across triplets. Start sites, ATG or GTG, were further adjusted by locating the presence of ribosome-binding sites (RBSs) approximately 6-10 nucleotides upstream of the start site and based around the sequence GGAGG. A BLAST search of the National Center for Biotechnology Information (NCBI) database using a protein query was used to annotate these putative open-reading frames based on homology.

2.13 PCR targeting of cosmids

REDIRECT© technology; 'PCR targeting system in *Streptomyces coelicolor* A3(2)' by Bertolt Gust, Tobias Kieser and Keith Chater (2002) was used as a basis to carry out PCR targeting. This technology is based on the PCR targeting system described by Datsenko and Wanner and was further developed by Bertolt Gust (Datsenko and Wanner 2000; Gust *et al.* 2003). The protocol was further adapted for use with *Planomonospora* through some minor modifications.

2.13.1 Construction of integrative cosmids

The B4-1 cosmid (pIJ12321) was targeted to introduce an integrase (*intΦC31*), phage attachment site (*attP*), origin of transfer (*oriT*) and apramycin resistance cassette (*aac(3)IV*), creating pIJ12323. Cosmid F13-1 (pIJ12322) was similarly targeted to create pIJ12324. This was achieved through the replacement of *neo* on the SuperCosI backbone with the 5247 bp *oriT-attP-int-aac(3)IV* cassette.

E. coli BW25113/pIJ790 containing the SuperCosI cosmid to be targeted was grown in 10 ml L broth containing 50 µg/ml carbenicillin, 50 µg/ml kanamycin and 25 µg/ml chloramphenicol overnight at 30 °C with shaking at 250 rpm. 10 ml SOB (without Mg) containing carbenicillin, kanamycin, chloramphenicol and 100 µl 1 M L-arabinose were inoculated 1 in 100 with this overnight culture and was incubated with shaking at 30 °C for ~4 hours until OD₆₀₀ ~0.6. Electrocompetent cells were generated from the induced culture as described in 2.5.8. 50 µl of the cell suspension were mixed with ~100 ng gel purified 5247 bp *SspI* fragment derived from pIJ10702 (pMJCOS1). Electroporation was carried out as 2.5.8. Typically 500 µl aliquots of the transformation mix were plated onto two plates of L agar containing 50 µg/ml apramycin and 100 µg/ml carbenicillin and were incubated overnight at 37 °C.

After incubation overnight, colonies were scraped from one transformation plate in 1 ml H₂O. The cells were pelleted by centrifugation and the cosmid DNA purified using the standard cosmid preparation protocol described in 2.5.2. 1 µl of the resulting cosmid DNA was used to transform 50 µl chemically competent *E. coli* DH5α, as described in 2.5.9. Typically 200 µl and 800 µl of the transformation were plated on to L agar containing 50 µg/ml apramycin and 100 µg/ml carbenicillin and were incubated overnight at 37 °C. Resulting transformants were picked into 10 ml L broth with the same antibiotic selection and were grown overnight at 37 °C. Cosmid DNA was prepared from these overnight cultures following the standard cosmid preparation protocol described in Section 2.5.2. Cosmid DNA which had been successfully targeted with the pIJ10702 *Sspl* fragment was confirmed by *NotI* restriction digest which yields a diagnostic band-shift from 6807 bp to 8682 bp. Restriction digest with *Bam*HI was used to confirm the identity of the cosmid.

2.13.2 Construction of mutant cosmids by gene replacement

The mutation of a specific cosmid by gene replacement was carried out in *E. coli* BW25113/pIJ790 containing the cosmid to be targeted essentially as described in 2.13.1. The gene replacement cassette was amplified by high-fidelity PCR using the vector pIJ10700 containing the *hyg-oriT* cassette as a template. The primers were designed to anneal to the universal primer binding sites of 19 and 20 bp which flank the cassette, and have 39 bp extensions homologous to the sequence flanking the region of cosmid to be replaced (primer sequences are listed in Table 2.5). PCR was carried out as follows:

	Amount/ Concentration	Volume (µl)	Final Concentration
Template DNA	100 ng/µl	0.5	
Forward Primer	100 µM	0.5	1 µM
Reverse Primer	100 µM	0.5	1 µM
dNTPs	40 mM (10 mM each)	1	800 µM (200 µM each)
Expand HiFi Buffer 2 (Roche)	10x	5	1x
DMSO	100 %	2.5	5 %
Expand HiFi Taq (Roche)	2.5 U/µl	1	2.5 U
dH ₂ O		39	
Total		50	

94 °C 2 min	}	S
94 °C 45 s		
50 °C 45 s		
72 °C 90 s		
94 °C 45 s	}	HN
55 °C 45 s		
72 °C 90 s		
72 °C 5 min		

The PCR reaction was treated with 1 μ l *DpnI* (NEB; 20 U/ μ l) for 1 hour at 37 °C to remove template DNA. The PCR-amplified template was purified using the Qiagen PCR purification kit and was eluted in 25 μ l elution buffer heated to 55 °C. 1 μ l of the purified cassette was separated on a 1 % TBE agarose gel by electrophoresis to check the size and quality of the cassette DNA. 1 μ l (approximately 100 ng) of the purified cassette DNA was used to transform *E. coli* BW25113/pIJ790 containing the cosmid, and prepared as described in 2.13.1. Transformants were selected on DNA agar containing 40 μ g/ml hygromycin. Transformed cosmids were transferred to *E. coli* DH5 α as described in 2.13.1 and the correct targeting of the cosmid confirmed by PCR, using primers flanking the cassette insertion (listed in Table 2.5). The resulting confirmation PCR products were purified using a PCR purification kit, and the correct replacement of the gene in question was confirmed by sequencing as described in 2.6.7. To confirm that no other rearrangements had occurred in the targeted cosmids, targeted and WT cosmids were subjected to restriction digest analysis. Usually a couple of digests were performed individually with either; *NotI*, *BamHI*, *NcoI* or *PstI*, and fragments compared by separating on a 1 % TBE agarose gel by electrophoresis. The restriction digest patterns were compared to those from *in silico* digests of the respective WT and mutant cosmids (using NEB Cutter).

2.13.3 Construction of minimal gene set

Minimal gene sets were created by gene replacement of DNA regions outside of the biosynthetic gene cluster. The vector pIJ773 containing apramycin^R-*oriT* cassette was the template used to amplify the extended resistance cassette. The primers were designed to contain an additional restriction site for an enzyme which does not cut elsewhere on the cosmid. This site is located between the universal primer binding site and the 39 bp extension which has homology to the cosmid. To remove the genes upstream of the cluster, *SpeI* sites were introduced, while *XbaI* sites were introduced to remove the genes downstream of the cluster. Gene replacement was carried out in *E. coli* BW25113/pIJ790 containing cosmid B4-1 essentially as described in 2.12.2. The region of cosmid upstream and downstream of the biosynthetic gene cluster were targeted in turn. The targeting was

checked by PCR using primers annealing outside the resistance cassette, and by a diagnostic restriction digest. A preparative digest was performed on the targeted cosmid, using the enzyme which recognises the site engineered into the primers. The cosmid was gel extracted and religated to remove the resistance cassette. This was checked by PCR and restriction digest. The minimalised cosmid was targeted to replace *neo* on the SuperCosI backbone with the *oriT-attP-int-aac(3)IV* cassette as described in 2.13.1.

2.13.4 FLP-mediated recombination to generate scar mutants

E. coli DH5 α /BT340 was grown in 10 ml L broth containing 25 μ g/ml chloramphenicol overnight at 30 °C with shaking at 250 rpm. 10 ml L broth containing chloramphenicol were inoculated 1 in 100 with this overnight culture and was incubated with shaking at 30 °C for 3-4 hours until OD₆₀₀ ~0.6. Electrocompetent cells were generated from the induced culture as described in 2.5.8. 50 μ l of the cell suspension were mixed with ~100 ng cosmid containing the FRT-flanked cassette to be removed. Electroporation was carried out as in 2.5.8. Typically 200 μ l and 800 μ l aliquots of the transformation mix were plated onto DNA agar containing 25 μ g/ml chloramphenicol and 80 μ g/ml hygromycin and were incubated for 2 days at 30 °C. Approximately eight single colonies were picked and streaked on to L agar without antibiotics and incubated at 42 °C overnight to promote FLP-recombination. Single colonies from these plates were picked and streaked first onto DNA agar containing hygromycin and then onto L agar containing carbenicillin and kanamycin. The plates were incubated at 37 °C overnight and compared to identify clones sensitive to the selection for the cassette used for mutagenesis (hyg) but resistant to the selection for the cosmid backbone (carb and kan). Four to eight such clones were selected and the removal of the cassette confirmed by PCR using flanking primers and the PCR products checked by Sanger sequencing.

The scar cosmid has had the *oriT* removed along with the resistance marker. One option to introduce the scar cosmid into *P. alba* is to use protoplast transformation. Alternatively the scar cosmid can be targeted once more to re-introduce the *oriT* in order to facilitate the mobilisation of the construct by conjugal transfer. The *bla* gene on the SuperCosI backbone was targeted for replacement with *bla-oriT-resistance marker-bla*. The scarred construct with *oriT* and additional resistance marker introduced into the backbone in place of *bla* was conjugated into *P. alba*. Kanamycin-resistant *P. alba* colonies containing the entire scar cosmid integrated by a single crossover were isolated. These were restreaked onto to kanamycin-free ISP4, followed by screening for concomitant loss of kanamycin resistance, indicating a double crossover event has occurred to leave the 81 bp scar sequence in the *P. alba* chromosome in place of the gene of interest.

2.14 Conjugation methods

2.14.1 *Streptomyces*

E. coli ET12567/pUZ8002 harbouring the plasmid or cosmid to be transferred was grown overnight in 10 ml L broth containing the plasmid selection as well as 50 µg/ml kanamycin and 25 µg/ml chloramphenicol at 37 °C with shaking at 250 rpm. 100 µl of this overnight culture were inoculated into 10 ml L broth containing the same selection and was grown at 37 °C with shaking at 250 rpm until the culture reached an OD₆₀₀ of 0.4. The cells were pelleted at 3000 rpm for 10 min, then washed twice with 10 ml of L broth to remove antibiotics. The pellet was then resuspended in 1 ml of L broth. For each conjugation 10 µl (approximately 10⁸) *Streptomyces* spores were added to 500 µl 2x YT broth. Spores were heat shocked at 50 °C for 10 min and allowed to cool. Then 0.5 ml of the *E. coli* cell suspension was added to the spores. The mix was serially diluted and the 10⁻¹, 10⁻² and 10⁻³ dilutions were plated out onto SFM media containing 10 mM MgCl₂. The plates were incubated at 30 °C for 16-20 hours. Each plate was overlaid with 1 ml water containing 0.5 mg nalidixic acid (20 µl of 25 mg/ml stock to give a final concentration of 20 µg/ml in 25 ml agar plate) to kill *E. coli* donor cells and the selection for the plasmid as appropriate, as detailed in (Kieser *et al.* 2000). The incubation was continued at 30 °C. Single exconjugants were picked after 3-5 days growth and were streaked on to SFM containing 25 µg/ml nalidixic acid and the plasmid selection (typically 50 µg/ml apramycin).

2.14.2 *P. alba*

E. coli ET12567 pUZ8002 harbouring the plasmid or cosmid to be transferred, were prepared for conjugation as described in 2.14.1. For each conjugation, 10 µl (approximately 10⁶) *Planomonospora* spores were added to 500 µl ISP4 broth. Spores were heat shocked at 50 °C for 10 min to synchronise germination and allowed to cool. In some cases, spores were incubated at 37 °C 250 rpm for 2.5 hours after heat shock then vortexed to disperse germ tubes which may have clumped together. Then 0.5 ml of the *E. coli* cell suspension was added to the germinated spores and inverted to mix. 200 µl aliquots were plated onto ISP4 media containing 10 mM MgCl₂. The plates were incubated at 30 °C for 18-20 hours. Each plate was overlaid with 1 ml water containing 0.5 mg nalidixic acid and the appropriate selection for the plasmid. Typically this selection was 1.25 mg hygromycin (25 µl of 50 mg/ml stock to give a final concentration of 50 µg/ml in 25 ml agar plate) or 0.5 mg apramycin (10 µl of 50 mg/ml stock to give a final concentration of 20 µg/ml in a 25 ml agar plate). One plate from each conjugation was overlaid with 1 ml water containing 0.5 mg nalidixic acid only as a control. Incubation was

continued at 30 °C. Single exconjugants were picked after 10 days for small integrative constructs, or 2-4 weeks for homologous recombination with cosmids.

2.14.3 *Nonomuraea*

Nonomuraea mycelium was harvested from 24-48 hour old cultures by centrifugation at 4000 rpm for 10 min and was washed twice in ice cold 20 % glycerol. The mycelial pellet was resuspended in approximately half the original volume ice-cold 20 % glycerol. Mycelial samples were used immediately or were stored at -80 °C and defrosted on ice prior to conjugation. *E. coli* ET12567 pUZ8002 harbouring the plasmid or cosmid to be transferred, were prepared for conjugation as described in 2.14.1. 0.5 ml of mycelium was mixed with 0.5 ml of *E. coli* for each conjugation. 200 µl aliquots were plated onto V0.1 media containing 10 mM MgCl₂. The plates were incubated at 30 °C for 18-20 hours. Each plate was overlaid with 1 ml water containing 0.5 mg nalidixic acid and the appropriate selection for the plasmid (typically 1.25 mg apramycin). One plate from each conjugation was overlaid with 1 ml water containing 0.5 mg nalidixic acid only as a control. Incubation was continued at 30 °C. Single exconjugants were picked after 14 days for small integrative constructs, or 3-5 weeks for homologous recombination with cosmids.

2.15 Generation of mutants by homologous recombination in *P. alba*

The region of cosmid B4-1 or a truncated derivative encoding the gene to be deleted was replaced with the *hyg-oriT* cassette amplified by high-fidelity PCR from pIJ10700 as described in 2.12.2. The successful mutation of the cosmid to $\Delta orf::hyg-oriT$ was confirmed by PCR using the appropriate verification primer pair (listed in Table 2.5) along with restriction digest to confirm the integrity of the rest of the cosmid. The mutant cosmid was transferred to *E. coli* ET12567 pUZ8002 and the resulting strain used to transfer the cosmid to *P. alba* spores by conjugation as described in 2.14.2. Exconjugants were typically streaked for single colonies successively three times on selected on ISP4 containing 25 µg/ml nalidixic acid and 40 µg/ml hygromycin. These were sub-cultured in ISP4 containing 40 µg/ml hygromycin for 7 days at 30 °C to promote double-crossover recombination events. Mycelium from this culture was streaked to obtain single colonies on ISP4 containing 40 µg/ml hygromycin. Single colonies were assayed for growth on ISP4 containing 50 µg/ml kanamycin. Sensitivity to kanamycin indicated a double-crossover recombination event so these colonies were sub-cultured in ISP4 containing 40 µg/ml hygromycin for 7 days at 30 °C, from which mycelium was stored as a master cell bank at -80 °C. A further sub-culture in ISP4 containing 40 µg/ml hygromycin was grown

at 30 °C for 2-3 days for preparation of genomic DNA, as described in Section 2.10.1. This genomic DNA was analysed by PCR using the appropriate verification primer pair. The absence of the WT gene was also confirmed using gene internal primers (listed in Table 2.5).

2.16 Complementation of mutant phenotypes

Complementation of mutant phenotypes was carried out via the introduction of the deleted open-reading frame *in trans* with expression from the native promoter of the open-reading frame in question. The vector was pSET152, which contains the apramycin resistance marker and integrates at the Φ C31 integration site in *P. alba*.

To complement the deletion of *pspA*, the region between *pspX* and *pspA* (P_{pspA}) and the open-reading frame of *pspA* were amplified by high-fidelity PCR using primers PpspAF_BamHI and PspAR_XbaI. The resulting fragment was cloned into *Bam*HI and *Xba*I double digested pSET152 to create pIJ12532.

To complement the deletion of *pspX*, the region between *pspA* and *pspX* (P_{pspX}) and the open-reading frame of *pspX* were amplified by high-fidelity PCR using primers PpspXF_BamHI and PspXR_XbaI. The resulting fragment was cloned into *Bam*HI and *Xba*I double digested pSET152 to create pIJ12539.

To complement the deletion of *pspR*, the region between *pspQ* and *pspR* (P_{pspR}) and the open-reading frame of *pspR* were amplified by high-fidelity PCR using primers PpspRF_XbaI and PspRR_BamHI. The resulting fragment was cloned into *Bam*HI and *Xba*I double digested pSET152 to create pIJ12540.

To complement the other deletion mutants two further vectors were generated based on pSET152 containing either the intergenic region between *pspX* and *pspA* (P_{pspA} ; pIJ12533), or between *pspW* and *pspJ* (P_{pspJ} ; pIJ12536). These regions were amplified by high-fidelity PCR using primers PpspAF2_BamHI and PpspAR_SpeI_BamHI, and PpspJF_SpeI_BamHI and PpspJR_BamHI respectively. These primers were designed to allow introduction of the fragment into pSET152 using *Bam*HI, as well as introducing a unique *Spe*I site. The resulting fragment was first cloned into the commercial vector pGEM-T. Four white colonies were prepped for plasmid DNA and the insert amplified in high-fidelity PCR using pSET152F and pSET152R primers. The PCR product was sequenced and plasmids lacking spontaneous mutations were liberated from pGEM-T in a *Bam*HI digest. The isolated fragment was ligated into *Bam*HI digested, dephosphorylated pSET152. Colony PCR using one primer within the insert (1289FlanA or -3confirmR) and

one primer in the vector (pSET152F2) identified clones in which the insert was ligated in the correct orientation so that the introduced *SpeI* site is next to the *XbaI* site present in pSET152. This created pIJ12533 (pSET152-P_{pspA}-*SpeI*-*XbaI*) and pIJ12536 (pSET152-P_{pspJ}-*SpeI*-*XbaI*). The appropriate *psp* gene and its cognate putative RBS (GGAGG lying 9-12 bp upstream from the start codon) were amplified from pIJ12321 by high-fidelity PCR using the appropriate primer pair (listed in Table 2.5). Primers were designed to allow the ORF to be cloned into the *SpeI* site introduced with the promoter and the *XbaI* site already present in pIJ12533 and pIJ12536 such that they lay downstream of the respective promoter region.

All constructs were confirmed by Sanger sequencing. The resulting constructs, along with the empty vector controls (promoter only), were mobilised into the respective mutant strain via conjugation from *E. coli* ET12567 pUZ8002 as described in Section 2.14.2.

2.17 Bioassay methods

2.17.1 Solid bioassay

The indicator organism was *M. luteus* ATCC 4698. 50 µl of a glycerol stock were streaked onto L agar plates and incubated for 2 days at 30 °C. A single colony was used to inoculate L broth which was grown overnight at 30 °C with shaking at 250 rpm. This culture was diluted 1 in 25 into 50 ml L broth and grown at 30 °C with shaking until an OD of 0.4-0.6 was reached. This culture was diluted 1 in 10 into molten soft nutrient agar at 50 °C and approximately 10 ml used to overlay each 25 ml plate containing the growing producer organism. The plates were incubated at 30 °C overnight, or until defined halos were visible.

2.17.2 Liquid bioassay

The indicator organism was *M. luteus* ATCC 4698. An overnight culture was grown and subcultured as described in 2.13.1. This culture was diluted 1 in 10 into molten L agar at 50 °C and approximately 50 ml poured into 100 mm square petri dishes. 40 µl samples of spent culture supernatant were applied as 2x 20 µl to sterile antibiotic assay discs and allowed to dry. These were then placed onto plates containing the target organism and the plates incubated at 30 °C overnight, or until halos are clearly defined.

2.18 Mass Spectrometry

2.18.1 Matrix-Assisted Laser-Desorption Ionisation Time of Flight (MALDI-ToF)

2.18.1.1 Low mass range: 700 - 4000 Da

Culture supernatants were diluted 1 in 5 with 5 % formic acid (ARISTAR[®], VWR). The diluted sample (~0.8 µl) was spotted onto a PAC plate (Prespotted AnchorChip[™] MALDI target plate, Bruker Daltonics) and the spots were washed briefly with 8 µl 10 mM ammonium phosphate/ 0.1 % TFA according to the manufacturer. After drying, the samples were analysed by MALDI-TOF on a Bruker Ultraflex TOF/TOF. The instrument was calibrated using the prespotted standards (~200 laser shots). Samples were analysed using a laser power of approximately 25 % and spectra were summed from ~20 x 20 laser shots. The accuracy of MALDI-ToF MS was 20-50 ppm. Analysis of mass spectrometry data and figure preparation was carried out using FlexAnalysis software (Bruker Daltonics).

2.18.1.2 High mass range: 4000 - 12000 Da

Culture supernatants were diluted 1 in 5 with 5% formic acid (ARISTAR[®], VWR). Matrix (sinapic acid, Fluka/Sigma-Aldrich) dissolved in 90 % acetonitrile/10 % of 0.1 % trifluoroacetic acid (Rathburn Chemicals Ltd) at 2 mg/ml was spotted onto an AnchorChip[™] MALDI steel target plate (Bruker Daltonics). Samples (diluted culture supernatants) were then spotted onto the dried matrix spots and incubated for 3-5 min. The spots were washed shortly with 10 mM ammonium phosphate, 0.1 % TFA according to the manufacturer. Ubiquitin and insulin (Sigma-Aldrich) were spotted with sinapic acid as matrix in a similar way and used as standards for mass calibration. The samples were analysed by MALDI-TOF on a Bruker Ultraflex TOF/TOF using a method optimised for the mass range up to 12 kDa in linear mode. The data was processed in FlexAnalysis (Bruker Daltonics).

2.18.2 Electrospray ionisation (ESI)

For sequence analysis, culture supernatants were applied to a Qtof II[®] mass spectrometer (Micromass) using Picotip[®] emitters (Newobjective). The instrument was setup in positive ion mode and calibrated in the range of m/z 100-2200 using MS/MS fragments of *Glu-Fibrinopeptide B* (*Sigma*) as standards. Planosporicin was analysed both by full scan MS analysis for mass determination and by MS/MS fragmentation for

sequence analysis. Fragmentation was carried out at a collision energy setting of 45 V. A large number of spectra was collected and processed using the Masslynx® 4.1 software (Waters). For *de novo* sequencing the PepSeq option in Masslynx was used, and peaklists were generated using the Transform option.

2.19 Extraction Methods for planosporicin

2.19.1 Methanol extraction from *P. alba* mycelium

A 100 ml flask filled with 20 ml of AF/MS media was inoculated 1 in 10 with a *P. alba* seed culture and grown under normal culture conditions. After 4-10 days, mycelium from the entire culture collected by centrifugation at 4000 rpm for 15 min. A sample of supernatant was collected for analysis and the rest discarded. The cell pellet was extracted with 20 ml of MeOH and vortexed three times for 1 min. A 1 ml sample was taken and sonicated three times at 30 seconds in a Branson 2510 sonicator. The remaining ~19 ml of vortexed mycelium was centrifuged at 4000 rpm for 15 min and 1 ml of the supernatant was dried down in a miVac Duo concentrator (GeneVac). Likewise the vortexed and sonicated mycelium was centrifuged at 13000 rpm for 5 min and the supernatant dried down to ~100 µl. Extracts were stored at -20 °C.

2.19.2 Concentration of planosporicin with Diaion HP20 bead matrix

Concentration of planosporicin was carried out using a method based on (Castiglione *et al.* 2007). Supernatant was harvested from a culture of *Planomonospora* and filtered through 0.2 µm (Millipore). The supernatant was acidified to ~pH 3 with concentrated sulphuric acid. Diaion HP20 polystyrene resin (Mitsubishi Chemical Co.) was prepared by washing in 100 % methanol for 15 min followed by water for 15 min. The resin was added to the supernatant at 2.6 % v/v. Tubes containing supernatant and resin were incubated for two hours at 4 °C on a vertical rotating mixer. The resin was then recovered by centrifugation at 4000 rpm for 5 min. The bulk of the 'spent' supernatant was carefully removed with a pipette and stored for later analysis. The resin was washed once in 5 ml methanol/water 1:1 (v/v) for 1 min and the wash fraction stored for later analysis. The resin was eluted in 1 ml methanol/water-saturated butanol/water 9:1:1 (v/v/v) for 5 min at room temperature on a vertical rotating mixer. The eluant was stored for later analysis. The elution was repeated several times to give a series of fractions. 40 µl samples were tested for bioactivity by applying to antibiotic assay discs. Eluant samples could also be

concentrated by drying under vacuum and samples resuspended in 5 % formic acid for MALDI-ToF mass spectrometry.

2.19.3 Analytical High-Pressure Liquid Chromatography (HPLC)

HPLC was carried out on a HP Agilent 1100 machine using a C18 column with dimensions 4.6 mm x 250 mm (Waters). The column was equilibrated overnight to 65 % phase A (90 % ammonium formate (2 % v/v) and 10 % HPLC-grade acetonitrile) and 35 % phase B (30 % ammonium formate (2 % v/v) and 70 % HPLC-grade acetonitrile). A 40 µl injection of supernatant was loaded onto the column and eluted at a 1 ml/min flow rate with a linear gradient from 35 % to 90 % phase B in phase A for 30 min. UV spectroscopy data were collected at 268 nm. An autocollector was set to collect fractions when the UV 268 nm absorbance exceeded a gradient of 0.05 U/second and a threshold of 0.5 units. HPLC fractions were collected from 100 µl injections in a maximum volume of 800 µl per vial.

2.20 RNA methods

2.20.1 Isolation of RNA from *P. alba*

P. alba spores (10^8) were inoculated into 10 ml liquid ISP4 media in 25 ml conical flasks containing three glass beads and the appropriate antibiotic selection. These seed cultures were grown for one week at 30 °C with shaking at 250 rpm to allow dispersed mycelial growth. Mycelial density was estimated through measuring optical density at 450 nm. Precultures were used to inoculate 30 ml liquid AF/MS media in 250 ml conical flasks containing twelve glass beads to a starting OD₄₅₀ of 0.1-0.2.

Two methods were successfully used to extract RNA from *P. alba* mycelia at a variety of timepoints. In the first method, 3-4 ml culture were harvested, mixed with double the volume of RNAprotect solution (Qiagen), vortexed for 5 seconds and left for 5 min at room temperature. Mycelium was collected by centrifugation at 4000 rpm for 10 min; after the supernatant was discarded, the mycelial pellets were stored at -80 °C for up to 2 weeks before the extraction of RNA.

To extract RNA, mycelial pellets were gently resuspended in 200 µl TE buffer containing 15 mg/ml lysozyme and incubated at room temperature for 1 hour. 250 µl of resuspended mycelium were transferred to 2 ml screw-capped tubes to which 300 mg 0.1 mm diameter silica beads (Biospec Products Inc.) and 700 µl of RLT buffer without β-mercaptoethanol

(Qiagen) were added. Mycelium was lysed in a FastPrep machine (Thermo Scientific) for 30 seconds at a speed of 6.5. The cell debris was collected by centrifugation at 13000 rpm for 1 min at 4 °C. The resulting supernatant was cleared by two successive phenol chloroform-isoamylalcohol pH 8 extractions followed by a third extraction using only chloroform. In each case the cell lysate was added to a 2 ml tube containing 700 µl of the solvent, vortexed for 1 min and centrifuged at 13000 rpm for 10 min at 4 °C. The aqueous phase (~600 µl) was removed to an RNase-free 2 ml tube and 440 µl of 100 % ethanol added. The solution was mixed well then applied to an RNeasy mini column (Qiagen). The RNA was purified following the manufacturer's instructions and performing an on-column DNaseI digestion (Qiagen). RNA was eluted from the column twice with 30 µl of RNase-free distilled water. RNA was stored at -80 °C.

An alternative method of extracting RNA from *P. alba* involved harvesting 3-4 ml culture at the appropriate timepoint in 2 x 2 ml round-bottom tubes. Mycelium was collected by centrifugation at 13000 rpm for 1-2 min. ~1 ml of supernatant was decanted into a fresh tube for later analysis. The remaining supernatant was discarded and the mycelial pellet flash-frozen in liquid nitrogen. These pellets were moved to -80 °C for storage until the extraction of RNA.

To extract RNA, mycelial pellets were collected on dry ice. An RNAase-free pestle and mortar was cooled on dry ice with liquid nitrogen. Each mycelial pellet was ground using the pestle and mortar in a small volume of liquid nitrogen. The resulting powder was scraped into a 15 ml tube using RNAase-free plasticard. 1.5 ml TRI reagent (Sigma) was added to each tube and allowed to stand for 5-10 min at room temperature. Next 300 µl chloroform were added, vortexed for 15 seconds then left at room temperature for 2 min. The tubes were centrifuged for 15 min at 6000 rpm at 4 °C. The aqueous phase (~600 µl) was removed to an RNase-free 2 ml tube and made up to 900 µl with RLT buffer without β-mercaptoethanol (Qiagen). To this, 500 µl of 90 % ethanol were added. The solution was mixed well then applied to an RNeasy mini column (Qiagen). As described above, the RNA was purified following the manufacturer's instructions during which an on-column DNaseI digestion (Qiagen) was also performed. RNA was eluted from the column twice with 30 µl of RNase-free distilled water. RNA was stored at -80 °C. The yield of RNA depended on the timepoint sampled, but an average of 92 µg RNA per 3-4 ml of culture sampled was obtained.

2.20.2 DNaseI treatment of RNA

In addition to the on-column DNaseI digest, RNA was subject to a second DNaseI digest to ensure RNA samples were free from contamination with genomic DNA prior to RT-

PCR. 2.5 µg purified RNA were digested with 1.25 U RNase-free DNaseI (Invitrogen) in a volume of 25 µl at room temperature for 15 min. The reaction was inactivated through the addition of 2.5 mM EDTA and heated at 65 °C for 10 min. 8 µl of this reaction contained 0.73 µg RNA which was used as the template in the synthesis of cDNA (Section 2.20.4).

2.20.3 Confirming RNA quality

2.20.3.1 Checking for DNA contamination

A control PCR experiment was performed to confirm the absence of DNA contamination in RNA samples. General analytical PCR was carried out as described in Section 2.6.1 but with 150-300 ng RNA as a template and a total of 30 cycles. Primers *hrdB_RTf* and *hrdB_RTR* were used to amplify the constitutively expressed *hrdB* gene. 10 ng of *P. alba* gDNA was used as a positive control.

2.20.3.2 Checking RNA concentration

The concentration of RNA was quantified using Nanodrop (Thermo scientific). Each sample was diluted 1 µl in 4 µl H₂O for a more accurate reading. The 260/280 ratio was between 1.9 and 2.3, and the 260/230 ratio was greater than 1.8 and in most cases also greater than the 260/280 reading, as expected for pure RNA.

2.20.3.3 Checking RNA quality

A gel tank, tray and comb was soaked in 0.1 M NaOH overnight, then treated with RNAzap and rinsed in 2 x autoclaved water. 1 % agarose was made up in 1x TBE buffer and autoclaved before adding 0.5 µg/ml ethidium bromide. Approximately 1 µg RNA was separated on a 1 % TBE agarose gel by electrophoresis. This revealed both the 23S and 16S bands, and often the 5S band and larger mRNA aggregates.

2.20.4 cDNA synthesis

1µg RNA was converted to cDNA using Superscript III 1st strand synthesis Supermix (Invitrogen) following manufacturer's instructions. In addition a no enzyme blank was run for each RNA template, using standard taq PCR buffer. Reactions were carried out as follows:

Component	Volume (μ l)
2x Reaction Mix (Invitrogen)	10
Superscript III 1st strand RT enzyme mix (Invitrogen)	2
Template RNA (0.73 μ g)	8
RNase-free dH ₂ O	to 20 μ l

25 °C 10 min
 42 °C 120 min
 50 °C 30 min
 55 °C 30 min
 85 °C 5 min

The reactions were chilled briefly on ice then 1 μ l RNase H (Invitrogen) was added and incubated at 37 °C for 20 min. cDNA from each reverse transcriptase reaction was diluted 1 in 10 in H₂O and was determined by Nanodrop (Thermo Scientific) to contain approximately 80-110 ng cDNA. cDNA was either used immediately in Section 2.20.5 and 2.20.6 or stored at -80 °C until use.

2.20.5 Reverse transcriptase (RT)-PCR

Freshly diluted 1 in 10 cDNA was used as a template in standard PCR with Taq polymerase (Section 2.6.1) in a 25 μ l total reaction volume. RT-PCR primers listed in Table 2.5 were specially designed using strict rules. The length was 18-20 bp, GC content between 40-60 %, melting temperature between 55-60 °C (and a maximum different of 4 °C between primers in a pair). In addition 3' end T, hairpins and primer dimers were avoided. Furthermore the amplicon had a length of 80-150 bp, the GC content between 40-60 % and secondary structure avoided. For all PCR reactions an annealing temperature of 56 °C and 30 cycles of amplification were used. The PCR products were checked by adding 5 μ l of 10x Orange G to each reaction and analysing 10 μ l on a 2 % TBE agarose gel by electrophoresis. The positive control for each cDNA template used primers *hrdB_RTf* and *hrdB_RTR* annealing to a gene with 90 % nucleotide identity to *S. coelicolor hrdB*. The negative control used a 1/10 dilution of the RT reaction on RNA template without the RT enzyme.

2.20.6 Quantitative (q) RT-PCR

Freshly diluted 1 in 100 cDNA was used as a template in qPCR using SYBR GreenER qPCR SuperMix Universal (Invitrogen). Reactions were carried out as follows:

Component	Volume (μ l)
SYBR GreenER qPCR SuperMix Universal (Invitrogen)	12
DMSO (40 %)	3
Forward primer (10 μ M)	0.5
Reverse primer (10 μ M)	0.5
ROX	0.2
1:100 cDNA	2.5
RNase-free dH ₂ O	to 25 μ l

Primers used were those previously designed for RT-PCR and used in Section 2.20.5. In parallel to reactions using cDNA, each primer pair was also used in reactions on different dilutions of *P. alba* genomic DNA in the same 96-well plate to generate a standard curve for each gene analysed. Assuming the *P. alba* genome is 7.5 Mb (based on the statistics from 454 next generation sequencing), four genome dilutions were made with; 1.5×10^6 , 1.5×10^5 , 1.5×10^4 and 1.5×10^3 genomes. To set up each reaction, 2.5 μ l of the appropriate sample were pipetted into the bottom of each well of a white qPCR plate (ABI). Then 22.5 μ l of the appropriate master mix were pipetted into each well, taking care not to touch the tip against the well. In every qPCR run, all reactions were run in triplicate. At least 2 sets of 3 blank wells (no template) were included in each run. Negative controls used a 1/100 dilution of the RT reaction on RNA template without the RT enzyme as a template with primers hrdB_RTf and hrdBRTR. Once set up, the plate was sealed with an adhesive lid then covered in foil and kept at 4 °C for a maximum of 1 hour before being run in a Chromo4 machine (BioRad, CA, USA). The PCR program consisted of:

```

50 °C 2 min
95 °C 10 min
95 °C 15 s
56-58 °C 60 s   } x 40 cycles
95 °C 1 s
melting curve analysis
plate read

```

qPCR results were analysed using Opticon 2 Monitor software (MJ Research, Waltham, MA, USA). The quality of the data was assessed in two ways. For each well of the

reaction, a meltcurve was run. If the primers have amplified just one product then there should be one clean peak in the rate of change of fluorescence. Multiple peaks imply multiple products, from non-specific binding of the primers or the formation of primer dimers. For each well, a quantitation graph is available. These are used to check at which cycle number the amplification takes off from. For the standard curve, triplicates amplify at the same time, with the more concentrated gDNA samples taking off earlier. This is further clarified through adding 5 μ l of 10x Orange G to each reaction and analysing 20 μ l on a 2 % TBE agarose gel by electrophoresis. One band of the expected size should be visible.

The threshold cycle (Ct) value denotes how many PCR cycles are required for the sample fluorescence to reach the threshold level (i.e. the more target DNA present in a sample, the lower the Ct value). The Ct values were averaged from each triplicate (standard curve, samples and negative control) to obtain a mean. A standard curve for each primer pair was drawn by plotting Ct values against the number of genomes. A regression line was drawn and the R^2 value calculated. The primer efficiency was calculated using the equation: $E = (10^{(-1/\text{slope})} - 1) \times 100$. The standard curve was used to calculate the copy number of cDNA in sample. This number was normalised against the copy number of *hrdB* in each sample to take account of likely different concentrations of cDNA in each sample.

Chapter 3 : Structure of planosporicin

3.1 Introduction

The genus *Planomonospora* is a member of the *Streptosporangiaceae* family, with the type genus *Streptosporangium* (Goodfellow *et al.* 1990). This family was formed from a number of genera on the basis of 16S rRNA sequence similarity. In total, five validly described *Planomonospora* strains have been isolated from soil samples collected from five continents; Africa, Europe, Asia, North America and South America. These include *P. alba* from the Sudan and *P. sphaerica* from India (Mertz 1994), both of which were studied in this Chapter. Prior to this work, no antimicrobial compounds were reported to be produced by any of these five *Planomonospora* strains.

In 2004 a patent submitted by Vicuron Pharmaceuticals Inc. described another *Planomonospora* strain, *Planomonospora* sp. DSM 14920 (Losi *et al.* 2004). Over two years later, a paper described how extracts from this strain were found to contain a cell-wall active antibiotic that was not a β -lactam (Castiglione *et al.* 2007). Liquid chromatography coupled to mass spectrometry (LC-MS) identified a novel lantibiotic, referred to as NAI-97, which was patented as antibiotic ID 97518 in 2004 and subsequently named planosporicin three years later (Castiglione *et al.* 2007). Planosporicin shows a wide spectrum of *in vitro* antimicrobial activity against a range of Gram-positive pathogens that includes those resistant to methicillin and vancomycin (as discussed in Chapter 1). Given the increasing prevalence of multi-drug resistant pathogens, planosporicin and its derivatives may be of commercial interest.

This Chapter describes how a second *Planomonospora* strain was found to produce planosporicin. Conditions were optimised for its growth, sporulation and planosporicin production, which enabled studies to validate the structure of the lantibiotic.

3.2 Analysis of *P. alba* and *P. sphaerica*

3.2.1 Growth characteristics of *P. alba* and *P. sphaerica* on agar media

Prior to the commencement of this project, the only *Planomonospora* strain proven to produce planosporicin was *Planomonospora* sp. DSM 14920 (Castiglione *et al.* 2007). This strain is not publicly available, so two other *Planomonospora* strains were accessed

from culture collections to assess their ability to produce planosporicin (as described in Section 2.1 of Chapter 2).

P. alba and *P. sphaerica* were grown on a range of typical actinomycete agar media. Growth was most prolific on agar plates of ISP4, OBM, AF/MS and D/seed media (Figure 3.1). Growth of *P. alba* on ISP4 and OBM agar media resulted in production of a thin, white/cream aerial mycelium that entirely covered the yellow vegetative mycelium. However, growth morphology was substantially different on AF/MS and D/seed agar media. A thick, convoluted cream/yellow vegetative mycelium that lacked production of any aerial mycelium was observed. In contrast, growth of *P. sphaerica* on ISP4 and OBM agar media resulted in production of a pinkish/orange vegetative mycelium, almost entirely covered by a white/cream aerial mycelium bearing sporangia. Like *P. alba*, the morphology of *P. sphaerica* mycelium changed on AF/MS and D/seed agar media. *P. sphaerica* pinkish/orange vegetative mycelium was thick and convoluted on both agar media with scant white/cream aerial hyphae only visible on D/seed agar. Growth of both *Planomonospora* strains was slow on all of the agar media tested, with 5 to 10 days of incubation required to produce utilisable colonies.

3.2.2 Growth characteristics of *P. alba* in liquid media

Preliminary analysis of bioactive compounds produced by *P. alba* and *P. sphaerica* in Section 3.3.1 (see later) led to *P. alba* being the focus of this work. Thus the growth characteristics of *P. alba* were also observed in liquid medium. A steel spring was originally used to try to promote dispersed growth of the culture, however the mycelium grew in large potentially physiologically heterogeneous pellets (Figure 3.2 A-D). This could have been due to foaming prompted by increased agitation of the media by the steel spring, resulting in the deposition of mycelium onto the sides of the flask and aggregation into pellets. However the addition of 0.05 % Antifoam 289 (Sigma) appeared to have little effect on pellet formation (Figure 3.2 B). The addition of 5 % sucrose or 0.5 % proline to culture media had been shown previously to improve mycelial dispersion in two related *Streptosporangiaceae*; *Nonomuraea* and *Microbispora* (Foulston and Bibb 2010; Marcone *et al.* 2010b). However these additional ingredients did not significantly improve the dispersion of *P. alba* mycelium (Figure 3.2 C). The greatest increase in mycelial dispersal was observed when the steel spring was exchanged for glass beads to agitate the media (Figure 3.2 E-H). The addition of 0.05 % Antifoam 289, 5 % sucrose and 0.5 % proline had negligible further effects on dispersion of mycelium (Figure 3.2 F-H).

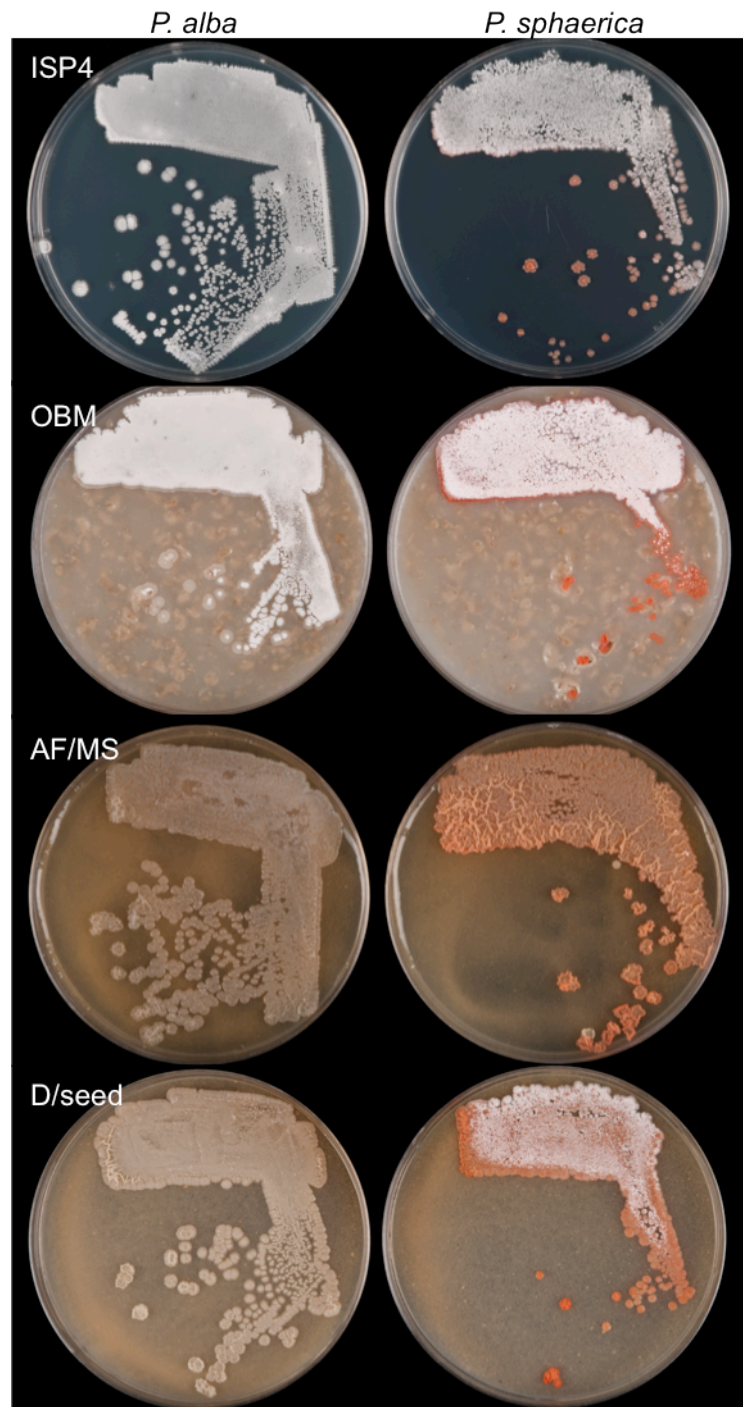


Figure 3.1 : Morphology of *P. alba* and *P. sphaerica* grown on agar media. Spores were streaked on different media (as labelled) and incubated for 7 days at 30 °C.

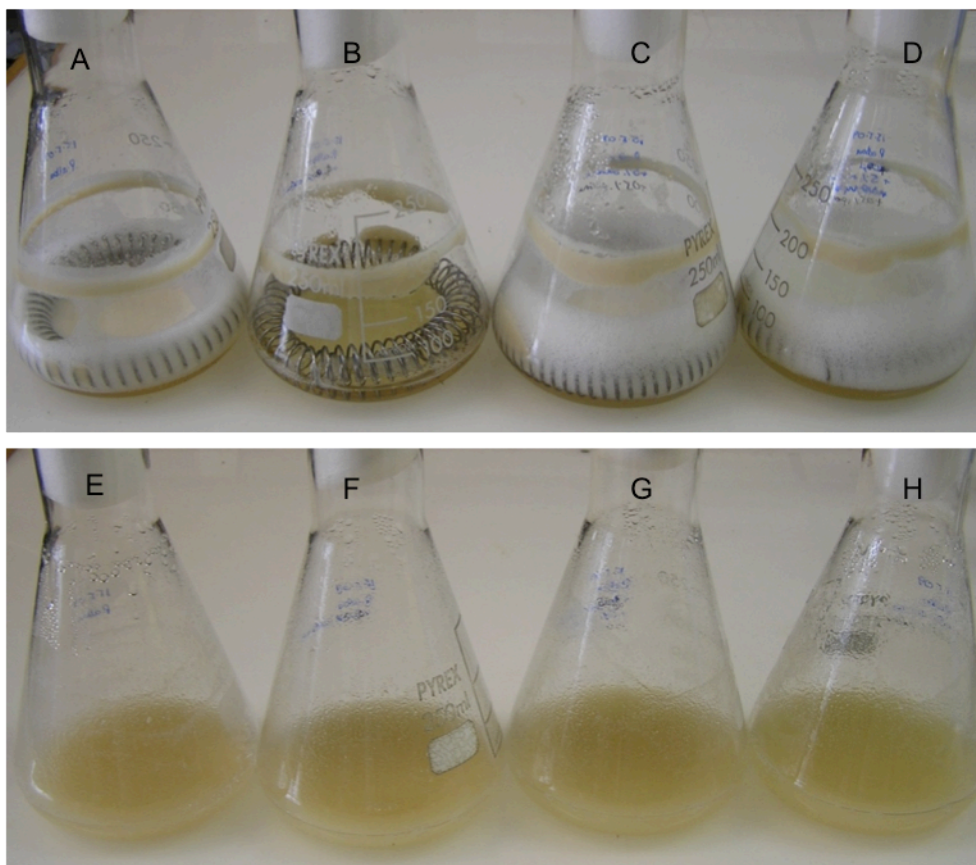


Figure 3.2 : Morphology of *P. alba* in liquid ISP4 media.

40 ml cultures were grown under different conditions for 3 days at 30 °C with shaking at 250 rpm. To promote dispersed growth, flasks A to D contain a stainless steel spring, flasks E to H contain glass beads. Flasks B and F contain 0.05 % Antifoam 289 (Sigma). Flasks C and G contain 5 % sucrose and 0.5 % proline. Flasks D and H contain 0.05 % Antifoam 289, 5 % sucrose and 0.5 % proline. All flasks were inoculated 1 in 10 from a 24 hour culture grown in 50 ml ISP4 with a stainless steel spring to promote dispersed growth.

The relatively dispersed growth of *P. alba* in liquid culture in the presence of glass beads allowed for an accurate assessment of growth rate. Growth was quantified through either a total protein assay or, more generally, by measuring the optical density of the culture at 450 nm. These methods gave essentially the same results (Figure 3.3). The growth curve of *P. alba* differed depending on the starting density of the cultures. In a typical growth curve, rapid growth commenced within 20 hours of inoculation and continued for a further 80 hours. Light microscopy revealed that the mycelial morphology changed over time from small, elongated clumps of mycelium, to dense mycelial masses with no fragmentation apparent.

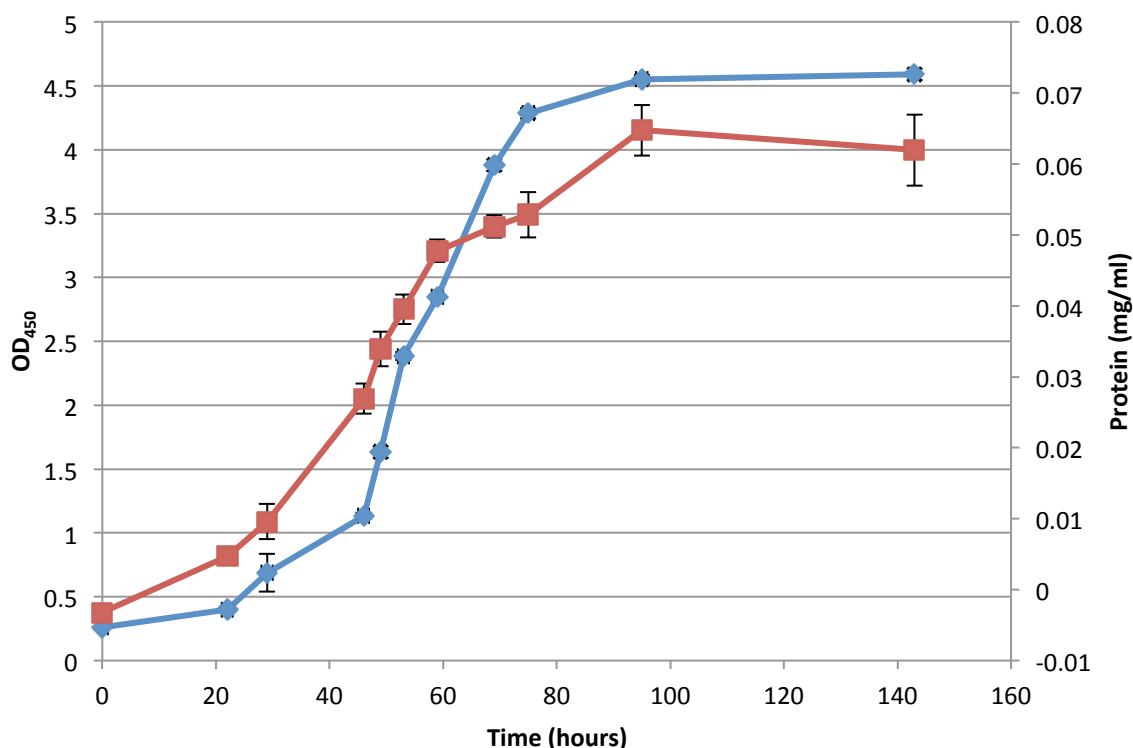


Figure 3.3 : Growth curve of *P. alba* measured by OD₄₅₀ and total protein assay.

A 48 hour preculture was used to inoculate 50 ml AF/MS cultures in triplicate to an OD₄₅₀ ~0.2. Flasks were cultured for 8 days at 30 °C with shaking at 250 rpm. At each time point, optical density at 450 nm was measured (blue diamonds) and 1 ml culture was used in a total protein assay (red squares). Standard deviation is displayed in error bars.

3.2.3 Scanning electron microscopy of *P. alba*

P. alba was grown on ISP4 agar medium to prompt sporulation. Colony structure and sporulation was observed by scanning electron microscopy (SEM). Magnification X 3,500 revealed bunches of clavate monospores (Figure 3.4 A). However these micrographs differ somewhat from those published when the species was first classified 16 years ago (Mertz 1994). In the Mertz paper, *P. alba* was grown for 14 days on ATCC medium 172 and a magnification of X 5,000 revealed spores in two parallel lines along the sporangiophore (Figure 3.4 B). This arrangement is illustrated and annotated in Figure 3.4 C (Miyadoh 1997). Yet in this work, *P. alba* did not sporulate well on ATCC medium 172, and most sporulation was observed on ISP4 medium. Even on ISP4 medium, some sporangia appear empty, with no evidence of spore release through opercula. This may imply that not all spores develop properly on this medium, or could simply be an artefact created during SEM.

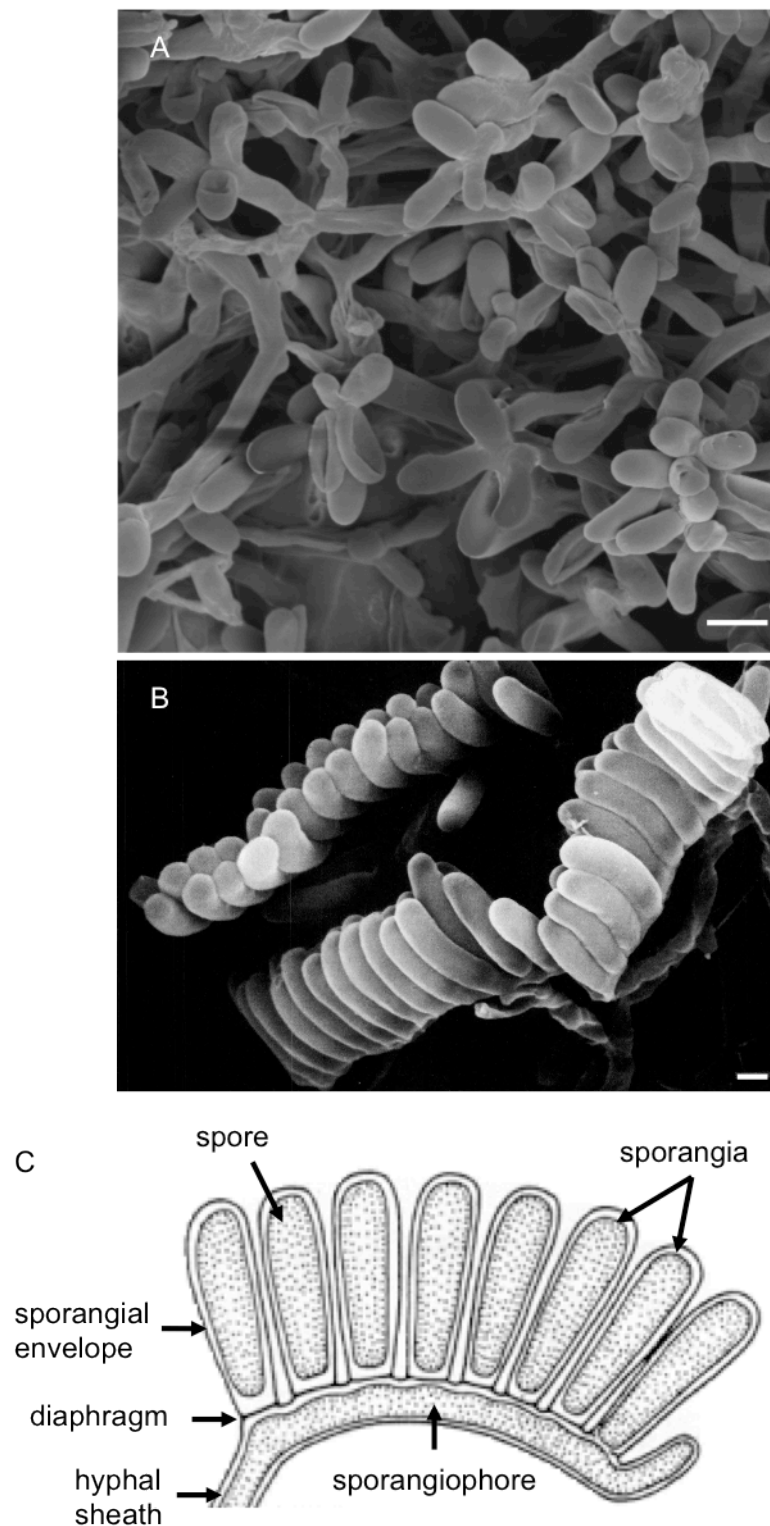


Figure 3.4 : Morphology of *P. alba* spores.

A; Scanning electron micrograph of *P. alba* grown on ISP4 media for 7 days at 30 °C, showing sporangia attached to sporangiophores. Magnification is X 3,500. Scale bar is 2 µm. **B;** Scanning electron micrograph of *P. alba* grown on ATCC medium 172 for 14 days at 30 °C, showing parallel rows of sporangia with attached sporangiophores. Magnification is X 5,000. Scale bar is 1.25 µm. From Mertz *et al.* 1994. **C;** Schematic of a sporangiophore bearing clavate monospores, adapted from Miyadoh, 1997.

3.2.4 16S rRNA

The 16S rRNA of *P. alba* was partially sequenced using the F27 and R1492 PCR primers (Mazza *et al.* 2003). The sequence is identical to that of *P. alba* JCN 9373 deposited in the National Center for Biotechnology Information (NCBI) protein database as NR_024778.1. M. Sosio kindly provided the 16S rRNA sequence of *P. sp.* DSM 14920, the NAI-97 producer strain. Multiple sequence alignment through CLUSTALW2 (Chenna *et al.* 2003) revealed that out of 1441 bp of aligned sequence, there is just one point mutation at base 1403 which is 'T' in *P. alba* and 'C' in *P. sp.* DSM 14920. A more comprehensive analysis included 16S rRNA partial sequence from *P. sphaerica* JCM 9374 (NR_024779.1), *P. parantospora* subsp. *antibiotica* JCM 3094 (NR_024730.1) and *P. venezuelensis* JCM 3167 (NR_024732.1). A phylogenetic tree based on partial 16S rRNA sequences is depicted in Figure 3.5. This reveals that *P. alba* and *P. sp.* DSM 14920 are very closely related within the *Planomonospora* genus, it seems they may represent two different strains of the same species.

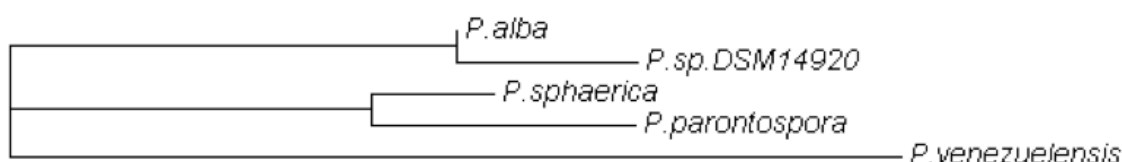


Figure 3.5 : Phylogram based on alignment of partial 16S rRNA sequences from five *Planomonospora* strains.

16S rRNA sequences were retrieved from the NCBI databank: NR_024778.1 (*P. alba* JCN 9373), NR_024779.1 (*P. sphaerica* JCM 9374), NR_024730.1 (*P. parantospora* subsp. *antibiotica* JCM 3094) and NR_024732.1 (*P. venezuelensis* JCM 3167). The alignment was performed using ClustalW2 (Chenna *et al.* 2003).

3.3 Planosporicin production

3.3.1 Detection by bioassay

Attempts to detect planosporicin production were initially carried out using a bioassay on agar plates with *Micrococcus luteus* as an indicator organism, as detailed in Chapter 2. Production of an antibiotic compound was observed by both *P. alba* and *P. sphaerica* on several media. Although *P. alba* often produced a bioactive compound on ISP4 and D/seed media, a halo of inhibition was most consistently observed on AF/MS medium (Figure 3.6). In order to characterise the nature of this bioactive compound, cultures of both strains were grown in AF/MS liquid medium and the supernatants assayed for

bioactivity against *M. luteus*. Only the supernatant obtained from *P. alba* retained its activity after passage through a 0.2 μM filter, indicating the production of a small bioactive molecule.

Supernatants obtained from different stages of a *P. alba* growth curve were also assayed for antibiotic production. Entry into stationary phase occurred after approximately 100 hours of incubation. At this point, but not before, the *P. alba* culture supernatant contained a compound that inhibited the growth of *M. luteus* (Figure 3.7). Thus the remainder of this Chapter is focused on the identification of the bioactive molecule produced by *P. alba*.

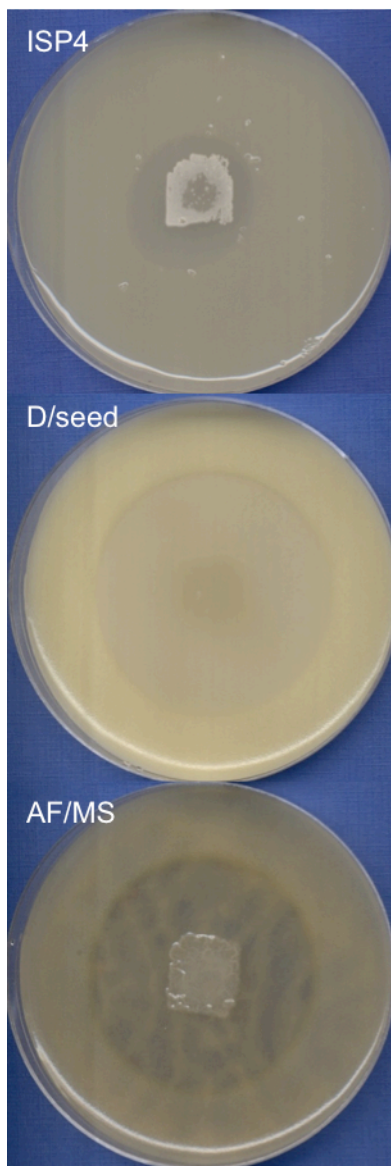


Figure 3.6 : *P. alba* solid bioassay.

P. alba was grown as a patch on plates of different media for 5 days before an overlay with the *M. luteus* indicator organism. Production of a bioactive compound by *P. alba* resulted in a halo of inhibition of *M. luteus* growth.

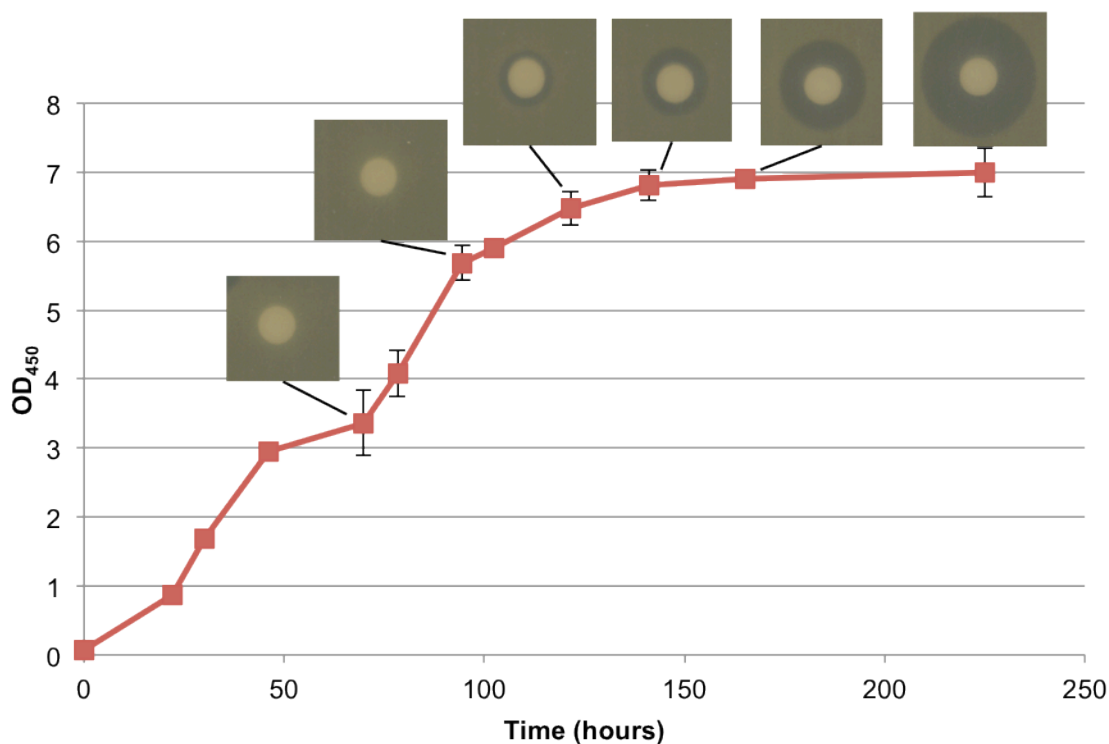


Figure 3.7 : Growth curve of *P. alba* in liquid medium.

A 7 day old preculture was used to inoculate 30 ml AF/MS cultures in duplicate to an OD₄₅₀ ~0.1; these were incubated for 10 days at 30 °C with shaking at 250 rpm. OD₄₅₀ was measured at regular intervals with standard deviation displayed in error bars. 40 µl of culture supernatant from each time point was applied to assay discs, which were laid onto a plate of *M. luteus*. Plates were incubated at 30 °C for 36 hours before being photographed.

3.3.2 Matrix-Assisted Laser Desorption/Ionisation Time-of-Flight analysis

To identify the bioactive compound, filtered culture supernatants were subjected to Matrix-Assisted Laser Desorption/Ionisation Time-of-Flight (MALDI-ToF) analysis. In MALDI, the sample is co-crystallised with a matrix compound that protects it from direct laser excitation, but allows vaporisation into the mass analyser. This soft-ionisation technique minimises sample damage, resulting in little fragmentation. The time-of-flight (ToF) mass analyser separates the ions according to their mass-to-charge ratio (m/z). Samples are acidified prior to application to the matrix, thus the masses labelled in spectra represent the protonated ion, $[M + H]^+$ or other adducts. The high resolution of the method separates the isotopes of the molecules, giving a multiplet of peaks making up the fine structure of each main peak. For all five elements that occur in peptides, the lightest isotope is also by far the commonest. The monoisotopic mass of a peptide is the mass it

would have if all constituent atoms were of the commonest isotope, thus this is the lowest mass in the multiplet.

The original reported mass for planosporicin was 2194 Da, calculated from the mass of the doubly charged ion $[M + 2H]^{2+}$ observed in electrospray ionisation mass spectrometry (ESI-MS) (Castiglione *et al.* 2007). However this was calculated from the second isotope, with a mass-to-charge ratio (m/z) of 1097.5. If instead the monoisotopic mass is used, with a m/z of 1096.9, the expected mass for planosporicin becomes 2191.8 Da (i.e. $((1096.9-1) \times 2)$). Likewise, theoretical calculations based on the presumed composition of the mature peptide predict a mass of 2191.7790 Da.

MALDI-ToF was carried out by Dr Gerhard Saalbach and Dr Mike Naldrett (JIC Proteomics Facility) in the range of 700-4000 Da. Culture supernatants of *P. alba* that possessed bioactivity revealed a compound with a monoisotopic mass of 2191.7 Da ($[M + H]^+ = 2192.7$; Figure 3.8), corroborating the theoretical calculations and the ESI-MS data published previously for planosporicin produced by *P. sp.* DSM 14920. It thus seemed likely that *P. alba* did indeed produce the same lantibiotic.

In the full scan spectrum range, multiple peaks were observed. These were attributed in part to sodium and potassium adducts of the compound, which are 23 and 39 Da larger, respectively, than the expected mass of the compound (Figure 3.8 and Table 3.1). Additional peaks likely represent oxidation of the sulphur atom in each of the five (Me)Lan bridges, creating compounds with incremental mass increases of 16 Da (Wilson-Stanford *et al.* 2009). In addition, oxidation of methionine and tryptophan may also occur during sample preparation for MALDI (Zhang *et al.* 2007). Planosporicin has one tryptophan at position six, so a total of six oxidations are possible. The most commonly observed peaks had a mass of 2192 Da, 2208 Da, 2214 Da and 2230 Da (Figure 3.8). These correspond to the protonated adduct, the oxidation of one (Me)Lan bridge, the sodium adduct and the potassium adduct, respectively. A similar profile was observed previously in the MALDI-ToF spectrum of the lantibiotic microbisporicin (Foulston and Bibb 2010).

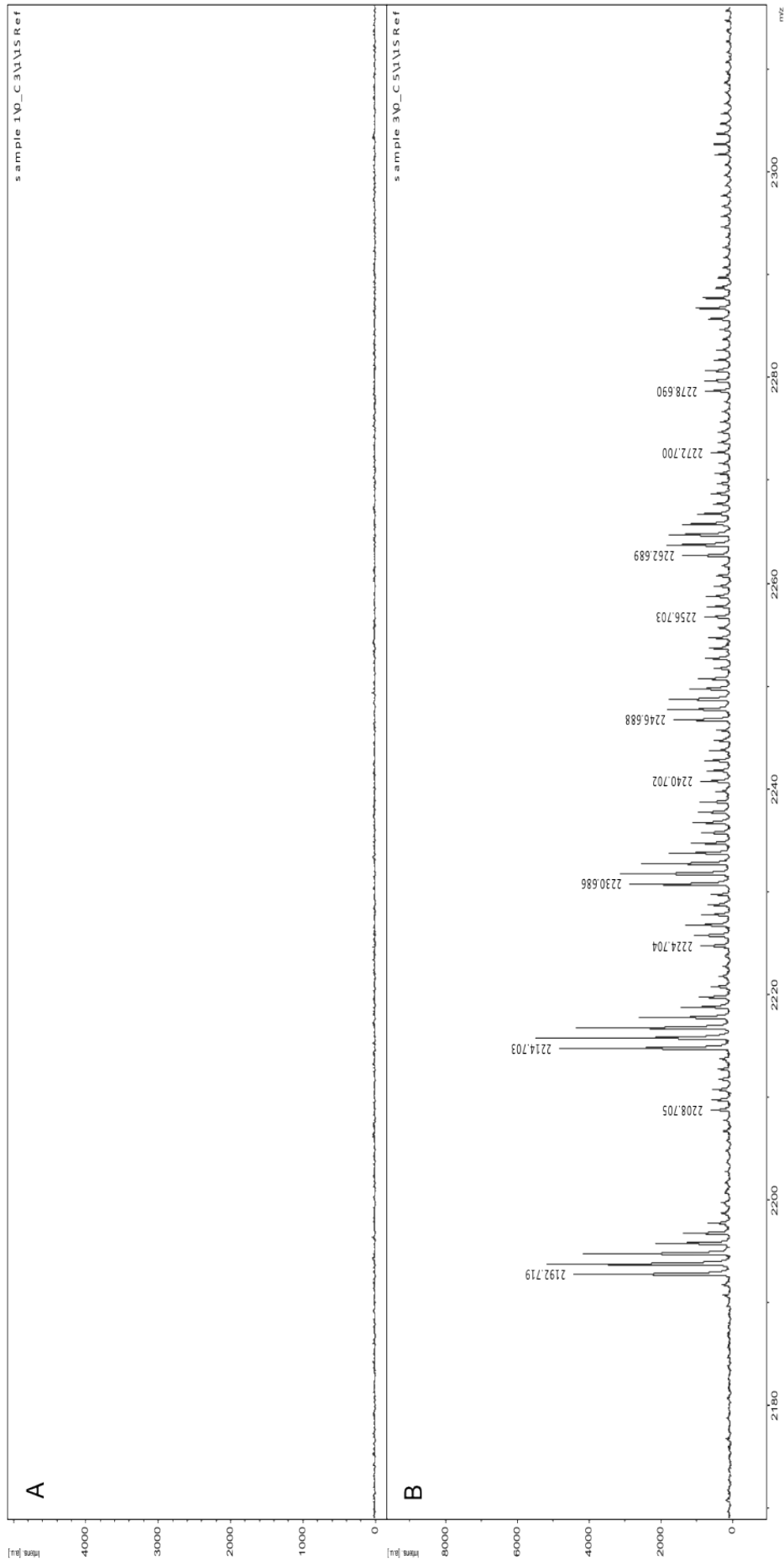


Figure 3.8 : MALDI-ToF mass spectrometry analysis.

A; Spectrum from AF/MS medium only control. **B;** Spectrum of filtered supernatant from *P. alba* cultured 9 days in AF/MS. The presence of atom isotopes within the compound means that each peptide gives a multiplet of small peaks each +1 Dalton higher than the monoisotopic mass. Intensity is given on the y-axis in arbitrary units (au) and the mass/charge ratio (m/z) on the x-axis. The monoisotopic mass of each m/z peak is labelled and the expected identity is listed in Table 3.1.

Planosporicin adduct	Predicted mass (Da)
$[M + H]^+$	2192.779
$[M + H]^+ + O$	2208.779
$[M + Na]^+$	2214.779
$[M + H]^+ + 2O$	2224.779
$[M + Na]^+ + O$	2230.779
$[M + K]^+$	2230.779
$[M + H]^+ + 3O$	2240.779
$[M + Na]^+ + 2O$	2246.779
$[M + K]^+ + O$	2246.779
$[M + H]^+ + 4O$	2256.779
$[M + Na]^+ + 3O$	2262.779
$[M + K]^+ + 2O$	2262.779
$[M + H]^+ + 5O$	2272.779
$[M + Na]^+ + 4O$	2278.779
$[M + K]^+ + 3O$	2278.779
$[M + H]^+ + 6O$	2288.779
$[M + Na]^+ + 5O$	2294.779
$[M + K]^+ + 4O$	2294.779
$[M + Na]^+ + 6O$	2310.779
$[M + K]^+ + 5O$	2310.779
$[M + K]^+ + 6O$	2326.779

Table 3.1 : The predicted masses of planosporicin adducts.

The expected m/z peaks for the protonated $[M + H]^+$, sodium $[M + Na]^+$ and potassium $[M + K]^+$ adducts of each compound are shown. The expected m/z peaks resulting from the oxidation of each lanthionine bridge are also shown as $+ nO$. m/z peaks frequently observed in spectra of *P. alba* supernatants are shaded in blue.

3.3.3 Purification of bioactive compounds

To confirm that the bioactive compound had exactly the same mass as planosporicin, a very pure and concentrated sample is needed. Several extraction methods were investigated. Vicuron Pharmaceuticals Inc. reported that planosporicin could be recovered

from the culture supernatant through extraction with water-immiscible organic solvents, particularly alkanols of at least four carbon atoms (Losi et al. 2004). However, extraction with butan-1-ol resulted in only a slight increase in the concentration of the bioactive peptide as measured by bioassay (data not shown). The use of Sep-Pak Vac (Waters) cartridges significantly increased the concentration of all compounds in the supernatant. Although a much larger halo of inhibition was observed, a lot of media components were also present, obscuring the MALDI-ToF spectra (data not shown). Diaion HP20 hydrophobic polystyrene resin was found to bind the bioactive compounds which were eluted with methanol:butanol:water (9:1:1) (Chapter 2). Bioassays of the supernatant before and after extraction revealed that the bioactive compound was concentrated in the resin eluent, affording a larger halo of inhibition (Figure 3.9 A and B). These eluted fractions were concentrated by drying, resuspended in 5 % formic acid and subjected to ESI-MS. The bioactive compound was subjected to full scan MS analysis for mass determination. MALDI-ToF was carried out by Dr Gerhard Saalbach and Dr Mike Naldrett (JIC Proteomics Facility). Eighty scans were collected and processed to obtain the exact monoisotopic mass of the double charged ion; $[M + 2H]^{2+}$ is 1096.9027 Da (Figure 3.9 C). This mass was used to accurately calculate the mass of the bioactive molecule present in the supernatant of *P. alba*; $(1096.9027-1) \times 2 = 2191.8054$ Da. This mass corresponded extremely well to the predicted mass of planosporicin of 2191.7790 Da, and suggested very strongly that *P. alba* did indeed make planosporicin.

3.4 Structural analysis of planosporicin

3.4.1 Genome scanning

The presence of (Me)Lan bridges prevents structural characterisation of planosporicin through Edman degradation. Thus a combination of ESI-MS, tandem mass spectrometry (MS/MS) and NMR was used to accumulate data to propose the structure of the 24 amino acid peptide (Figure 3.10 A) (Castiglione *et al.* 2007). The MS/MS data were published as a table of the m/z of the six major fragments. These fragments were annotated as defined parts of the mature peptide and used to justify the arrangement of Me(Lan) bridges in the proposed structure (Castiglione *et al.* 2007). However several fragments were proposed to result from cleavage at both ends of the peptide (e.g. a fragment corresponding to residues 6-23), which is unlikely in MS/MS (G. Saalbach, personal communication). Furthermore, the NOE connectivity matrix did not link the Pro-Gly motif to the rest of the molecule, suggesting that a revision of the published primary structure might also be necessary.

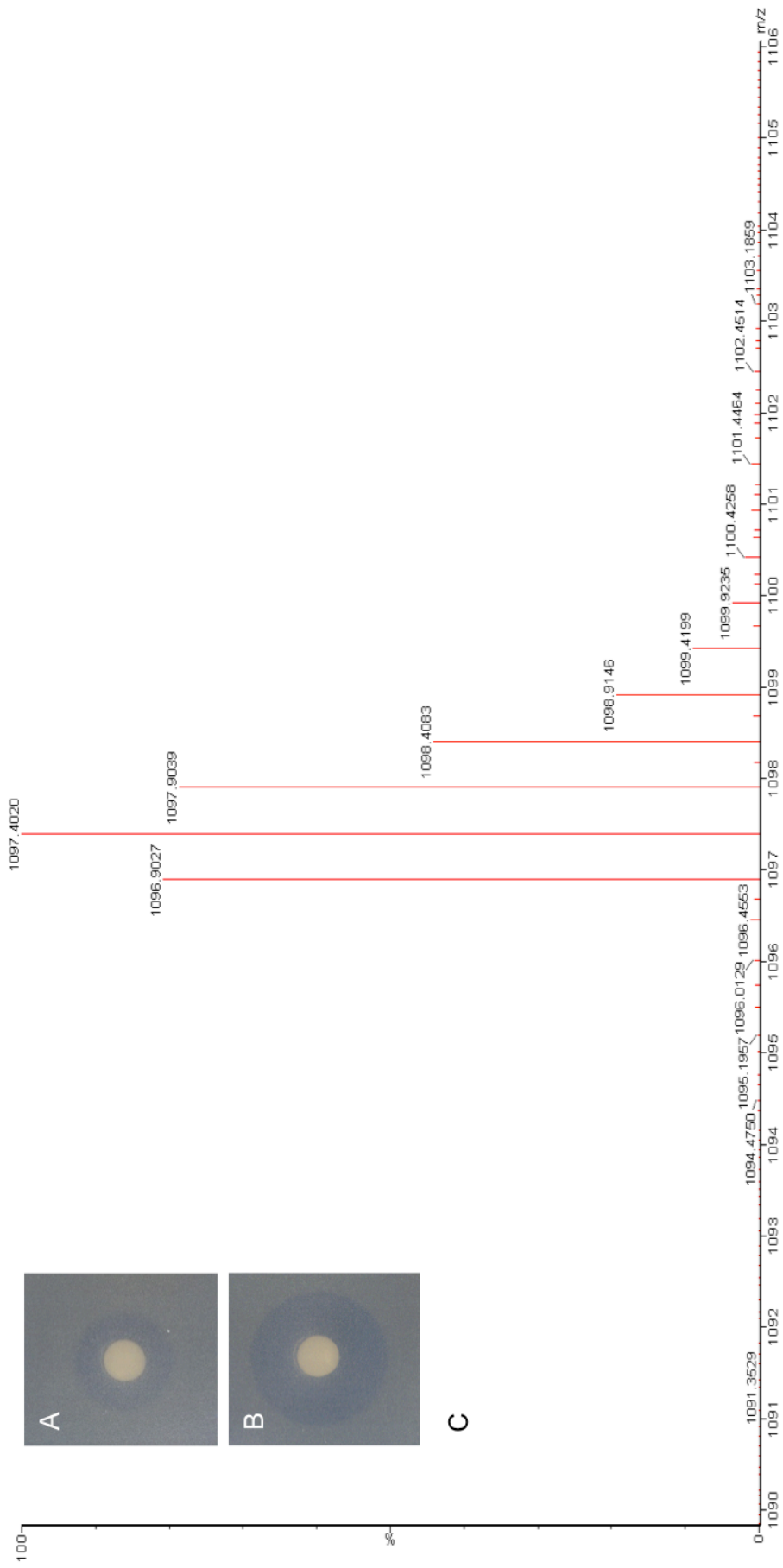


Figure 3.9 : Concentrated culture supernatant of *P. alba*.

A; Bioassay with *P. alba* supernatant after 6 days growth in AF/MS. **B;** Bioassay after elution of the same supernatant from Diaion HP20 resin. **C;** Precursor ion scan on material from B. 80 scans in from Electrospray ionisation (ESI) were combined and processed. Each peak represents the double charged ion $[M + 2H]^{2+}$ with m/z labelled above. The y-axis shows the intensity normalised to the highest peak and m/z is displayed on the on the x-axis.

High molecular weight *P. alba* genomic DNA was analysed through a 454-based genome scan (described further in Chapter 4). The published planosporicin tertiary structure published was used to predict a likely primary structure (Castiglione *et al.* 2007). The Dha and Dhb residues were predicted to be derived from serine and threonine residues, respectively. For each (Me)Lan bridge, it was initially assumed that the cysteine residue was always C-terminal to the serine or threonine residue, as found in many lantibiotics (Chatterjee *et al.* 2005). Thus the more N-terminal residue was assumed to be serine or threonine and the C-terminal residue was assumed to be cysteine. On this basis, the primary amino acid sequence of planosporicin was predicted (Figure 3.10 B). This protein sequence was used to search the 454 database using TBLASTN, one of the Basic Local Alignment Search Tool (BLAST) search programs that takes a given amino acid sequence and searches six-frame translations of a nucleotide sequence database for similar sequences (Altschul *et al.* 1990). A search of the *P. alba* 454 contig database using the predicted planosporicin propeptide sequence identified a contig encoding a 56-residue peptide (Figure 3.10 C). Closer inspection revealed that this peptide was likely to consist of a 32-residue leader peptide, followed by a 24-residue propeptide. The leader peptide contained no cysteine residues. The presence of a 'FQLD' motif between positions -17 and -14 indicated that this putative lantibiotic is likely to belong to 'Class I'. FNLD-type motifs are characteristic of type A lantibiotics, which also show particular conservation at positions -1, -2 and -4 of the leader peptide which are commonly arginine/glutamine, proline and alanine/isoleucine, respectively (Chatterjee *et al.* 2005). Planosporicin possesses the conserved proline at position -2, which implies that cleavage of the leader peptide will be through a separate LanP protease rather than a chimeric LanT transporter.

The propeptide sequence identified through genome scanning differed significantly from the predicted sequence. Firstly, the Pro-Gly motif located at the C-terminus of the published structure was revealed to be located in the middle of the peptide as Pro₉ and Gly₁₀ (Figure 3.10 B and C). Secondly, it appeared that the five (Me)Lan bridges did not have uniform directionality. As yet, all thioether crosslinks of type A lantibiotics are incorporated by LanC enzymes solely in the N-to-C terminal direction, i.e. with the cysteine residue located C-terminal to the serine or threonine residue (Chatterjee *et al.* 2005). However, the second Lan bridge of this peptide appeared to be derived from a cysteine residue that was located N-terminal to the serine residue. This reversed directionality is seen in some type B lantibiotics. For example, cinnamycin contains two MeLans in which the cysteine is positioned N-terminally to the threonine (Kaletta *et al.* 1991).

The identification of the FQLD motif indicative of type A lantibiotics was at odds with the reverse orientation of one Lan bridge. Thus further structural analysis was undertaken to determine the conformation of the (Me)Lan bridges within the molecule.

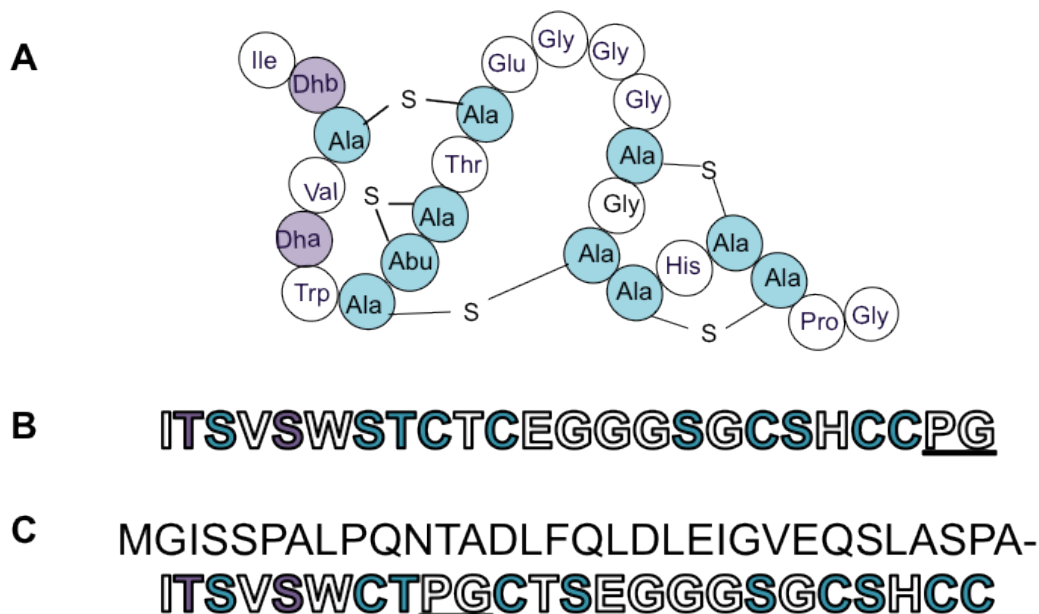


Figure 3.10 : Structure of planosporicin.

Residues are; unmodified (white), dehydrated (purple), dehydrated and cyclised (blue). **A**; The original structure proposed in 2007 by Castiglione *et al.* **B**; The predicted primary amino acid sequence based on the structure in **A**. **C**; The 56 residue peptide identified from *P. alba* by genome scanning. The 32 residue leader peptide (black) is followed by a 24 residue propeptide. Underlined is the Pro-Gly motif.

3.4.2 Tandem mass spectrometry (MS/MS)

MS/MS involves multiple rounds of mass spectrometry, with fragmentation occurring between each round. This requires a pure and concentrated sample. The concentration of planosporicin in filtered *P. alba* supernatant was too low for direct analysis. Therefore, filtered *P. alba* culture supernatant was concentrated using a hydrophobic polystyrene resin (detailed in Section 3.3.3). Several batch elutions were combined, dried down, resuspended in 5% formic acid and subjected to MALDI-ToF mass spectrometry. This extraction method was found to concentrate and partially purify planosporicin-associated masses from the *P. alba* supernatant.

During MS/MS, fragmentation is induced through collision-induced dissociation with an inert gas, resulting in bond breakage. The products are referred to as a-, b- or c-ions

(representing different points of cleavage with respect to the peptide bond in the peptide chain) when the charge is retained by the amino-terminal fragment and x-, y- or z-ions (again representing different cleavage positions) when it is retained by the carboxy-terminal fragment (Figure 3.11 A). Most commonly, breakage occurs through the lowest energy pathway, and consequently cleavage of the amide bond is favoured energetically (Steen and Mann 2004). Thus the fragments most commonly observed are b-ions when charge is retained on the N-terminus and y-ions when charge is retained on the C-terminus of the peptide. Other fragment ions are possible due to the further loss of ammonia or water (Steen and Mann 2004).

MS/MS was carried out by Dr Gerhard Saalbach and Dr Mike Naldrett (JIC Proteomics Facility). The 2191.8054 Da peptide was isolated, a round of fragmentation was carried out and a product ion scan assembled (Figure 3.11 B). The fragments obtained through MS/MS of a lantibiotic are often difficult to interpret due to the presence of the (Me)Lan bridges. As indicated above, peptides most frequently break once through the amide bond under low energy collision, but other cleavages can be induced when a higher energy collision is used. Thus in these experiments 45 V was used compared to the usual 30 V. This larger fragmentation force resulted in the cleavage of many different bonds, including the (Me)Lan bridges.

Through tabulating all the possible b-type and y-type ions that could be generated, identification of a majority of the high intensity peaks was possible (Table 3.2). Notably there were no y-ions corresponding to the loss of residues 1 to 3, and there were no fragments for the additional loss of C-terminal residues. Thus, unlike the data published in 2007 by Castiglione *et al.*, double breaks leading to losses at both termini were not detected. Three of the main MS/MS peaks correspond to fragments containing intact rings (Figure 3.12). The peak at m/z 1470.11 is only consistent with a structure in which Ala₃ is bridged to a residue N-terminal to Abu₈. Likewise the peaks at m/z 1029.05 and 1130.07 could only be formed if Ala₁₃ was bridged to a residue located C-terminal to its position. This indicated that the originally published tertiary structure of planosporicin was incorrect and that in fact Lan bridges connect Ala₃-Ala₇ and Ala₁₃-Ala₂₀, and not Ala₃-Ala₁₃ and Ala₇-Ala₁₈ as initially proposed.

Subsequently, Maffioli *et al.* (2009) repeated the NMR experiments of Castiglione *et al.* (2007) and performed additional MS/MS analysis on planosporicin produced by *P. sp.* DSM 14920. This new MS/MS data correlates well with that obtained in these studies for planosporicin produced by *P. alba* (Table 3.2). The results of this analysis confirmed beyond reasonable doubt that the peptide produced by *P. alba* was indeed planosporicin. The lantibiotic from both *Planomonospora* strains fragmented in precisely the same way, thus the same conclusions on the bridging pattern were reached independently.

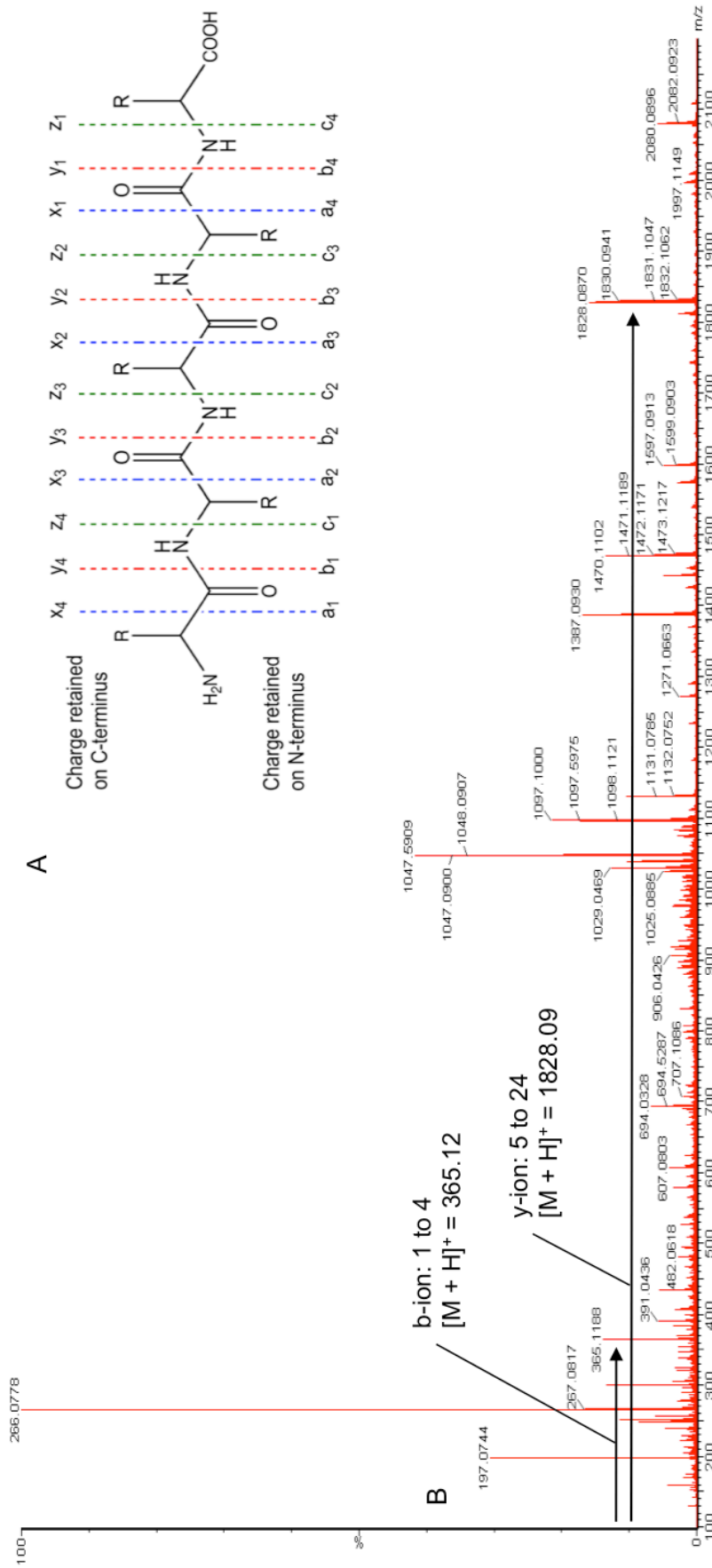


Figure 3.11 : Product ion scan.

A; Chemical structure of a peptide with the designation for fragment ions generated when the peptide backbone is fragmented using Roepstorff-Fohlmann-Biemann nomenclature. The lowest energy bond is the amide bond whose cleavage is marked in red, creating y-ions when the charge is retained by the carboxy-terminal fragment and b-ions when the charge is retained by the amino-terminal fragment. **B**; Product ion scan obtained from MS/MS analysis of the concentrated culture supernatant of *P. alba*. Each peak represents the singly charged ion $[M + H]^+$ with m/z labelled above. The y-axis shows the intensity normalised to the highest peak and m/z is displayed on the on the x-axis. Two peaks are labelled with the residues of planosporicin expected within that fragment.

<i>m/z</i>	fragment	<i>m/z</i> (Maffioli <i>et al.</i> 2009)
1828.09	5-24(SH)	1829.9
1387.08	9-24(SH)	1387.6
2076.15	1-23[-Ala(OH)SH]	2076.9
1470.11	8-24	1470.7
1029.05	13-24	
1130.07	12-24	1130.7
2079.08	2-24	
1996.13	3-24	

Table 3.2 : The major fragments from MS/MS of the 2191.8054 Da peptide planosporicin. The fragments of planosporicin produced by *P. alba* are listed as $[M + H]^+$ ions in the left hand column in descending order of intensity. In bold are fragments corresponding to the breakage of a peptide bond between lanthionine bridges (depicted in Figure 3.12). The right hand column lists the published masses as a result of MS/MS on planosporicin produced by *P. sp.* DSM 14920.

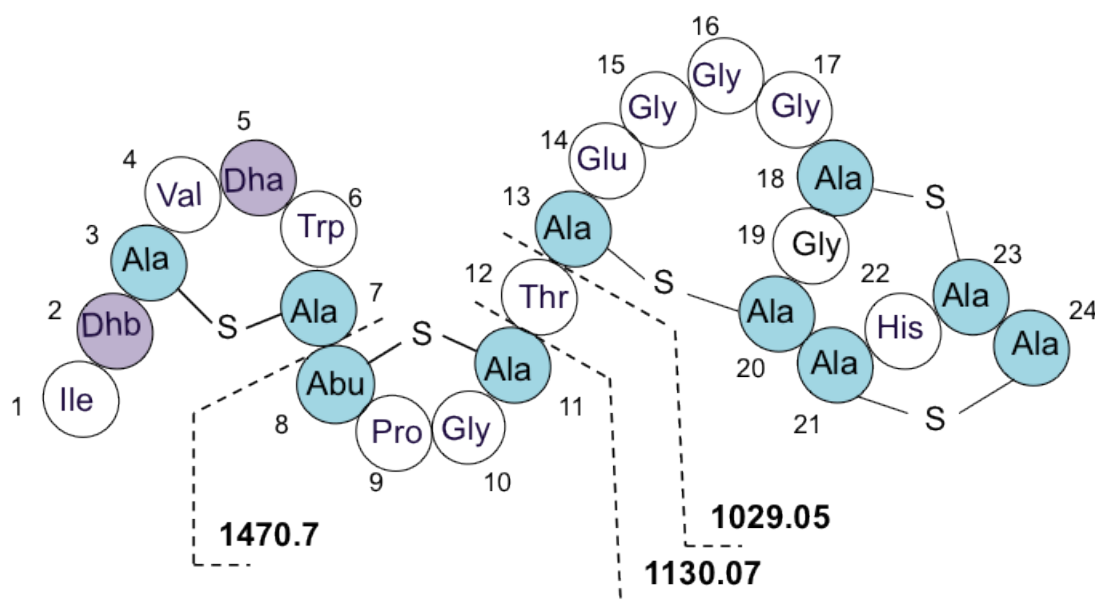


Figure 3.12 : Revised structure of planosporicin based on peptide fragments from MS/MS. The quoted masses correspond to $[M + H]^+$ y-ions extending from the point of cleavage (the dashed lines) to the C-terminus of the peptide.

3.5 Similarity to other lantibiotics

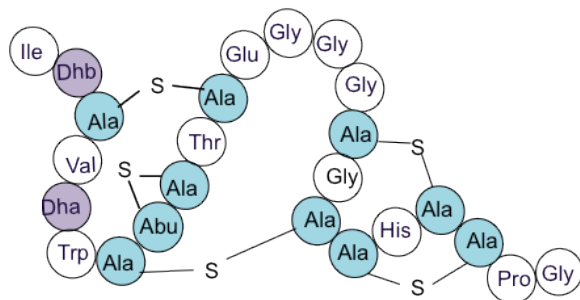
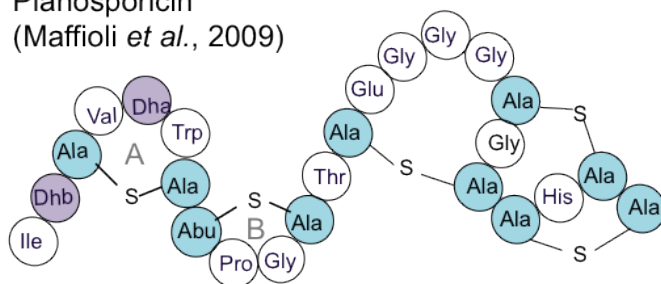
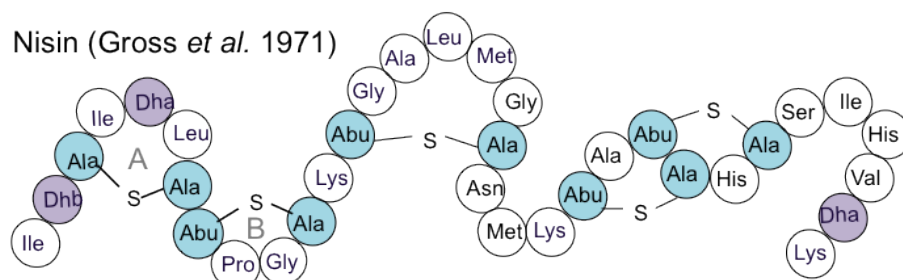
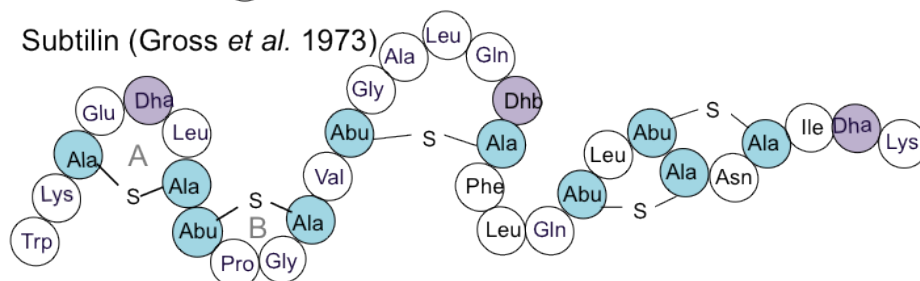
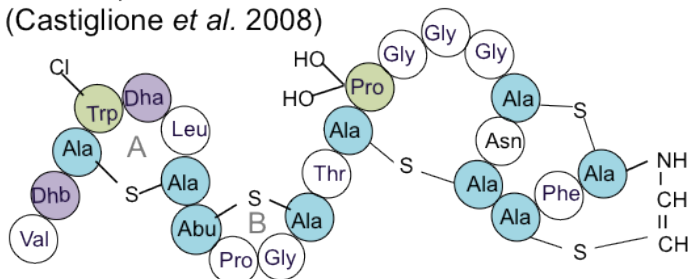
3.5.1 Nisin

The confirmation of the structure of planosporicin produced by *P. sp.* DSM 14920 (Maffioli *et al.* 2009) and *P. alba* (this work) enables a comparison to other lantibiotics. Nisin is the most extensively studied of all lantibiotics. The two common forms of nisin are nisin A and Z, which differ by a single amino acid at position 27, which is histidine in nisin A and asparagine in nisin Z (de Vos *et al.* 1993). More recently, two further variants, nisin Q and nisin F were also isolated from strains of *L. lactis* (Zendo *et al.* 2003; de Kwaadsteniet *et al.* 2008). Nisin Q differs at four positions (Val₁₅, Leu₂₁, Asn₂₇, and Val₃₀) from nisin A (Ala₁₅, Met₂₁, His₂₇, and Ile₃₀). Nisin F is most similar to nisin Z, differing at position 30, which is valine in nisin F, compared to isoleucine in nisin Z. Another two natural variants of nisin are produced by *Streptococcus uberis*; nisin U and nisin U2 (Wirawan *et al.* 2006). In total, the six nisin variants differ in up to 10 out of 34 amino acids present in nisin A. The structures of nisin A and Z were determined in 1971 and 1991, respectively (Gross and Morell 1971; Mulders *et al.* 1991). This was later confirmed by genetic analysis of the prepropeptide, *nisA* (Buchman *et al.* 1988).

The first identified form of nisin, nisin A, is used for comparison in this work. ClustalW2 was used to align the planosporicin prepropeptide sequence with nisin (NisA) to reveal regions of conserved sequence (Chenna *et al.* 2003). There are conserved regions in both the leader peptide and the propeptide (Figure 3.13). As noted in Section 3.4.1, the planosporicin leader peptide contains an FQLD motif similar to the FNLD motif first identified in nisin. The leader peptide is also similar to that of nisin in the region involved in cleavage to release the mature lantibiotic. The proline at position -2 and the alanine at position -4 are conserved in planosporicin. However the arginine at position -1 is not conserved in planosporicin, yet in nisin, mutation of this residue led to reduced processing of the peptide by the NisP leader peptidase (van der Meer *et al.* 1994). The most striking similarity is between the propeptides. The N-terminal 11 residues of planosporicin are identical to those of nisin except for valine at position 4 and tryptophan at position 6, which are isoleucine and leucine in nisin, respectively (Figure 3.13). This similarity in primary structure is reflected at the tertiary level. Nisin consists of N-terminal and C-terminal domains connected by methionine at position 21, acting as a flexible hinge region (Figure 3.14). This structure has subsequently been observed in several other lantibiotics, including subtilin produced by *Bacillus subtilis* (Figure 3.14). The N-terminal domain of these lantibiotics consists of two (Me)Lan rings termed the A and B rings. The revised structure of planosporicin has (Me)Lan rings in the same configuration at the N-terminus (Figure 3.14).



Figure 3.13 : An alignment of the planosporicin prepropeptide (PspA) with eight others. Namely nisin (NisA), epidermin (EpiA), microbisporicin (MibA), subtilin (SpaA), gallidermin (GdmA), actoracin (ActA), clausin (ClsA) and megateracin (MegA). Alignment constructed using ClustalW2 (Chenna *et al.* 2003). The partially conserved FNLD motif is in red. A putative conserved cleavage site is in green. The N-terminal 11 amino acids of the mature lantibiotic are in blue. The arrow indicates the putative cleavage site of the prepropeptide. The amino acid positions referred to in the text are numbered starting from this cleavage site.

Planosporicin (Castiglione *et al.*, 2007)Planosporicin
(Maffioli *et al.*, 2009)Nisin (Gross *et al.* 1971)Subtilin (Gross *et al.* 1973)Microbisporicin
(Castiglione *et al.* 2008)**Figure 3.14** : A comparison of lantibiotics with structural similarity to planosporicin.

Residues are; unmodified (white), dehydrated (purple), dehydrated and cyclised residues (blue), other modifications (green). Abbreviations refer to; amino butyric acid (Abu), dehydroalanine (Dha), lanthionine (Ala-S-Ala), dehydrobutyrine (Dhb) and methyl lanthionine (Abu-S-Ala). N-terminal A and B rings are labelled.

3.5.2 Microbisporicin

The species *Microbispora corallina* is currently represented by four strains; DSM 44681 and DSM 44682 (Nakajima *et al.* 1999), NRRL 30420 (Lee 2003) and ATCC PTA-5024 (Lazzarini *et al.* 2005). Two of these strains have been reported to make lantibiotic compounds. *M. corallina* NRRL 30420 was reported to make the compounds MF-BA-1768_{α1} and MF-BA-1768_{β1} (Lee 2003). *M. corallina* ATCC PTA-5024 was reported to produce a lantibiotic patented as antibiotic 107891 (Lazzarini *et al.* 2005). Three years later it was revealed that this compound was discovered during the same screen that led to the discovery of planosporicin (Castiglione *et al.* 2008). In this paper, microbisporicin was described as two factors; microbisporicin A1 and A2. The compounds produced by *M. corallina* NRRL 30420 and ATCC PTA-5024 were initially thought to be two different lantibiotics. However it is now known that the bioactive compound produced by both strains is the lantibiotic microbisporicin. In 2010 the *mib* gene cluster was identified from *M. corallina* sp. NRRL 30420 (Foulston and Bibb 2010). This revealed microbisporicin is a single prepeptide produced by a single biosynthetic gene cluster and that the composition of the microbisporicin complex varies with the culture conditions used.

The primary structures of planosporicin and microbisporicin are highly similar (Figure 3.13). As would be expected, the highest similarity is between the propeptide, as selection is maintained at the level of peptide structure, while the leader peptide has more freedom for diversification. Thus while the prepeptides share 57 % identity end-to-end, the propeptides share 75 % end-to-end identity. In total, there are only six residues which differ between the two 24 amino acid peptides. In the N-terminal domain, differences occur at positions 1, 4 and 6, which change from isoleucine, valine and tryptophan in planosporicin to valine, tryptophan and leucine in microbisporicin. In the C-terminal domain, differences occur at positions 14, 19 and 22 changing glutamate, glycine and histidine in planosporicin to proline, asparagine and phenylalanine in microbisporicin. Despite this similarity, the originally published structure of planosporicin and microbisporicin differed markedly. In contrast, the revised structure of planosporicin revealed similar ring topology to microbisporicin (Figure 3.14). Both lantibiotics have one MeLan bridge linking Abu₈-Ala₁₁. Both share the three Lan bridges linking Ala₃-Ala₇, Ala₁₃-Ala₂₀ and Ala₁₈-Ala₂₃. However whereas planosporicin possesses a fourth Lan bridge linking Ala₂₁-Ala₂₄, microbisporicin has an AviCys residue (Figure 3.14). Furthermore, both lantibiotics have dehydrated residues at positions two and five. However in microbisporicin the tryptophan at position 6 is chlorinated and the proline at position 14 is dihydroxylated. These unusual post-translational modifications are not observed in planosporicin as these are two of the six residues which are not conserved between the two lantibiotics.

Degradation of peptides commonly occurs through specific proteases which attack either the amino or carboxy termini. Most lantibiotics (such as planosporicin, microbisporicin and nisin in Figure 3.14) have a dehydrated residue (Dha/Dhb) as the second amino acid of the mature peptide to prevent degradation by amino proteases (van der Donk, personal communication). Likewise, the carboxy terminus is also protected through specific modifications. Planosporicin has a Lan bridge formed from Cys₂₄ following nucleophilic attack on the dehydrated Ser₂₁. This protects the lantibiotic from degradation by carboxypeptidases, although the carboxy terminus is still exposed. This C-terminal (Me)Lan is also seen in other lantibiotics, including salivaricin A (Ross *et al.* 1993) and mutacin II (Chen *et al.* 1998). In contrast, the AviCys modification in microbisporicin removes the carboxy terminus and introduces an alkene group, conferring additional rigidity that may confer additional protection from proteases.

3.5.3 Actoracin

The NCBI protein database was searched using BLASTP and the planosporicin propeptide (i.e. without the leader peptide) as a protein query (Altschul *et al.* 1990). This revealed that planosporicin shares 75 % identity end-to-end with a putative propeptide from *Actinomyces* sp. oral taxon 848 str. F0332. Importantly, eight out of eleven N-terminal residues are conserved including those that form the A and B rings of planosporicin. The structure of this putative lantibiotic was predicted on the basis of the known structure of microbisporicin and named 'actoracin' (Figure 3.15) (Foulston 2010). Actoracin was predicted to have the same bridging pattern as planosporicin combined with the AviCys of microbisporicin. This propeptide corresponds to the 24 C-terminal amino acids of an 86 residue peptide encoded by gene ID ZP_06162152, which presumably encodes the complete prepropeptide. This prepropeptide was named ActA, and alignments with other prepropeptides reveals the presence of an 'FNLD' type motif in the leader peptide, characteristic of type AI lantibiotics (Figure 3.13).

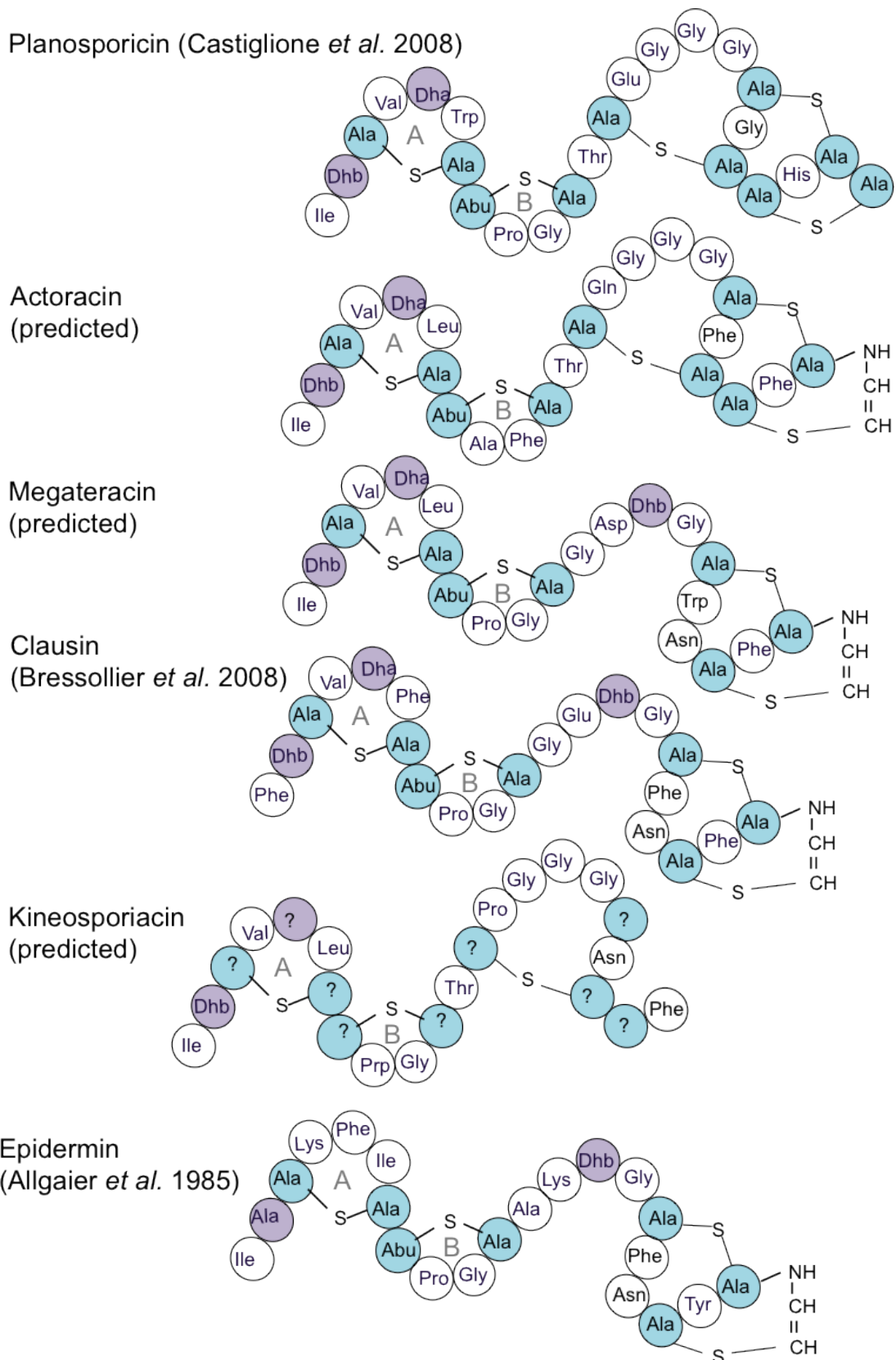


Figure 3.15 : A comparison of lantibiotics with structural similarity to planosporicin.

Residues are; unmodified (white), dehydrated (purple), dehydrated and cyclised residues (blue). Abbreviations refer to; amino butyric acid (Abu), dehydroalanine (Dha), lanthionine (Ala-S-Ala), dehydrobutyrine (Dhb) and methyl lanthionine (Abu-S-Ala). N-terminal A and B rings are labelled.

3.5.4 Clausin

Another putative peptide with a significant alignment to the planosporicin propeptide was identified in the genome sequence of *Bacillus megaterium* QM B1551. This 22 amino acid putative propeptide exhibited 63 % end-to-end identity with planosporicin. Although two residues shorter, there is a marked similarity at the N-terminus, where 10/11 residues are conserved (Figure 3.15). The only change is at position 6, which is leucine, as in microbisporicin and nisin. The structure of this putative lantibiotic, termed here megateracin, was modelled on the structure of epidermin, which is also a 22 amino acid lantibiotic with an N-terminus similar to planosporicin (Figure 3.15). In the NCBI database, this entry is annotated as a 'putative antibiotic protein' with locus tag BMQ_pBM70151, but in this work the prepropeptide is referred to as MegA.

A related organism, *Bacillus clausii*, has been marketed as Enterogermina® by Sanofi-Aventis as a probiotic to combat gastrointestinal disorders. The peptide responsible for its biological activity was recently identified, characterised and named clausin (Bressollier *et al.* 2007; Bouhss *et al.* 2009). Thus a different species of *Bacillus* produces a lantibiotic structurally similar to megateracin from *B. megaterium* (Figure 3.15). Clausin was recently reported to interact with lipid intermediates of the cell wall biosynthesis pathway by binding to the pyrophosphate moiety of lipid I and lipid II (Bouhss *et al.* 2009). In the *B. clausii* genome sequence, the *lanA* gene encoding the prepropeptide of clausin had not been identified by the genome annotation software. However, the primary structure of clausin from *B. clausii* can be deduced from its published tertiary structure (Bressollier *et al.* 2007; Bouhss *et al.* 2009). Previous analysis of the *B. clausii* KSM-K16 genome in the region 369440-3707200 in Artemis (Rutherford *et al.* 2000), identified a putative prepropeptide from 3706154-3706766 which was named ClsA and was predicted to produce a mature lantibiotic with the same sequence as clausin (Foulston 2010).

Thus it appears that two species of *Bacillus* produce two structurally similar lantibiotics. The tertiary structure of clausin corresponds to the *lanA* previously identified from the *B. clausii* genome (Foulston 2010). While clausin is not yet entered in the NCBI database, a BLASTP search identified a similar lantibiotic from *B. megaterium*. The two lantibiotics produced by the *Bacillus* sp. described in this Section differ at four out of 22 positions; 1, 6, 13 and 17. Figure 3.13 depicts the alignment of the MegA and ClsA prepropeptides with those of other lantibiotics. The presence of a 'FNLD' type motif suggests that these lantibiotics are of type AI. The strongest similarity in primary sequence is to the lantibiotic epidermin, produced by *Staphylococcus epidermidis* and its structural analogue gallidermin, produced by *Staphylococcus gallinarum*. This is reflected in the ring topology with three (Me)Lan bridges and an AviCys at the C-terminus (Figure 3.15). At the N-terminus, nine out of eleven amino acids match those of planosporicin, corresponding to

the A and B rings, whereas C-terminal similarity is reduced, sharing only one out of the three Lan bridges observed in planosporicin (Figure 3.15).

3.5.5 Kineosporiacin

Another planosporicin-like lantibiotic is produced by *Kineosporia* sp., an actinomycete of the *Kineosporiaceae* family. A patent was filed for this compound in which it is reported that Edman sequencing was blocked after the first residue, probably due to the presence of (Me)Lan bridges in the molecule (Shimizu and Masaki 2004). Thus the lantibiotic was chemically modified to introduce modifications that made the peptide accessible to Edman sequencing (Meyer *et al.* 1994). However this method did not discriminate between cysteine, serine and threonine residues incorporated into (Me)Lan bridges or between dehydrated serine and threonine residues, and therefore only the partial amino acid sequence of the mature compound was reported (Shimizu and Masaki 2004). Since no genome sequence for any *Kineosporia* sp. was available, it was not possible to identify the biosynthetic gene cluster. Instead, the structure of the putative lantibiotic was predicted by modelling the partial amino acid sequence on the structure of microbisporicin and was named kineosporiacin (Figure 3.15) (Foulston 2010). The N-terminus of this molecule is likely to form the A and B-rings found in planosporicin. Of the confirmed residues, only position 6 (a leucine) differs from planosporicin (a tryptophan).

3.6 Discussion

P. alba is an actinomycete that forms a branched, non-fragmenting substrate mycelium and aerial hyphae that bear clavate spore vesicles each containing a single spore (Mertz 1994). This Chapter confirmed that *P. alba* produces the pentacyclic lantibiotic planosporicin in which 50 % of the 24 amino acids are subject to post-translational modification. A recent *in silico* analysis compared all known lantibiotic structures (Rink *et al.* 2005). Although no strict sequence motifs were defined that govern the modification, statistical analysis demonstrates that dehydratable serines and threonines are more often flanked by hydrophobic than by hydrophilic amino acids. It was also indicated that serine residues escape dehydration more often than threonines. However in planosporicin, the only unmodified Ser/Thr residue is a threonine at position 12.

Lantibiotics are commonly classified on the basis of their structure and the type of enzymes that introduce post-translational modifications. Type AI lantibiotics are those which are elongated and use individual modification enzymes, whilst type BII are globular

and use a bifunctional enzyme to create (Me)Lan bridges. The revision of the planosporicin structure from that published by (Castiglione *et al.* 2007) to that of (Maffioli *et al.* 2009) involved both a shift of two amino acids and a reorganisation of two Lan bridges (Figure 3.14). This has implications for the classification of planosporicin. The original classification as a type B lantibiotic reflected the presence of two thioether crosslinks between Ala₃–Ala₁₁ and Ala₇–Ala₁₈ constraining the lantibiotic into a compact shape. The structural revision gives planosporicin an elongated shape, characteristic of type A lantibiotics. The change in bridging pattern resulted in an N-terminus consisting of rings Ala₃–Ala₇ and Abu₈–Ala₁₁ that resemble the A and B rings of nisin (Figure 3.14). Thus the overall elongated structure of planosporicin and its N-terminal similarity to nisin suggests that it is a type A lantibiotic.

The N-terminus of nisin is dominated by the presence of the A and B Me(Lan) rings (Figure 3.14). NMR studies have revealed these two rings wrap around the pyrophosphate group of lipid II (Hsu *et al.* 2004). The backbone of the nisin peptide chain forms five hydrogen bonds to the pyrophosphate. Thus there is no involvement of the specific side-chains of each amino acid component. The high degree of N-terminal homology of planosporicin to nisin implies a similar mode of action. Homology modelling revealed the A and B rings can accommodate a variety of side-chain compositions apart from at the conserved positions 3 and 7-11 (Hsu *et al.* 2004). Planosporicin possesses the conserved Ser₃ and LCTPGC motif at positions 6-11, indicating that it too can form a pyrophosphate cage around lipid II.

Analysis of other similar known and putative lantibiotics (actoracin, megateracin, clausin and kineosporiacin) with significant alignments to planosporicin indicated that these molecules may form a separate subclass of type AI lantibiotics. The type member of this group is microbisporicin produced by *M. corallina* (Castiglione *et al.* 2008; Foulston and Bibb 2010). Microbisporicin is distinguished from planosporicin by several additional post-translational modifications and a greater potency against Gram-positive bacterial pathogens. Microbisporicin is now in late pre-clinical trials with Sentinella Pharmaceuticals Inc. The increased activity of microbisporicin compared to planosporicin could be due to the presence of a chlorinated tryptophan, dihydroxyproline and AviCys group. It is possible that these additional modifications increase the affinity of microbisporicin to lipid II, increasing its efficacy as an antimicrobial compound. In particular, the difference in the C-termini of the two lantibiotics has an effect on the charge of the molecules. The free carboxy terminus of planosporicin confers an additional negative charge whereas the carboxy terminus of microbisporicin is used in the AviCys modification. This difference may be responsible for the difference in potency observed between the two molecules. Both lantibiotics are thought to target lipid II in the cell wall, so it can be envisaged that the

additional negative charge on planosporicin has the effect of repelling the peptide away from the negative cell membrane. In contrast, microbisporicin lacks this negative charge, enabling a greater affinity with the cell membrane, facilitating the binding of lipid II and conferring increased potency.

The confirmation of the structure of planosporicin presented in this Chapter gives scope for further work on the structure-activity relationships of this promising subclass of lantibiotics. This information adds to the wealth of knowledge about lantibiotics which can be used to define rules for the rational design of modified peptides (Rink *et al.* 2005). A comparison of the biosynthetic gene clusters for planosporicin and microbisporicin will likely reveal similarities and differences in their synthesis. Thus the identification of the planosporicin biosynthetic gene cluster is the focus of Chapter 4 and bioinformatic analysis of the constituent genes follows in Chapter 5.

3.7 Summary

P. alba sporulated best on ISP4 agar medium. SEM pictures revealed a cluster of clavate monospores on each sporangiophore.

P. alba produced planosporicin most consistently on AF/MS agar medium, resulting in a halo of growth inhibition of the assay organism *M. luteus*.

MALDI-ToF and ESI mass spectrometry analyses of culture supernatants of *P. alba* grown in AF/MS liquid medium revealed peaks corresponding in mass to planosporicin.

A hydrophobic polystyrene resin was used to concentrate and partially purify planosporicin from culture supernatant.

454 genome scanning revealed the correct primary structure of planosporicin.

MS/MS analysis predicted the subsequently revised and correct tertiary structure of planosporicin.

Planosporicin is a type A lantibiotic with close N-terminal similarity to nisin and microbisporicin.

Planosporicin is likely to bind the pyrophosphate moiety of lipid II through the A and B-rings at its N-terminus.

Chapter 4 : Identification of the planosporicin biosynthetic gene cluster

4.1 Introduction

The original aim of this project was to identify the gene cluster responsible for the biosynthesis of planosporicin. Planosporicin had previously been identified as a secondary metabolite of *P. sp.* DSM 14920 (Losi *et al.* 2004). Its *in vitro* activity against a range of Gram-positive human pathogens, activity at physiological pH and low toxicity *in murinae* implied a potential clinical role as a novel antimicrobial (Castiglione *et al.* 2007).

During the 'golden age' of antibiotics from the 1940s to the 1990s, the main method of discovery was through the detection of bioactive compounds in extracts from natural sources (Zerikly and Challis 2009). The discovery of planosporicin described in Chapter 3 is one example of antibiotic discovery using this approach. Nowadays, genome scanning can also be used to identify cryptic gene clusters in combination with genetic engineering techniques to define their products. This additional strategy takes advantage of new technologies to find novel metabolites. A recent example of a lantibiotic identified through genome mining is lichenicidin. The two-peptide lantibiotic was isolated from *Bacillus licheniformis* ATCC 14580 after identifying a gene with homology to *IanM* (Begley *et al.* 2009). The success of such techniques has led to the development of further tools for genome mining. BAGEL2 identifies putative bacteriocins using additional features such as the genomic context of biosynthetic genes in addition to their homology to characterised genes (de Jong *et al.* 2010).

Once the gene cluster has been identified, the next step is cloning and characterisation. As genetic engineering techniques began to complement screening of cell extracts, a number of new methods emerged during the 1980s. The generation of non-producing mutants allows characterisation of the biosynthetic pathway. In these mutants, fragments of the WT genome can be screened for their ability to complement this mutation and thus identify biosynthetic genes. This method was used in *Streptomyces griseus* to identify mutants blocked in the pathway of streptomycin biosynthesis (Hara and Beppu 1982). Alternatively, the resistance genes of a biosynthetic gene cluster could be cloned directly into a heterologous expression host and screened for resistance to the bioactive compound. This method was used to identify the tylosin-resistance genes *tlrA* and *tlrB*

from *Streptomyces fradiae* on a restriction fragment which conferred resistance to tylosin on *Streptomyces lividans* (Cox *et al.* 1986). Often these resistance genes can subsequently be used as a probe to locate nearby biosynthetic genes.

Many variations on a theme have developed since these early examples, each with their own advantages and disadvantages. Thus the method used is rationalised on a case-by-case basis depending on the information known about the system and the technology available. In 2003, the first lantibiotic gene cluster was cloned from a high GC organism; the cinnamycin biosynthetic gene cluster was cloned from *S. cinnamoneus* (Widdick *et al.* 2003). A radiolabelled probe was generated using sequence data for the part of the cinnamycin gene cluster from *Streptoverticillium griseoverticillatum*. Southern analyses identified a fragment of the gene cluster in *S. cinnamoneus* which was used as the template to design primers to clone overlapping DNA fragments that were subsequently cloned together. These fragments were sequenced and shown to correspond to the 17 kb cinnamycin biosynthetic gene cluster.

However this primer walking technique is time-consuming. A popular alternative is the generation of a stable large insert library from a complex genome. Several different cloning vectors have been used, but perhaps most popular are cosmids, fosmids, bacterial artificial chromosomes (BACs) and P1-derived artificial chromosomes (PACs). The choice of system often depends on the size of the biosynthetic gene cluster, which varies dramatically depending on the complexity of the biosynthetic pathway. The enzymatic requirements of secondary metabolites varies depending on their mechanism of synthesis. Non-ribosomal peptides and polyketides are synthesised by non-ribosomal peptide synthetases (NRPSs) and polyketide synthases (PKSs), respectively. Non-ribosomal peptides are biosynthesised through the sequential condensation of amino acid monomers, whereas polyketides are synthesised from the repetitive addition of two carbon ketide units derived from thioesters of short carboxylic acids such as acetate. NRPSs and PKSs are large, multi-enzymatic, multi-domain megasynthetases. Consequently the biosynthetic gene cluster can be over 100 kb in size. The 80.7 kb gene cluster for simocyclinone D8 was cloned from *Streptomyces antibioticus* Tü60 40 on six overlapping cosmids (Trefzer *et al.* 2002). Simocyclinone D8 consists of four different moieties, one of which is a polyketide. Consequently the gene cluster is highly complex with 49 protein coding sequences including a type II PKS. The large size of this cluster prevents cloning into a single cosmid. In contrast, the entire daptomycin NRPS was identified in a 128 kb region of the *S. roseosporus* genome cloned in a BAC through heterologous expression in *S. lividans* (Miao *et al.* 2005). Ribosomal peptides such as lantibiotics follow the canonical route of mRNA translation by ribosomes. Thus few lantibiotic clusters are larger than 20 kb. Consequently it is reasonable to predict that the

entire planosporicin gene cluster can be cloned within the ~40 kb insert DNA of a typical cosmid, as will be attempted in this Chapter.

Once a library is obtained, there are several screening methods. Oxazolomycin A is a peptide-polyketide hybrid compound. The oxazolomycin biosynthetic gene cluster was identified from *Streptomyces albus* JA3453 by screening a cosmid library with a degenerate probe (Zhao *et al.* 2006). The 79.5 kb cluster covered four overlapping cosmids, consisting of 20 open reading frames (ORFs) that encode both NRPSs and PKSs. However, screening libraries with degenerate probes is not the most reliable method, since non-specific hybridisation can yield false positives (Claesen and Bibb 2010). The simocyclinone cluster was identified through hybridisation with three different probes; two were fragments from the landomycin and urdamycin biosynthetic gene clusters and one was an amplified product of a likely simocyclinone biosynthetic gene from *S. antibioticus* (Trefzer *et al.* 2002).

Recently, the method of screening such libraries has been adapted to avoid the pitfalls of a degenerate probe. This is particularly useful when identifying an unknown gene cluster from an organism whose codon usage is also not well known. The microbisporicin gene cluster was identified from *M. corallina* through probing a cosmid library with a specific probe (Foulston and Bibb 2010). Information to amplify this probe was obtained from genome scanning data generated through next generation sequencing. Likewise, the cypemycin gene cluster was identified from *Streptomyces* sp. OH-4156 through similar methods. Solexa-based genome scanning revealed a contig containing the structural peptide that was used as a probe for cosmids containing the entire cypemycin gene cluster (Claesen and Bibb 2010). This Chapter describes the construction of a *P. alba* cosmid library which was probed for the planosporicin gene cluster using sequence information obtained through genome scanning.

4.2 454 sequencing

454 genome scanning of *P. alba* was initially carried out by the University of Liverpool. Although the data was of fairly poor quality, regions commonly found in many lantibiotic gene clusters were identified. The published planosporicin structure provided a basis for predicting the propeptide sequence, as described in Chapter 3. A search of the *P. alba* 454 contig database using a protein query (TBLASTN) revealed that the LanA prepropeptide was encoded fairly centrally within the 1729 bp contig01289. Analysis in Frameplot 2.3.2 predicted the likely protein-coding region of this high-GC content DNA (Bibb *et al.* 1984). Similar TBLASTN searches using SpaB (the dehydratase enzyme of the subtilin gene cluster) as a query revealed that the 5' end of *lanB* was also present on

contig01289, while the middle of *lanB* covers the entirety of the 559 bp contig03088. PCR was able to link contig01289 to contig03088, filling the 112 bp gap between the two contigs. These two contigs were used to design probes for the *psp* cluster as described in Section 4.4 (Figure 4.1).

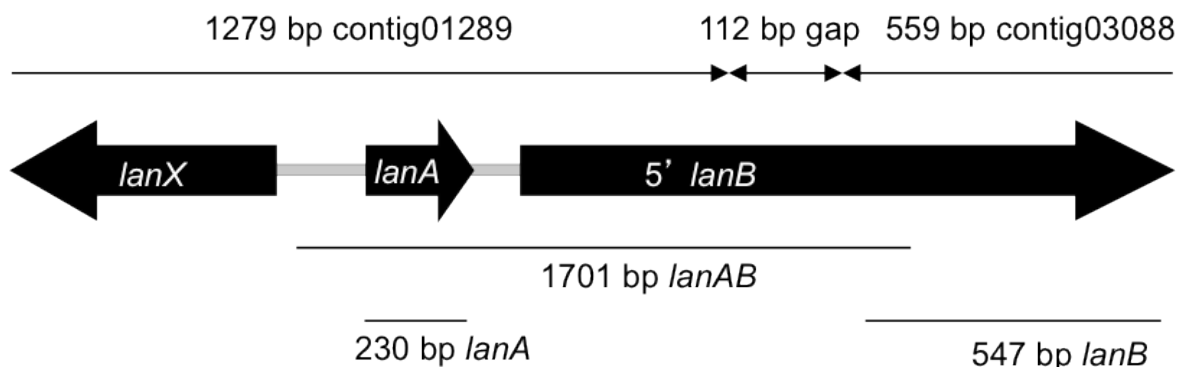


Figure 4.1 : A schematic depicting three genes proposed to be part of the planosporicin biosynthetic gene cluster annotated using the generic designation '*lan*'.

The sequence is covered by two contigs generated from 454 genome scanning data linked by PCR. Three probes; *lanA*, *lanAB* and *lanB* were amplified on the basis of this data.

Thus preliminary scanning identified small sections of the planosporicin gene cluster and PCR linked these sections together to form a larger, contiguous sequence. TBLASTN searches were also carried out for other common *lan* genes. However some are too similar to housekeeping genes to prove useful. For example *lanT* bears close homology to many ATP-binding cassette (ABC) transport systems that are difficult to assign to specific functions. However the ABC transporters conferring immunity typically contain an 'E-loop', instead of the well-conserved Q-loop in the nucleotide-binding domain which distinguishes them from others (Okuda *et al.* 2010). A TBLASTN search with MibE and MibF revealed that the 2552 bp contig00809 likely contained lantibiotic immunity genes. A gene similar to MibE was located centrally within this contig, and like MibE was predicted to contain six membrane spanning domains. The 5' terminus of a gene similar to MibF was also located on this contig. These may form the self-resistance mechanism to protect *P. alba* from its own lantibiotic. The observed similarity of planosporicin to microbisporicin discussed in Chapter 3 inspired TBLASTN searches for unusual enzymes found in the microbisporicin cluster, such as MibD, the enzyme catalysing S-[(Z)-2-aminovinyl]-D-cysteine formation and MibH, the FAD-dependent tryptophan halogenase. However, no hits were found in either case.

Subsequently, 454 next generation sequencing was repeated on the Norwich Research Park at The Genome Analysis Centre. *P. alba* genomic DNA was fragmented into 400-600 bp fragments, quantified, amplified by emulsion PCR and sequenced on a quarter of a plate. This generated 72 Mb of data which was assembled into 1017 contigs. These contigs contain 7.3 Mb of data, so the reads gave approximately 10-fold coverage of the genome. The Newbler assembler predicted a genome size of 9.1 Mb, but as 99 % of reads were assembled into contigs, it seems likely that the *P. alba* genome is approximately 7.5 Mb. A full statistical analysis of this assembly is given in Chapter 2. In brief, there were 944 contigs at least 500 bp long, the largest of which was 140 kb. Half of the entire assembly was contained in contigs or scaffolds equal to or larger than 17 kb. The mean GC content of the assembled data was 73.9 mol%GC (Figure 4.2). All contigs were stored in a database from which TBLASTN searches identified regions common in most lantibiotic clusters. In total, three large contigs: 00232 (~8 kb), 00214 (~8.5 kb) and 00436 (~4 kb) appeared to cover the entire *psp* cluster with just small gaps of <100 bp between the contigs.

4.3 Cosmid library preparation

A cosmid library was constructed as described in Chapter 2. A cosmid is essentially a plasmid that contains *cos* sites from lambda phage. Cosmids frequently also contain a bacterial origin of replication, an antibiotic resistance gene and a cloning site. Thus they can replicate as plasmids in bacteria, and transfected cells can be identified through their ability to survive on the appropriate antibiotic selection. The presence of *cos* sites enables packaging of the linear cosmid into phage heads. This packaging step selects for large inserts, providing a clear advantage over BAC and PAC vectors. Once packaged the cosmid can be stably introduced into cells by transduction. Consequently, cosmids are able to contain 30 to 50 kb of DNA, while most plasmid cloning vectors are generally able to carry only up to 20 kb.

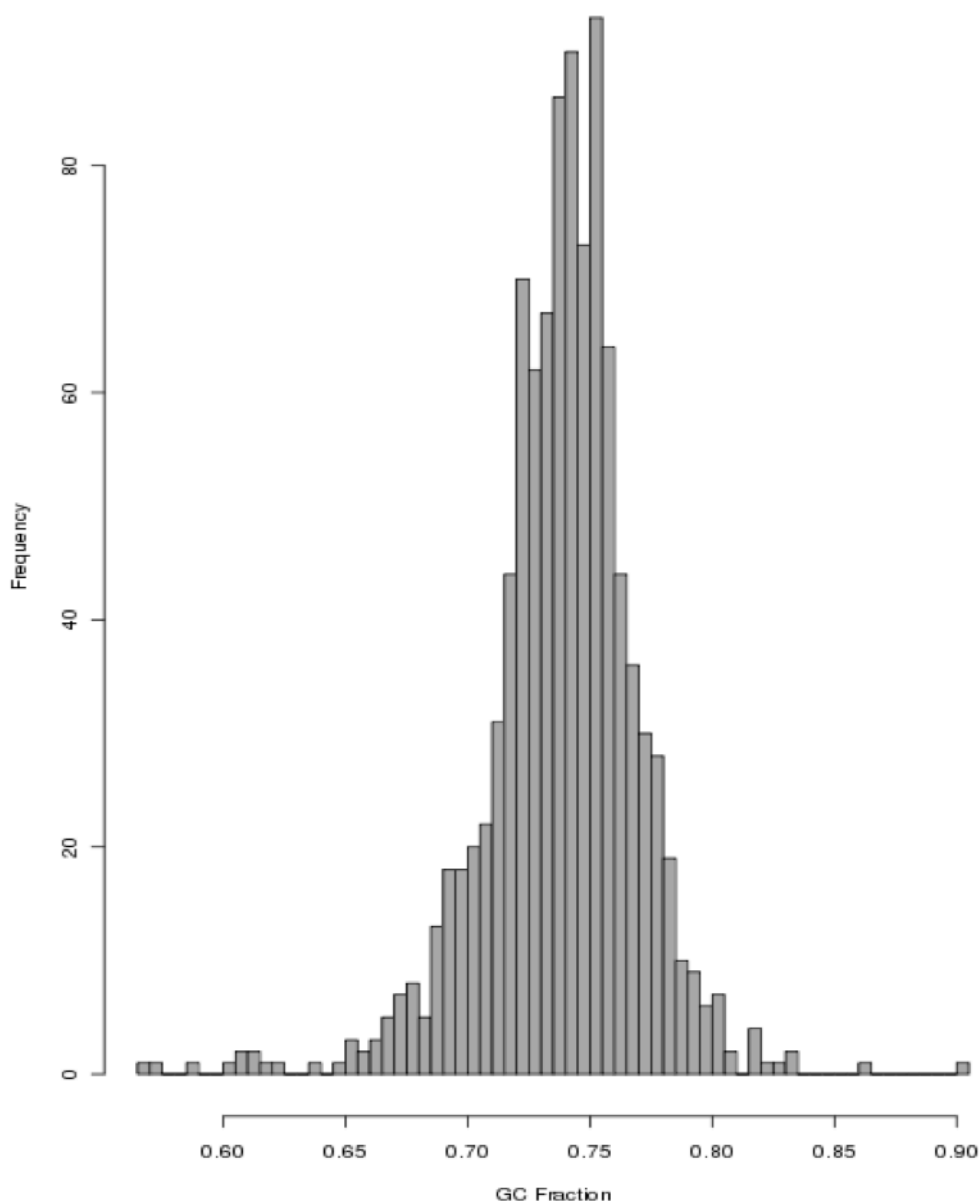


Figure 4.2 : Histogram of the GC content for the assembled contigs from 454 sequencing of *P. alba* genomic DNA.

4.3.1 Genomic DNA

The cosmid library used *P. alba* genomic DNA (gDNA) as insert DNA. A critical stage is the isolation of high molecular weight gDNA at a high concentration as described in Chapter 2.11. The extracted *P. alba* gDNA was subjected to pulsed field gel electrophoresis and the size was estimated to range from 100-300 kb (Figure 4.3). A partial restriction digest reduced the *P. alba* gDNA to 30-50 kb, an optimal size for phage packaging (Figure 4.3). Subsequent dephosphorylation of the digested fragments reduced the risk of cloning two non-contiguous segments of genomic DNA into the same cosmid vector.

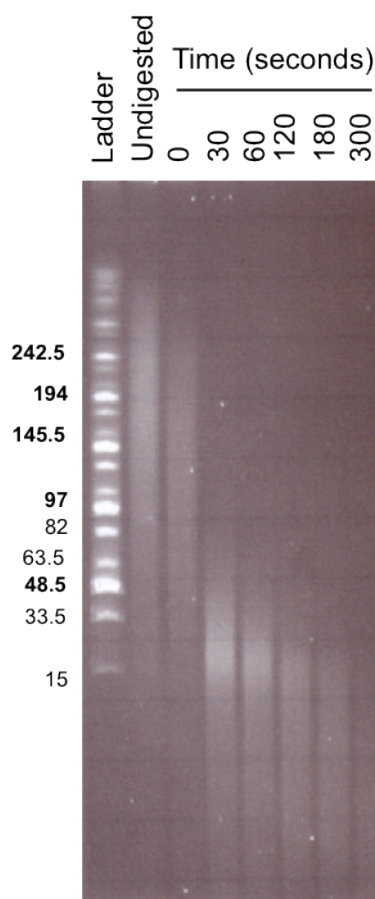


Figure 4.3 : High molecular weight *P. alba* gDNA analysed by PFGE.

gDNA was subjected to partial digestion with *Sau3AI*. Increasing incubation time resulted in smaller fragments of gDNA. A negative control of gDNA with no *Sau3AI* treatment (undigested) was also run. Size estimation is given by MidRange I PFG Marker (NEB) with band size in kb indicated next to the gel (bold figures represent major bands).

4.3.2 Cosmid library construction

The construction of the cosmid library was a multi-step cloning process and is described in Chapter 2 and depicted as a schematic in Figure 4.4. The first step involved digesting the vector DNA to generate two vector arms. This was accomplished by digestion with two different enzymes. First the cosmid vector SuperCosI (Stratagene) was linearised by digestion with the single cutter *XbaI* and the linear vector was dephosphorylated. Then the vector was digested with a second enzyme, *BamHI*, which is isocaudomeric to the *Sau3AI* used for the partial digestion of gDNA as described in 4.3.1. The identical overhangs generated by *BamHI* and *Sau3AI* allowed the insertion of the digested gDNA into the vector. The ligation product was packaged *in vitro* into phage heads using packaging extracts derived from *E. coli* lysogens. The packaged cosmids were titred, then scaled-up to transfect the appropriate *E. coli* host to generate a library of several thousand colonies.

Of these, 3072 colonies were robotically arrayed into eight 384 well plates to create the cosmid library. If the *P. alba* genome is approximately 7.5 Mb in size, and the average cosmid insert is 40 kb then a library of 3072 clones gives over 16-fold coverage of the genome. Thus we can expect to find approximately 16 library clones containing at least part of the planosporicin biosynthetic gene cluster.

4.4 Clone identification

4.4.1 Probe preparation

Genome scanning data described in Section 4.2 was used to generate probes to hybridise to the cosmid library. Contig01289 and contig03088 were linked by PCR, creating a 2.4 kb fragment containing the planosporicin *lanA* and the 5' terminus of its *lanB*. The sequence of this fragment was used to design three pairs of primers to amplify probes spanning 230 bp of *lanA*, 547 bp of *lanB* and 1701 bp of *lanAB* (Figure 4.1). Each probe was initially amplified from *P. alba* gDNA using the Expand High Fidelity PCR System (Roche). This kit contains *Taq* DNA Polymerase and a polymerase with proofreading activity. *Taq* adds a single adenine nucleotide to the 3' end of amplified DNA fragments, creating an A-tail. This A-tail was used to ligate the amplified fragment into the commercial vector pGEM-T (Promega). pGEM-T is a linear construct due to digestion with *EcoRV* and has 3' terminal thymidine overhangs added at both termini. The *EcoRV* digestion cleaved *lacZ* α , encoding the α fragment of the LacZ protein which is required to complement *in vivo* a defective C-terminal segment of β -galactosidase (LacZ). The 3' terminal thymidines provided a compatible overhang for the PCR products. The products of the ligation reaction were introduced into *E. coli* DH5 α by transformation and plated onto LB media containing 100 μ g/ml carbenicillin, 0.5 mM Isopropyl β -D-1-thiogalactopyranoside (IPTG) and 80 μ g/ml X-gal (galactose linked to a substituted indole). IPTG functions as the inducer of the Lac operon. If the pGEM-T vector recircularises, a functional LacZ α fragment is formed that combines with the Ω subunit of the LacZ protein encoded in the *E. coli* chromosome to form a functional β -galactosidase enzyme. The colourless X-gal would then be metabolised by the β -galactosidase to form 5-bromo-4-chloro-indoxyl which is spontaneously oxidized to 5,5'-dibromo-4,4'-dichloro-indigo, an insoluble blue pigment, yielding a blue colony. In contrast, white colonies would indicate the presence of an insert ligated within the pGEM-T vector, preventing expression of a functional β -galactosidase enzyme. Several white colonies from each ligation were sequenced to confirm the lack of mutations potentially introduced by PCR. Each probe was amplified with the Expand High Fidelity PCR System, using a sequenced plasmid as template.

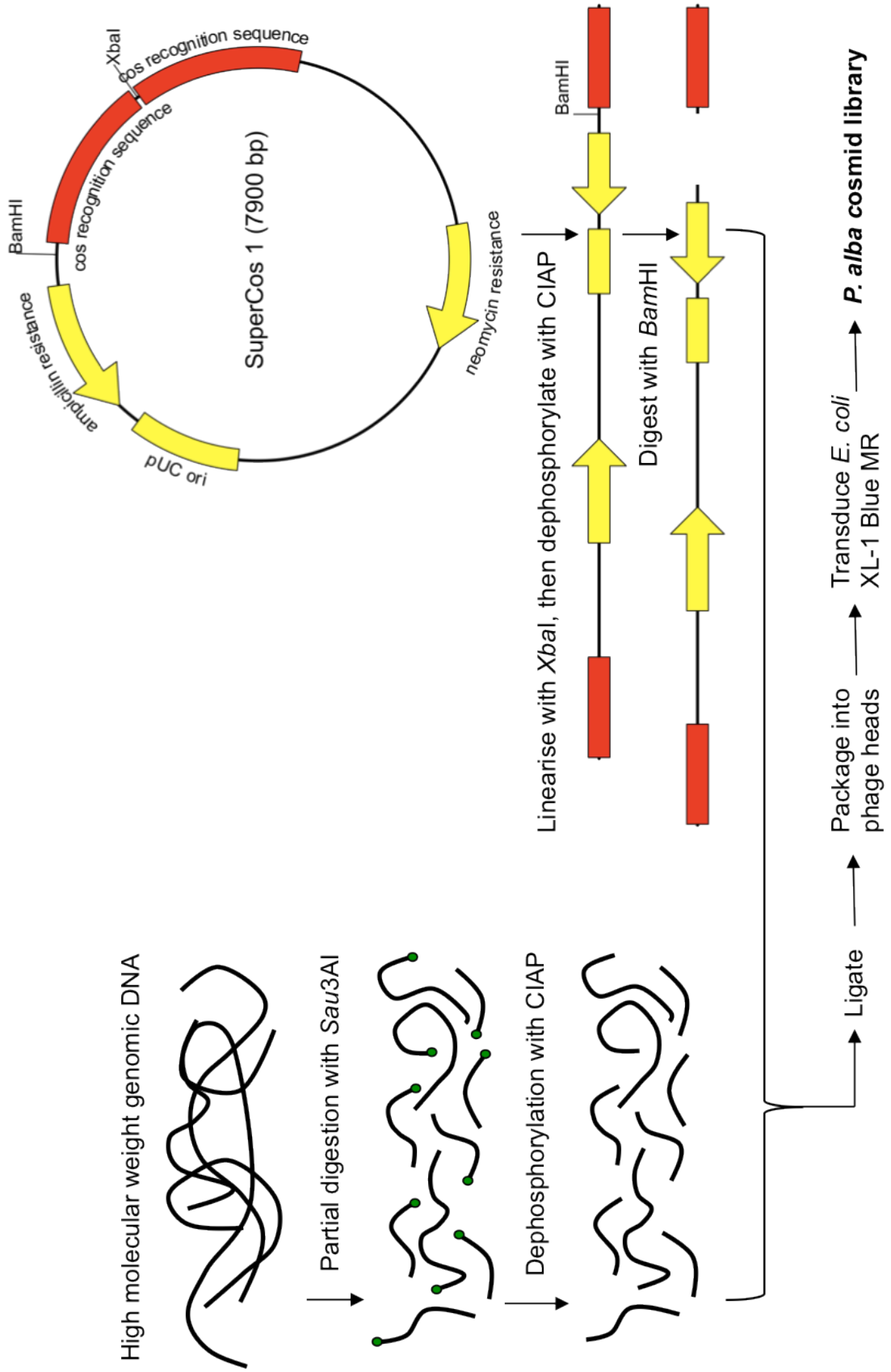


Figure 4.4 : Schematic of cosmid library construction.

4.4.2 Hybridisation to cosmid library

The 3072 clones of the cosmid library were transferred from eight different 384 well plates to a nylon membrane. Each cosmid was spotted twice in a characteristic duplex pattern to enable identification of which plate each hybridising clone came from. The three PCR products from Section 4.4.1 were randomly labelled with ^{32}P - α -dCTP to enable detection through photostimulated luminescence (PSL). Each labelled probe was hybridised individually with a separate copy of the cosmid library that was then exposed to a phosphor imaging plate. The radioactive signal was stored in the phosphor plate then later read by a phosphorimager. PSL releases the stored energy through stimulation with visible light, to produce a luminescent signal read by the phosphorimager, revealing a highly sensitive image of the original pattern of radiation.

Eight clones gave a distinctive double spot pattern with all three probes (Figure 4.5). A further five clones only hybridised well to a couple of the probes, implying they only contained part of the gene cluster. Thus probing the library with three probes located within just 1830 bp of the *P. alba* genome revealed 13 clones putatively containing at least part of the planosporicin biosynthetic gene cluster.

4.5 Clones picked from the library

4.5.1 PCR on *P. alba*

The eight positive clones identified in 4.4.2 were picked from the library and grown to amplify the cosmid DNA for isolation as described in Chapter 2. The presence of the planosporicin biosynthetic genes was confirmed by PCR using primers to amplify *lanA* (1289FlanA and 1289RlanA) and *lanB* (3088F2 and 3088R2). Additionally, the primers FlanEF and RlanEF were designed to anneal within contig00809 to amplify the putative immunity gene *lanE*. The 221 bp fragment was successfully amplified from all eight cosmids, indicating that it most likely corresponds to an ABC transporter in the planosporicin biosynthetic gene cluster. Long-range PCR was also attempted. A primer annealing to the cosmid backbone was used in combination with a primer annealing to a lantibiotic gene. Combinations of end_F and end_R with primers for *lanA* or *lanE* all failed to amplify a product. Likewise attempts to link *lanA* with *lanE* by long-range PCR also failed. It is likely these genes are located too distantly from each other to allow the amplification of a PCR product.

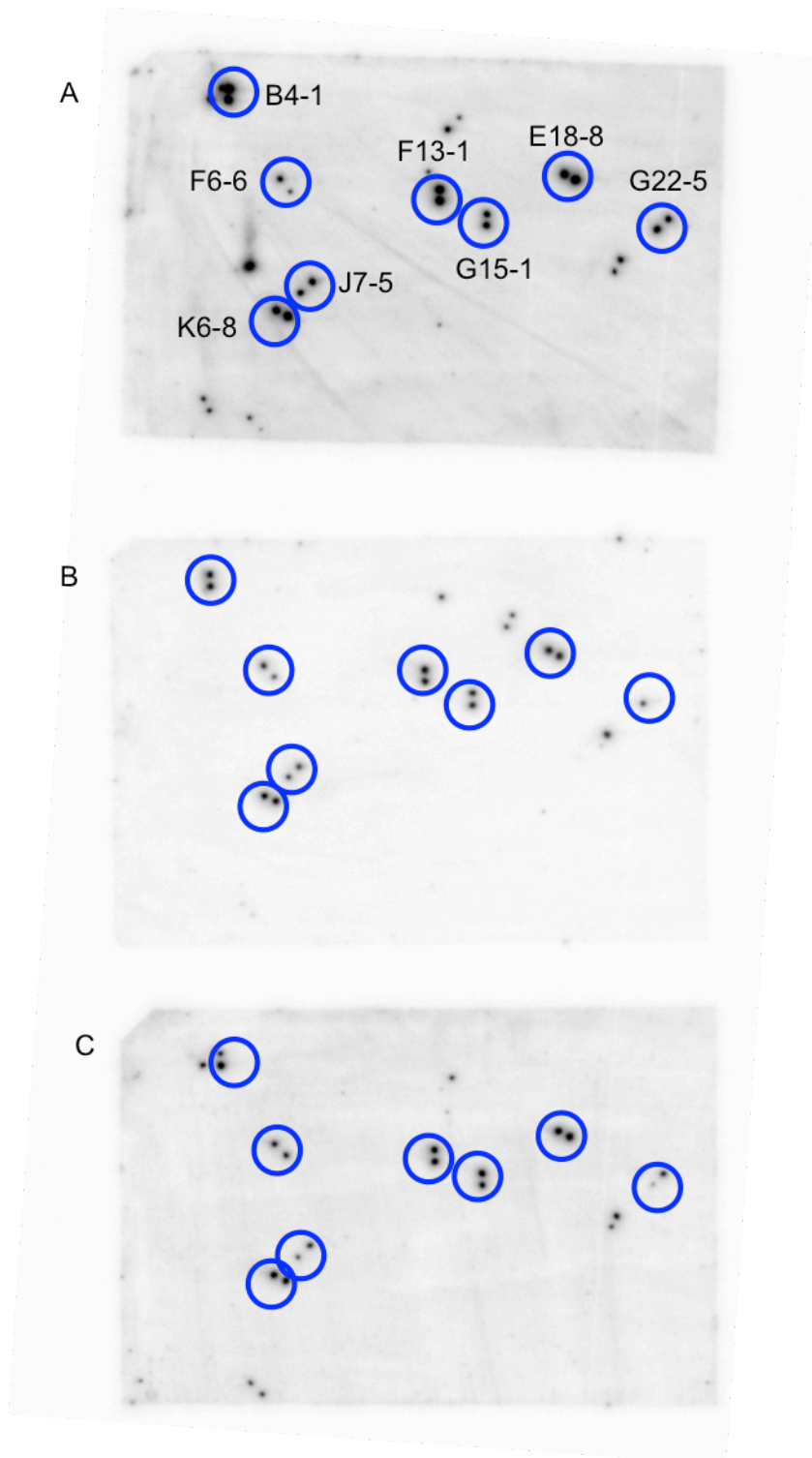


Figure 4.5 : Hybridisation of the *P. alba* cosmid library with three different probes. Cosmids taken forward for further investigation are ringed and labelled. **A**; Probe for 1701 bp spanning *lanA* and *lanB*. **B**; Probe for 230 bp of *lanA*. **C**; Probe for 547 bp of *lanB*.

4.5.2 End-sequencing of *P. alba* cosmids

Primers annealing to the cosmid vector backbone were used to sequence into the insert DNA of all eight cosmids that might have contained the entire planosporicin biosynthetic gene cluster. Primers end_F and end_R from table 2.6 were used to obtain ~800 bp of sequence data from each cosmid. This end-sequence data was used in a BLASTN search of the *P. alba* 454 database to identify contigs corresponding to the ends of each cosmid. Analysis of these contigs in Frameplot enabled the identification of likely ORFs from which the translated protein was used in a BLASTP search of the NCBI protein database to determine if lantibiotic genes are located close to the ends of the cosmid inserts. The absence of likely lantibiotic biosynthetic genes would imply that the entire planosporicin gene cluster is centrally located within the cosmid insert. End-sequence data from cosmids G15-1, J7-5 and K6-8 had significant alignment to either a permease or ATP-binding component of an ABC-type transporter. Based on studies with microbisporicin, it seemed likely that the planosporicin biosynthetic gene cluster would contain at least two ABC-type transporters; one to export the lantibiotic after synthesis and one that functioned as an immunity mechanism. For this reason these cosmid clones may not have contained the entire gene cluster. In contrast, the sequence at one end of cosmid B4-1 had homology to a *Streptosporangium roseum* dihydrofolate reductase, while the other end had homology to an *Actinosynnema mirum* methyl-accepting chemotaxis protein. Neither of these proteins have an obvious role in lantibiotic biosynthesis so this cosmid was a strong candidate for containing the entire gene cluster.

4.5.3 Restriction digest on *P. alba* cosmids

Further information as to the insert DNA contained in each cosmid was gleaned from restriction digests with different enzymes. All eight cosmids identified in Section 4.4.2 displayed some common fragments when digested with *Bam*HI or *Pst*I (Figure 4.6). Cosmids F13-1 and F6-6 appeared almost identical according to the size of the restriction fragments generated. Cosmids B4-1, G22-5 and E18-8 also had strikingly similar band patterns. Thus cosmids B4-1 and F13-1 were chosen for sequencing. These cosmids contained all planosporicin biosynthetic genes identified so far (Section 4.5.1), but none were present at the ends of the cosmid (Section 4.5.2). Despite these similarities, they gave different restriction digest patterns. The total size of each cosmid was estimated by addition of all the digest fragments. Both cosmids were approximately 40 kb in size, giving an estimated insert size of 32 kb, likely large enough to contain the entire biosynthetic gene cluster. For the rest of this thesis, cosmid B4-1 is referred to as plJ12321 and cosmid F13-1 is referred to as plJ12322.

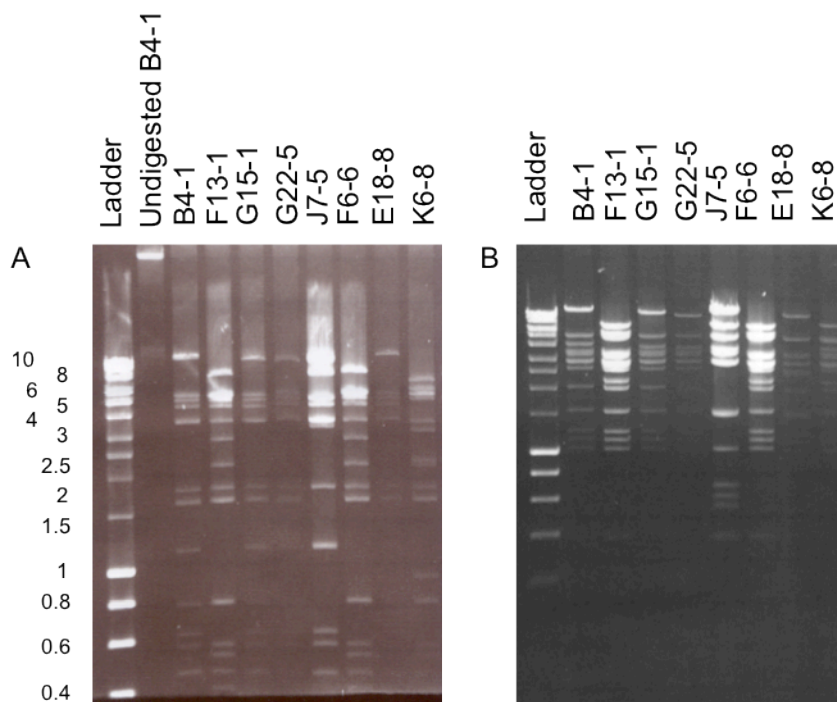


Figure 4.6 : Restriction digest patterns of different *P. alba* cosmids.

A; Digest with *Bam*HI. **B**; Digest with *Nco*I. Size estimation is given by Hyperladder I (Bioline) with band size in kb indicated next to the gel.

4.6 Cosmid sequencing and annotation

4.6.1 Sequencing

P. alba cosmids B4-1 (pIJ12321) and F13-1 (pIJ12322) were sequenced using Sanger sequencing at the DNA sequencing facility at the University of Cambridge. The insert DNA in pIJ12321 was 37760 bp, while pIJ12322 contained 37916 bp. Although very similar in size, the segment of cloned *P. alba* gDNA differed slightly between the two cosmids. pIJ12321 has the 1830 bp of sequence used to identify the clone located centrally within the insert DNA, whereas in pIJ12322 this sequence is located towards the right hand border of the insert DNA with the vector backbone. ClustalW2 was used to align the insert gDNA from pIJ12321 and pIJ12322 (Chenna *et al.* 2003). This identified one discrepancy, an additional cytosine residue in the pIJ12322 sequence. This potential frameshift mutation would have destroyed a stop codon present in the pIJ12321 sequence. Correspondence with the Cambridge Sequencing facility confirmed that this additional cytosine was an error during the assembly of the pIJ12322 sequence in Consed (Gordon *et al.* 1998).

4.6.2 Annotation

The full insert DNA of pIJ12321 was annotated using Artemis software (Rutherford *et al.* 2000). Instrumental to the correct assignment of ORFs was analysis of the GC content. The high-GC content of actinomycete genomes skews the usage of codons. The triplet code allows most freedom at the third codon position. Consequently the first, second and third nucleotide positions of codons show distinct differences in GC content, changing from intermediate to low to high GC content (Bibb *et al.* 1984). The GC-frameplot tool within the Artemis program tracks the GC content at each position within a moving window of a stipulated number of codons. Analysis of the GC content at each position reveals both where ORFs may be situated and their directionality (Figure 4.7). The start site of each ORF was adjusted to an ATG or GTG with a potential ribosome-binding site (RBS) nearby. The sequence GGAGG was used as the ideal RBS, located approximately 5 to 15 bp upstream of the start codon (Kieser *et al.* 2000). Thus the final annotation uses the start codon with a putative RBS upstream which creates the largest ORF.

The protein sequences from putative ORFs were compared to the NCBI protein database using the BLASTP search program (Altschul *et al.* 1990). Proteins with a high percentage similarity to those encoded by the planosporicin gene cluster were used to assign a putative function to each ORF.

It is interesting to note that a number of genes appear translationally coupled due to overlapping start and stop codons. This is common in gene clusters as a method of coordinating regulation. Translational coupling is common where the stoichiometric ratio of protein expression is essential for optimal functionality. Elsewhere in the cluster, there are larger gaps between ORFs. This gives scope for regulation of transcription through RNA secondary structure. Hairpin loops in mRNA may act as transcription terminators. These alter the stoichiometry of protein expression; coding regions before the stem loop are transcribed at a higher rate compared to those after the stem loop. This may be crucial when an enzyme and its substrate are expressed from the same operon. One enzyme can process many substrates so the stoichiometric ratio needs to be altered accordingly to optimise efficiency.

4.7 *P. sp.* DSM 14920 cosmids

4.7.1 PCR on *P. sp.* DSM 14920 cosmids

The NAICONS consortium constructed a second *Planomonospora* cosmid library with gDNA from *P. sp.* DSM 14920. In this strain, planosporicin is referred to as NAI-97. NAICONS identified four partially overlapping cosmids that appeared to contain planosporicin biosynthetic genes. In this work, PCR using cosmids 4B8 and 9A7 as templates amplified all three *lan* genes; *lanA*, *lanB* and *lanE*, while cosmid 4H8 appeared to lack *lanE* and cosmid 7C2 to lack *lanB* (Figure 4.8).

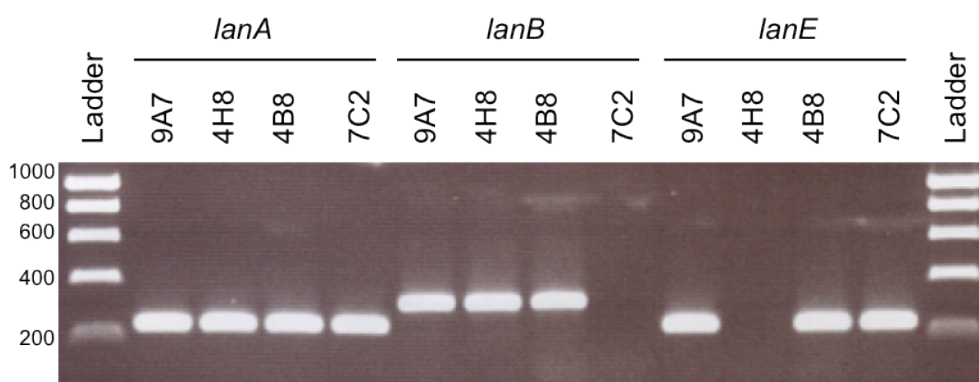


Figure 4.8 : PCR analysis to determine which *P. DSM 14920* cosmids were likely to contain the entire planosporicin gene cluster.

Cosmids 9A7, 4H8, 4B8 and 9A7 were used as templates to amplify *lanA* (1289FlanA and 1289RlanA; 230 bp), *lanB* (3088F2 and 3088R2; 303 bp) and *lanE* (FlanEF and RlanEF; 221 bp) from the *psp* gene cluster. PCR products were run on a 1 % agarose gel by electrophoresis. The ladder is Hyperladder I (Bioline) with band sizes annotated in bp.

The sequence and annotation of pIJ12321, described in detail in Chapter 5, depicts the organisation of the *psp* cluster in *P. alba*. Genome scanning of *P. sp.* DSM 14920 identified the planosporicin prepropeptide and analysis of a few kilobases upstream and downstream revealed a gene organisation very similar to that of *P. alba* (M. Sosio, personal communication). To investigate if sequences flanking the biosynthetic gene cluster in each strain show synteny, primers were designed to amplify across the predicted left and right borders of the *psp* cluster. These primers amplified a fragment of the same size whether *P. alba* or *P. sp.* DSM 14920 DNA was used as a template (Figure 4.9). Thus gene order appears to be conserved between the *psp* clusters from *P. alba* and *P. sp.* DSM 14920.

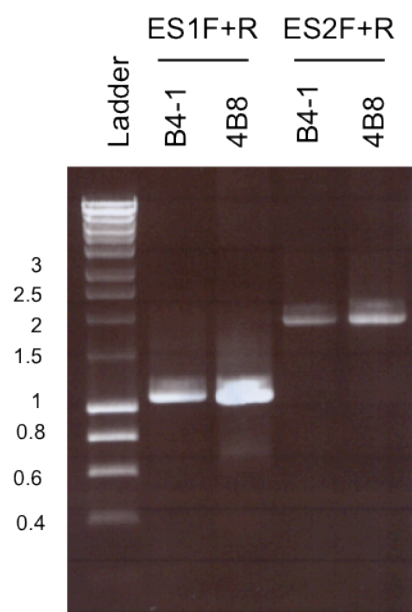


Figure 4.9 : PCR analysis to check the similarity of *P. DSM 14920* cosmid 4B8 to *P. alba* cosmid B4-1, both encoding the planosporicin gene cluster.

Cosmid DNA was used as a template to amplify regions across the borders of the planosporicin gene cluster. LongAmp polymerase (NEB) was used to amplify two regions. ES1F and ES1R amplify 1073 bp across the left hand border. ES2F and ES2R amplify 1883 bp across the right hand border. PCR products were run on a 1 % agarose gel by electrophoresis. The ladder is Hyperladder I (Bioline) with band sizes annotated in kb.

4.7.2 Restriction digest on *P. sp. DSM 14920* cosmids

Interestingly, restriction digests comparing *P. alba* cosmids with those from *P. sp. DSM 14920* appeared quite different. Figure 4.10 depicts the band pattern created by a *Bam*HI or *Nco*I digest. Several restriction fragments are common across all eight *P. alba* cosmids. The majority of these are absent from the digest pattern of the *P. sp. DSM 14920* cosmids. This implies that although the amino acid sequences of the encoded proteins is conserved, there is considerable variation at the nucleotide level in the planosporicin biosynthetic gene clusters from the two *Planomonospora* species.

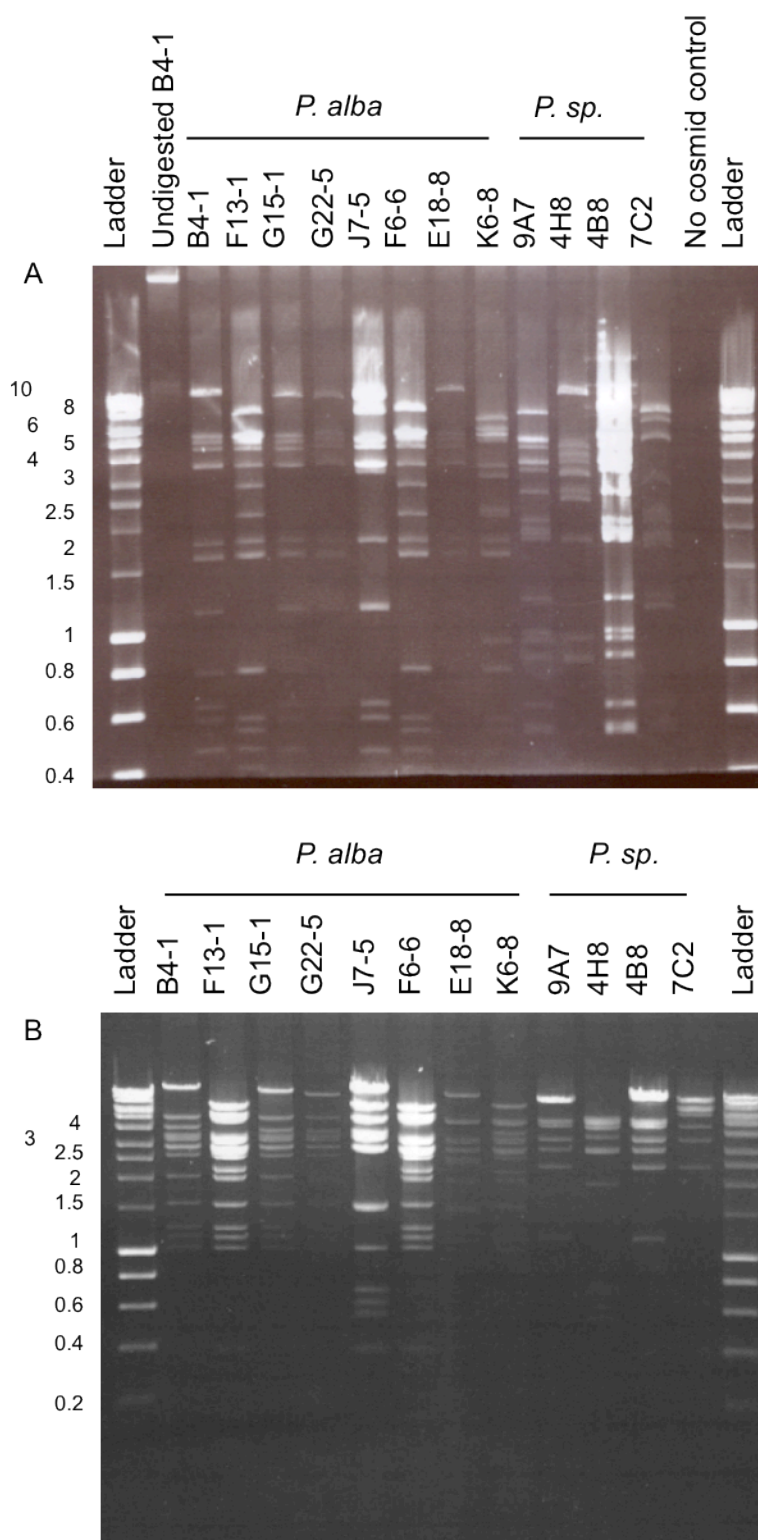


Figure 4.10 : Restriction digest patterns of eight *P. alba* cosmids and four *P. sp.* DSM 14920 cosmids with two different enzymes.

A; Digest with *Bam*HI. **B**; Digest with *Nco*I. Restriction fragments were run on a 1 % agarose gel by electrophoresis. Size estimation is given by Hyperladder I (Bioline) with band size in kb indicated next to the gel.

4.7.3 End-sequencing of *P. sp.* DSM 14920 cosmids

The *P. sp.* DSM 14920 cosmid library was made with a derivative of the SuperCosI vector used for the *P. alba* cosmid library. Consequently the same primers; end_F and end_R were used to sequence into the ends of the insert DNA. Cosmid 7C2 yielded end-sequence data, which a BLASTN search of the *P. alba* 454 database yielded contig01289 as a top hit. The presence of *lanA* and *lanB* at the end of this insert implied that part of the lantibiotic biosynthetic gene cluster would be missing from the cosmid. Cosmids 4B8 and 9A7 gave end-sequences which a BLASTP search of the NCBI protein database revealed similarity to *Streptosporangium roseum* genes for primary metabolism, implying the entire cluster may be present within these cosmids. For the rest of this work, cosmid 4B8 is referred to as pIJ12325 and cosmid F13-1 is referred to as pIJ12326.

4.8 Discussion

This Chapter describes how genome scanning through 454 sequencing created contigs that enabled BLAST-type searches for genes similar to those commonly found in lantibiotic gene clusters. This information allowed the amplification of probes which were hybridised against a *P. alba* cosmid library. From the selection of positive hits, a combination of PCR, restriction digests and end-sequencing were used to deduce which cosmids likely contained the entire cluster which were then sent for sequencing. The next Chapter continues along this theme with the bioinformatic characterisation of the cluster and a full discussion of the likely functions of individual genes.

Further work has fully justified the methodology chosen to identify the planosporicin gene cluster. An activity based screen could have been performed by integrating the *P. alba* cosmid library into a heterologous host. However, choice of host would have been difficult, as there would have been no guarantee that the cluster would have been expressed at a level that allowed planosporicin detection. Likewise, use of heterologous expression to identify which cosmids out of those which give a positive hit after hybridisation with a radioactive probe would likely have proven a futile exercise, as Chapter 6 details how attempts made to heterologously express the gene cluster proved slow and cumbersome.

Chapter 3 described how the predicted planosporicin propeptide sequence based on the 2007 structure of planosporicin had a number of errors (Castiglione *et al.* 2007). 454 sequence data revealed the C terminal Pro and Gly residues were actually located at positions 9 and 10 respectively. This was further confirmed in this Chapter through the amplification and sequencing of the *lanA* prepropeptide from *P. alba* as a probe to identify positive clones from the cosmid library. A retrospective investigation into the methods

used to identify the planosporicin biosynthetic gene cluster enables the conclusion that this genome scanning approach was appropriate. The differences in the precursor peptide compared to the published structure (Chapter 3) and lack of heterologous expression in *Streptomyces* (Chapter 6) would most likely have confounded other methods.

4.9 Summary

454 genome scanning revealed the *lanA* gene for planosporicin within the *P. alba* genome

A cosmid library was generated from *P. alba* gDNA

Three probes for the planosporicin gene cluster were amplified using 454 sequence data.

Eight out of 3072 clones hybridised to all three probes

Cosmids B4-1 (pIJ12321) and F13-1 (pIJ12322) were sequenced

pIJ12321 contained the entire 15.3 kb cluster centrally located within 37 kb of *P. alba* gDNA

Two cosmids from *P. sp.* DSM 14920 appear likely to contain the planosporicin gene cluster

Chapter 5 : The planosporicin biosynthetic gene cluster

5.1 Introduction

In Chapter 4, two cosmids were sequenced which revealed that pIJ12321 contains a 37.76 kb insert with the planosporicin structural gene, *pspA*, located centrally on the cloned insert. pIJ12322 is similar in size but the structural gene is located towards the right hand border of the insert so that the C-terminus of *pspV* is missing. Thus for the purpose of this work, it is the annotation of pIJ12321 which is described.

This Chapter details the bioinformatic analysis of the planosporicin biosynthetic gene cluster. This analysis will identify proteins from other organisms that are highly similar in sequence. These proteins are likely to exhibit conservation in which changes at a specific position of an amino acid sequence preserve the physico-chemical properties of the original residue. As function is conserved on the protein level, it is protein, not nucleotide, sequences that will be analysed. It is sometimes assumed that if two proteins have highly similar sequences then they are homologous, however sequence similarity does not always arise through common ancestry. Convergent evolution may result in two sequences being similar if they have both been selected to bind a particular substrate. Instead, these are functionally analogous sequences that have evolved separately. In this work the similarity between protein sequences will be quantified through 'percent identity', the extent to which two amino acid sequences are invariant.

In this Chapter, each of the proteins encoded within the proposed planosporicin biosynthetic gene cluster were used as a query in a search against all sequences in the current NCBI database to assign putative functions on the basis of sequence identity. In general, one type of database search, BLASTP, was used (Altschul *et al.* 1990). However in some cases a Position-Specific Iterative BLAST (PSI-BLAST) was used to find distant relatives of a protein (Altschul *et al.* 1997). This method first generates a list of closely related proteins from which the significant features are summarised and used as a query against the protein database to identify further proteins with a significant alignment. The process is iterated until no further proteins are found. Another bioinformatic tool was used for multiple sequence alignments, namely ClustalW2 (Chenna *et al.* 2003). Further

genomic analysis was performed using the Artemis Comparison Tool (Sanger Institute, Hinxton, UK).

5.2 Features of the minimal gene set

The annotation and bioinformatic analysis of pIJ12321 enabled a prediction of the likely minimal set of genes required for planosporicin biosynthesis, depicted in Figure 5.1. This annotation began in Chapter 3 with the identification of *pspA* encoding the planosporicin prepropeptide. A cluster of 14 genes surrounding *pspA* is likely to be responsible for planosporicin biosynthesis and they are discussed in Sections 5.3, 5.4 and 5.5. The cluster runs from 14970bp (*pspE*) to 30259bp (*pspV*) of the insert DNA of pIJ12321, representing a ~15.3 kb cluster. The GC content of the *P. alba* insert DNA cloned into pIJ12321 is 73.58 mol%, while the GC content of the cluster from *pspE* to *pspV* is slightly higher at 74.26 mol%. Interestingly, the rare codon 'TTA' occurs in the coding region of *pspX*, suggesting that production of planosporicin may be controlled by the *bldA* tRNA and therefore developmentally regulated (Leskiw *et al.* 1991). A detailed bioinformatics analysis of these genes along with others that may form part of the minimal gene set follows in the subsequent Sections and is summarised in Table 5.1.

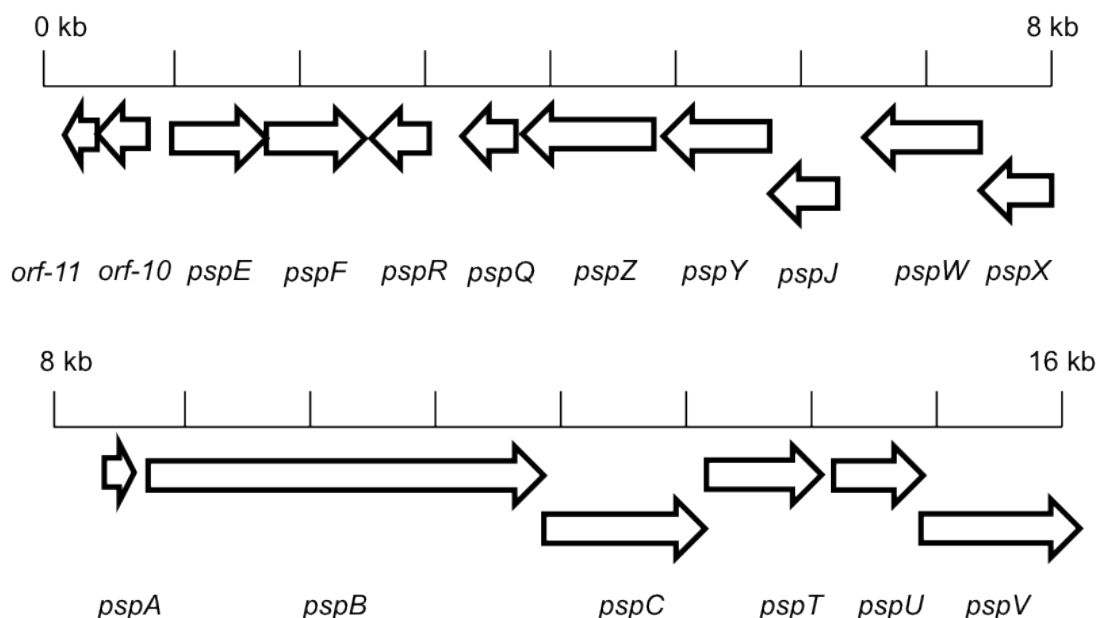


Figure 5.1 : The putative planosporicin gene cluster with each gene shown to scale. Translationally coupled genes are separated vertically.

ORF	<i>psp</i> nomenclature	length (aa)	Best match (other than <i>mib</i> cluster) found by BLASTP searches				Best match in <i>mib</i> cluster				Proposed function in <i>psp</i> cluster
			Organism	Accession number	Probability score	Annotated function	% Identity (across x amino acids)	Accession number	Probability score	Annotated function	
-11	?	78	<i>Streptomyces</i> sp. SPB78	ZP_072721_33.1	85.1 bits (209)	predicted protein	56 % (72)	N/A	N/A	N/A	Hypothetical protein
-10	?	102	<i>Ochrobactrum intermedium</i> LMG 3301	ZP_046804_25.1	33.9 bits (76)	protease Do	33 % (79)	N/A	N/A	N/A	Protease
-9	<i>pspE</i>	269	<i>Thermobispora bispora</i> DSM 43833	YP_003654_013.1	287 bits (734)	hypothetical protein Tbis_3431	59 % (268)	ADK32561.1	276 bits (707)	MibE; putative lantibiotic ABC transporter permease protein	ABC transporter (permease)
-8	<i>pspF</i>	249	<i>Thermobispora bispora</i> DSM 43833	YP_003654_012.1	304 bits (779)	ABC transporter-like protein	67 % (234)	ADK32562.1	295 bits (754)	MibF; putative lantibiotic ABC transporter ATP-binding protein	ABC transporter (ATP binding protein)
-7	<i>pspR</i>	260	<i>Symbiobacterium thermophilum</i> IAM 14863	YP_075969_016.1	47.8 bits (112)	two-component response regulator	38 % (82)	ADK32551.1	77.0 bits (188)	MibR; putative DNA-binding protein	Response regulator
-6	<i>pspQ</i>	138	<i>Thermobispora bispora</i> DSM 43833	YP_003654_016.1	142 bits (359)	hypothetical protein Tbis_3434	56 % (129)	ADK32550.1	140 bits (352)	MibQ; putative lipoprotein signal peptide-containing protein	Lipoprotein

Table 5.1 : Putative proteins of the planosporicin biosynthetic gene cluster.

Each amino acid sequence was used as a query in a NCBI Blast search (Altschul *et al.* 1990) against the non-redundant protein database (as of July 2011). The best match to the *mib* cluster and the best match from all other proteins are listed.

ORF	<i>psp</i> nomenclature	length (aa)	Best match (other than <i>mib</i> cluster) found by BLASTP searches				Best match in <i>mib</i> cluster				Proposed function in <i>psp</i> cluster	
			Organism	Accession number	Probability score	Annotated function	% Identity (across x amino acids)	Accession number	Probability score	Annotated function		% Identity (across x amino acids)
-5	<i>pspZ</i>	301	<i>Eggerthella lenta</i> DSM 2243	YP_003181_816.1	333 bits (854)	ABC transporter-like protein	60 % (293)	ADK32548.1	304 bits (779)	MibZ; putative ABC transporter ATP-binding protein	59 % (288)	ABC ATPase
-4	<i>pspY</i>	264	<i>Thermobispora bispora</i> DSM 43833	YP_003654_018.1	117 bits (292)	hypothetical protein Tbis_3436	40 % (245)	ADK32547.1	110 bits (274)	MibY; putative ABC permease	39 % (213)	ABC permease
-3	<i>pspJ</i>	213	<i>Thermobispora bispora</i> DSM 43833	YP_003654_019.1	64.3 bits (155)	hypothetical protein Tbis_3437	36 % (172)	ADK32546.1	57.0 bits (136)	MibJ; hypothetical protein	36 % (181)	Membrane protein
-2	<i>pspW</i>	314	<i>Thermobispora bispora</i> DSM 43833	YP_003654_015.1	98.6 bits (244)	hypothetical protein Tbis_3433	43 % (201)	ADK32552.1	90.5 bits (223)	MibW; putative anti-sigma factor	31 % (210)	Anti-sigma factor
-1	<i>pspX</i>	185	<i>Thermobispora bispora</i> DSM 43833	YP_003654_014.1	192 bits (489)	ECF subfamily RNA polymerase sigma-24	57 % (181)	ADK32553.1	199 bits (507)	MibX; RNA polymerase, sigma-24 subunit, ECF subfamily	59 % (184)	ECF sigma factor
1	<i>pspA</i>	56	<i>Actinomyces</i> sp. oral taxon 848 str. F0332	ZP_061621_52.1	45.8 bits (107)	conserved domain protein	55 % (42)	ADK32554.1	61.6 bits (148)	MibA; microbisporicin structural preprotein	67 % (46)	Pre-propeptide

Table 5.1 : Putative proteins of the planosporicin biosynthetic gene cluster (continued).

ORF	<i>psp</i> nomenclature	<i>psp</i> length (aa)	Best match (other than <i>mib</i> cluster) found by BLASTP searches				Best match in <i>mib</i> cluster				Proposed function in <i>psp</i> cluster	
			Organism	Accession number	Probability score	Annotated function	% Identity (across x amino acids)	Accession number	Probability score	Annotated function		% Identity (across x amino acids)
2	<i>pspB</i>	1067	<i>Desmospora</i> sp. 8437	ZP_084650_87	433 bits (1113)	lantibiotic biosynthesis protein	30 % (1039)	ADK32555	904 bits (2337)	MibB; lantibiotic dehydratase	52 % (1046)	Dehydratase
3	<i>pspC</i>	467	<i>Actinomyces</i> sp. oral taxon 848 str. F0332	ZP_061621_54.1	295 bits (755)	conserved hypothetical protein	42 % (443)	ADK32556.1	409 bits (1052)	MibC; lantibiotic cyclase	53 % (450)	Cyclase
4	<i>pspT</i>	322	<i>Streptomyces violaceusniger</i> Tu 4113	ZP_076059_13.1	235 bits (600)	daunorubicin resistance ABC transporter ATPase subunit	42 % (293)	ADK32558.1	355 bits (910)	MibT; ABC transporter ATP-binding domain	63 % (297)	ABC ATPase
5	<i>pspU</i>	265	<i>Thermobispora bispora</i> DSM 43833	YP_003652_761.1	145 bits (367)	ABC transporter protein	35 % (217)	ADK32559.1	255 bits (651)	MibU; ABC-type multidrug transport system permease component	61 % (231)	ABC permease
6	<i>pspV</i>	438	<i>Nocardioopsis dassonvillei</i> subsp. <i>dassonvillei</i> DSM 43111	YP_003682_582.1	141 bits (355)	hypothetical protein Ndas_4690	41 % (305)	ADK32560.1	201 bits (512)	MibV; hypothetical protein	42 % (324)	Hypothetical protein

Table 5.1 : Putative proteins of the planosporicin biosynthetic gene cluster (continued).

5.3 Biosynthetic enzymes

5.3.1 PspB

The 1067 residue PspB protein bears most resemblance to the lantibiotic dehydratases (which are commonly ~1000 residues). It contains three conserved domains, two of which correspond to the N and C termini of lantibiotic dehydratases (pfam04737 and pfam04738) and a third which commonly also occurs in longer LanB proteins (TIGR03891). The protein with most similarity to PspB was MibB, the microbisporicin dehydratase from *M. corallina* (Table 5.1). The top non-*mib* result was a putative lantibiotic dehydratase from *Desmospora* sp. 8437 (Table 5.1). This genus is in the family *Thermoactinomycetaceae*, Gram-positive endospore-forming bacteria in the Firmicutes phylum. This strain was isolated from human blood and subject to whole genome shotgun sequencing as part of the Human Microbiome Project. Nucleotide region 42690-45878 of the database entry contains the PspB-like protein. Downstream lies a putative LanC as well as several ABC transporters; however no putative prepropeptide had been annotated.

The next four proteins with significant alignments to PspB are from different species of *Bacillus*. Namely *Bacillus clausii* KSM-K16 (28 % over 1098 residues), *Bacillus megaterium* QM B1551 (29 % over 1082 residues), *Bacillus cereus* AH 1273 (28 % over 1033 residues) and *Bacillus mycooides* DSM 2048 (29 % over 1033 residues). The gene that encodes the PspB-like protein in *B. megaterium* (pBM70150) lies downstream of a gene encoding a protein which resembles PspA (pBM70151) identified in Chapter 3 and named megateracin. Likewise the gene encoding the PspB-like protein in *B. mycooides* (bmyco0001_53840) lies downstream of a putative *lanA* gene (bmyco0001_53830). Both of these putative 22 amino acid lantibiotics show a good alignment to clausin (*B. clausii*), gallidermin (*S. gallinarum*) and epidermin (*S. epidermidis*), as well as to the 24 amino acid lantibiotic planosporicin (*P. alba*). The context of these four *Bacillus* clusters is discussed further in Section 5.6.3.

Other proteins with good alignments to PspB include lantibiotic dehydratases from several streptomycetes, including *Streptomyces violaceusniger* Tu 4113 (34 % over 1099 residues) and *Streptomyces noursei* ATCC 11455 (35 % over 1125 residues). Other actinomycetes, such as *Actinomyces* sp. oral taxon 848 str. F0332, contain a protein with similarity to PspB; 32 % identity over 1074 residues. Interestingly, the PspB-like protein in *Actinomyces* sp. lies downstream of a protein identified due to its similarity to PspA and named actoracin (described in Chapter 3).

PspB has lower levels of similarity to LanB proteins from characterised biosynthetic gene clusters. There is just 23 % identity to SpaB and EpiB, the dehydratase enzymes involved

in subtilin and epidermin biosynthesis, respectively. PspB has just 22 % identity to MutB and NisB, the enzymes responsible for dehydrations in mutacin III and nisin, respectively. All four of these LanB proteins correspond to type AI lantibiotic gene clusters. The low level of similarity is to be expected, as members of the LanB family commonly share only ~25 % sequence identity (Chatterjee *et al.* 2005). On the basis of similarity to so many LanB enzymes, PspB is proposed to catalyse the formation of Dha and Dhb residues in planosporicin through the dehydration of specific serine and threonine residues, respectively.

Over the years, much *in vivo* work has been carried out to characterise the substrate specificity of LanB enzymes. The use of lantibiotic expression systems to produce lantibiotic analogues demonstrate the substrate promiscuity of these enzymes. For example the modification of non-lanthionine peptides by NisB and NisT when heterologously expressed as a chimera with the nisin leader peptide indicates the substrate promiscuity of LanB enzymes (Rink *et al.* 2007c). Likewise, the LanM enzymes responsible for the dehydration and cyclisation of type AI and type II lantibiotics demonstrate similar promiscuity which has been exploited in evolution. The planktonic marine cyanobacteria *Prochlorococcus* MIT9313 was found to encode a single LanM enzyme responsible for the dehydration and cyclisation of up to 29 different prepropeptides (Li *et al.* 2010). However this promiscuity must be tempered with a certain level of selectivity. Planosporicin contains an unmodified threonine residue at position 12 which is conserved in microbisporicin and actoracin (described in Chapter 3) whereas the threonine at position 2 is dehydrated to form dehydrobutyrine in all three lantibiotics.

5.3.2 PspC

pspC encodes a 467 amino acid protein which contains the conserved domain pfam05147 (LanC-like proteins) and shows strong similarity to lantibiotic cyclases. The two proteins with the most significant alignment to PspC were both proteins from actinomycetes. These are MibC, the lantibiotic cyclase responsible for the formation of (Me)Lan bridges in microbisporicin from *M. corallina*, and a conserved hypothetical protein from *Actinomyces* sp. oral taxon 848 str. F0332 (Table 5.1). The 'hypothetical' protein from *Actinomyces* sp. occurs downstream of the actoracin prepropeptide and likely LanB and is almost certainly the LanC enzyme in this biosynthetic gene cluster. The next five hits correspond to various members of the Firmicutes phylum. These include *Desmospora* sp. 8437 (33 % over 432 residues), *Paenibacillus alvei* (31 % over 448 residues), *Paenibacillus polymyxa* E681 (32 % over 462 residues), *B. cereus* AH 1273 (31 % over 434 residues) and *B. mycoides* DSM 2048 (31 % over 434 residues). In addition, both *B. clausii* KSM-K16 and

B. megaterium QM B1551 contain proteins with high similarity to PspC; 31 % (over 462 residues) and 30 % (over 447 residues), respectively. This supports the prediction that the planosporicin-like lantibiotics clausin and megateracin are produced by these strains (as described in Chapter 3) and is discussed further in Section 5.6.3. LanC enzymes from characterised lantibiotic biosynthetic gene clusters with similarity to PspC include SpaC (29 % over 454 residues), NisC (26 % over 401 residues), EpiC (25 % over 399 residues) and MutC (21 % over 430 residues).

NisC has been demonstrated to be necessary *in vivo* for the introduction of (Me)lan bridges into nisin (Koponen *et al.* 2002). This was further substantiated *in vitro* where purified NisC was demonstrated to introduce all five thioether crosslinks into nisin (Li *et al.* 2006). Subsequently NisC was subjected to site-directed mutagenesis of conserved residues (His₂₁₂, Arg₂₈₀, Asp₁₄₁, and Tyr₂₈₅) as well as of the ligands to the zinc ion in the active site (Cys₂₈₄, Cys₃₃₀, and His₃₃₁). His₂₁₂ and Asp₁₄₁ as well as all three zinc ligands were found to be essential for cyclisation activity (Li and van der Donk 2007).

LanC enzymes are zinc metalloproteins, in which the bound metal is proposed to activate the thiol moiety of the cysteine residue for nucleophilic addition onto the dehydrated serine or threonine (Li *et al.* 2006). PspC contains the conserved histidine and cysteine residues which are proposed to coordinate a zinc ion involved in the catalysis of (Me)Lan bridge formation in planosporicin (Figure 5.2).

5.3.3 PspTU

ATP-Binding Cassette (ABC) transporters use the hydrolysis of ATP without protein phosphorylation to energise the transport of substrates across cell membranes. ABC transporters contain two conserved regions; a highly conserved nucleotide-binding domain (NBD) and a less conserved transmembrane domain (TMD). The NBDs bind and hydrolyse ATP to supply the active energy for transport. Substrate specificity is determined by two TMDs that also provide the translocation pathway across the bilayer.

The 322 residue PspT protein contains the conserved domain cl09099, signifying it is a member of the P-loop NTPase Superfamily. PspT contains the Walker A motif (GxxxxGK[S/T], where x is any residue), and the Walker B motif (hhhh[D/E], where h is a hydrophobic residue) (Figure 5.3 A). The Walker A motif forms a phosphate-binding loop which binds the beta-gamma phosphate moiety of ATP while the Walker B motif binds the Mg^{2+} cation. In addition, three other conserved motifs are present in PspT. The H-loop/switch region contains a highly conserved histidine residue that is also important in the interaction of the ABC domain with ATP (Figure 5.3 A). The ABC transporter signature is a motif (also known as the LSGGQ motif) that is specific to the ABC transporter family (Figure 5.3 A). The Q-loop/lid, located between Walker A and the signature, consists of a glutamine residue residing in a flexible loop (Figure 5.3 A). This is thought to connect the NBD to the TMD to couple ATP hydrolysis to the conformational changes of the TMD during substrate translocation. Together the Walker A, Walker B, Q-loop and H-loop regions form the nucleotide-binding site (Schneider and Hunke 1998).

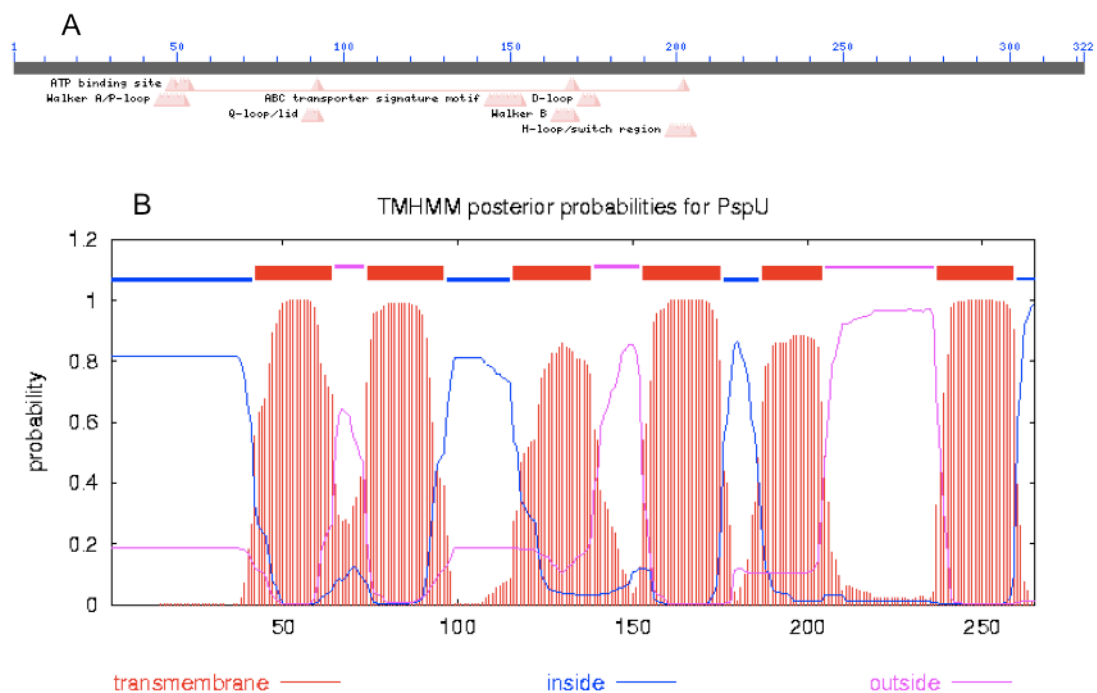


Figure 5.3 : Features of the ABC transporter PspTU.

A; Putative conserved domains detected in PspT (using BLASTP (Altschul *et al.* 2001)). **B**; Predicted transmembrane helices in PspU (using TMHMM (Krogh *et al.* 2001)).

The protein with the most significant alignment to PspT is the MibT ABC transporter ATP-binding domain from *M. corallina* (Table 5.1). This protein is proposed to have a role in the export of the lantibiotic microbisporicin after biosynthesis (Foulston and Bibb 2010).

Another lantibiotic export protein with high similarity to PspT is CinT (44 % across 282 residues). The CinT protein from *S. cinnamoneus* is likely to export the lantibiotic cinnamycin out of the mycelium (Widdick et al. 2003). A number of other proteins with high similarity to PspT are annotated as the ATPase subunit of a daunorubicin resistance ABC transporter, DrrA. Daunorubicin was initially isolated from *Streptomyces peucetius* and is a chemotherapeutic of the anthracycline family (Stutzman-Engwall and Hutchinson 1989). These PspT-like proteins were identified from three different phyla, which included a couple of Actinobacteria; *S. violaceusniger* Tu 4113 (42 % across 293 residues) and *Thermobispora bispora* DSM 43833 (43 % across 303 residues). In addition there was a thermophilic anaerobe able to produce hydrogen as a waste product from carbon monoxide and water, the Firmicute *Carboxydotherrmus hydrogenoformans* Z-2901 (38 % across 301 residues), and a couple of members from the Chloroflexi phylum; a thermophilic, green non-sulfur bacterium *Thermomicrobium roseum* DSM 5159 (40 % across 299 residues), and a mesophilic, anaerobic bacterium *Dehalogenimonas lykanthroporepellens* BL-DC-9 (39 % across 298 residues). The only protein with a different annotated function amongst the eight proteins with most significant similarity to PspT was a protein from *Streptomyces sviveus* ATCC 29083 (40 % identity across 298 residues). This protein was annotated as the nodulation ABC transporter NodI, which in nitrogen-fixing root nodule bacteria works with NodJ to export a variety of modified carbohydrate molecules as signals to plant hosts to establish root nodules.

The TMD consists of alpha helices embedded in the membrane bilayer. Most TMDs consist of a total of 12 α -helices with six α -helices per monomer to form a transmembrane-spanning permease component which functions to recognise substrates and undergo conformational changes to enable transport across the membrane (Young and Holland 1999). Using the TMHMM Server v. 2.0, the 265 amino acid protein PspU was predicted to contain six transmembrane (TM) helices (Figure 5.3 B) (Krogh et al. 2001). The best alignment of PspU was to MibU, the ABC-type multidrug transport system permease component from *M. corallina* (Table 5.1). Other hits included ABC transporter proteins from a number of actinomycetes; *T. bispora* DSM 43833, *Thermomonospora curvata* DSM 43183 (in the same suborder, Streptosporangineae, as *Planomonospora*), *Acidimicrobium ferrooxidans* DSM 10331 and *Sphaerobacter thermophilus* DSM 20745 (although in 2004 *Sphaerobacter* was reclassified to the class Thermomicrobia in the phylum Chloroflexi (Hugenholtz and Stackebrandt 2004)). In addition, three of the top eight hits are annotated as the membrane protein of the daunorubicin resistance ABC transporter, DrrB. Notably the CinH protein from *S. cinnamoneus* also had a significant alignment to PspU (34 % over 235 residues).

On the basis of the observed similarity to ABC transporters involved in the export of secondary metabolites, PspTU is proposed to function as the exporter responsible for the efflux of planosporicin out of *P. alba* mycelium after biosynthesis. Like CinTH and MibTU, PspTU is proposed to function as a tetramer of two dimers. As with both CinH and MibU, no putative peptidase domain was identified in PspU. This supports the idea that two classes of lantibiotic exporters exist, as discussed in Chapter 1. The division lies not with the classification of the lantibiotic, but rather is a reflection of the phylum of the producing organisms. All known lantibiotic gene clusters characterised so far from the actinomycete phylum contain exporters in which the NBD and TMD domains are located on separate proteins, whereas lantibiotic clusters from low-GC producers use a single protein containing both the NBD and TMD domains. Thus PspTU is another example from a high-GC organism that encodes the two domains on separate proteins.

5.3.4 PspYZ

As with PspT, the 301 amino acid protein PspZ contains the conserved domain cl09099 corresponding to the Walker A, Walker B, Q-loop and H-loop regions which form the nucleotide-binding site (Figure 5.4 A).

Although no putative conserved domains were detected in PspY, it had a high percentage identity with a number of TMD proteins. The TMHMM Server v. 2.0 predicted that the 264 amino acid protein PspY contained six TM helices (Figure 5.4 B) (Krogh et al. 2001). A BLASTP search revealed just 16 proteins with a significant alignment which ranged from 40 % across 245 residues for Tbis_3436 from *T. bispora* DSM 43833 to 37 % across 140 residues for Elen_1457 from *Eggerthella lenta* DSM 2243. When compared with the search results using PspZ as query, all 16 hits for PspY also have a neighbouring gene that has a significant alignment with PspZ. A majority of these PspYZ-like proteins are encoded by the order Actinomycetales in the Actinobacteria phyla; Tbis_3436 and 3435 from *T. bispora* DSM 43833, MibYZ from *M. corallina*, Sros_1169 and 1168 from *Streptosporangium roseum* DSM 43021, Franean1_3994 and 3995 from *Frankia* sp. EAN1pec, and ZP_06162160 and 06162161 from *Actinomyces* sp. oral taxon 848 str. F0332. Other orders included the Bifidobacteriales with Bbr_1071 and 1072 from *Bifidobacterium breve* UCC2003, and Blon_1404 and 1405 from *Bifidobacterium longum* subsp. infantis ATCC 15697, and the order Coriobacteriales with Elen_1457 and 1458 from *E. lenta* DSM 2243 and two genes from *Eggerthella* sp. 1_3_56FAA. Together these account for 10 of the PspY-like proteins. Five of the remaining matches correspond to the Chloroflexi *Ktedonobacter racemifer*, an aerobic and mesophilic bacterium. The final PspY-like protein is from an uncultured bacterium division (OP8) from an environmental

sample. For eleven of these PspYZ-like proteins, enough genome sequence was available to investigate the gene context of these ABC transporters. This is summarised in Figure 5.5 and discussed further in Section 5.6.2.

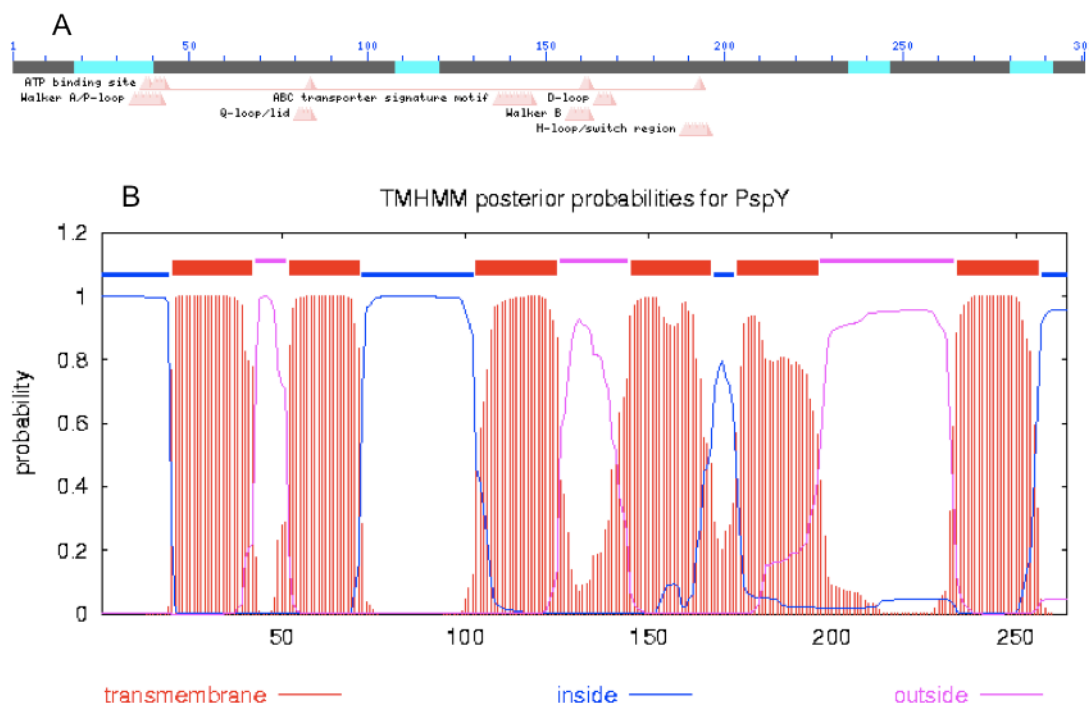


Figure 5.4 : Features of the ABC transporter PspYZ.

A; Putative conserved domains detected in PspZ (using BLASTP (Altschul *et al.* 2001)). Areas in blue were omitted from the domain search due to compositional bias. **B**; Predicted transmembrane helices in PspY (using TMHMM (Krogh *et al.* 2001)).

This analysis does not enable a conclusion as to the function of PspYZ. The presence of similar proteins in gene clusters that also encode putative prepeptide nearby (*M. corallina* and *Actinomyces* sp.) implies that PspYZ does play a role in planosporicin biosynthesis. The level of similarity between the NBD and TMD domains of PspTU and PspYZ differs significantly. While both PspT and PspU bear most similarity to MibTU (63 % and 61 %, respectively, with all other hits ~40 %), PspZ (NBD) has 59 % identity to MibZ while PspY (TMD) has only 39 % (and all other hits have a similar % identity). This could reflect a difference in the substrate recognised. While PspU and MibU are proposed to recognise similar lantibiotics for export, it could be that PspY and MibY are recognising very different substrates. On this basis it can be concluded that all the PspY-like proteins shown in Figure 5.5 are transporting different, as yet unknown, substrates.

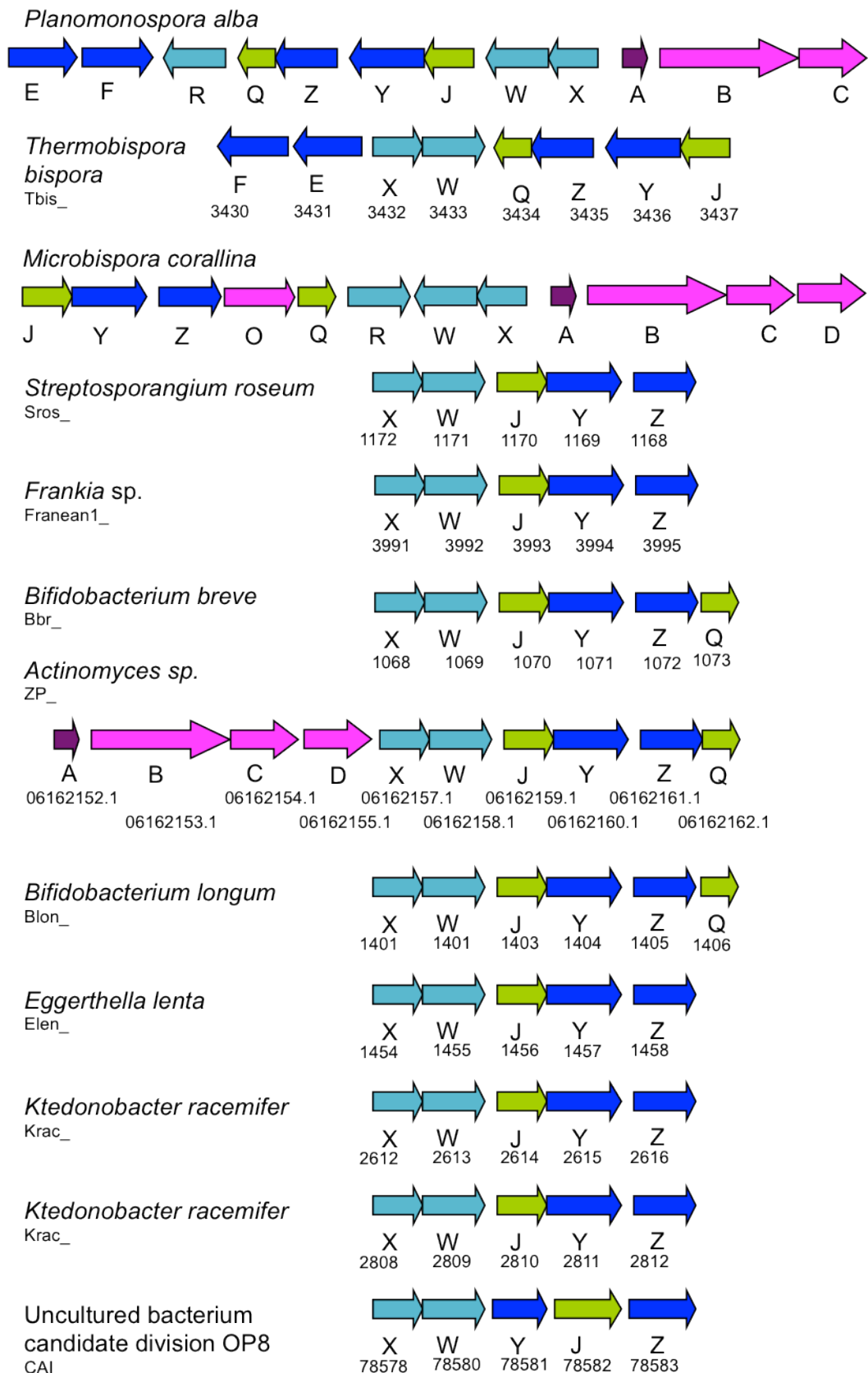


Figure 5.5 : A schematic showing the syntenous arrangement of the proteins bearing significant similarity to those encoded by *pspX*, *pspW*, *pspJ*, *pspY* and *pspZ*.

Genes are coloured and labelled according to the respective genes from the *psp* cluster, with the locus tag given below.

5.3.5 PspEF

As seen in both PspT and PspZ, the 249 residue PspF sequence contains the characteristic regions which form the nucleotide-binding site of ATP-binding proteins (Figure 5.6 A). Like PspU and PspY, the 268 residue PspE had a high percentage identity with a number of ABC transporter permease proteins. The TMHMM Server v. 2.0 predicted that PspE contained six TM helices (Figure 5.6 B) (Krogh *et al.* 2001). Thus, like PspTU and PspYZ, PspE and PspF are thought to function coordinately as an ABC transporter in which PspE is the membrane embedded permease unit and PspF is the ATP-binding cassette.

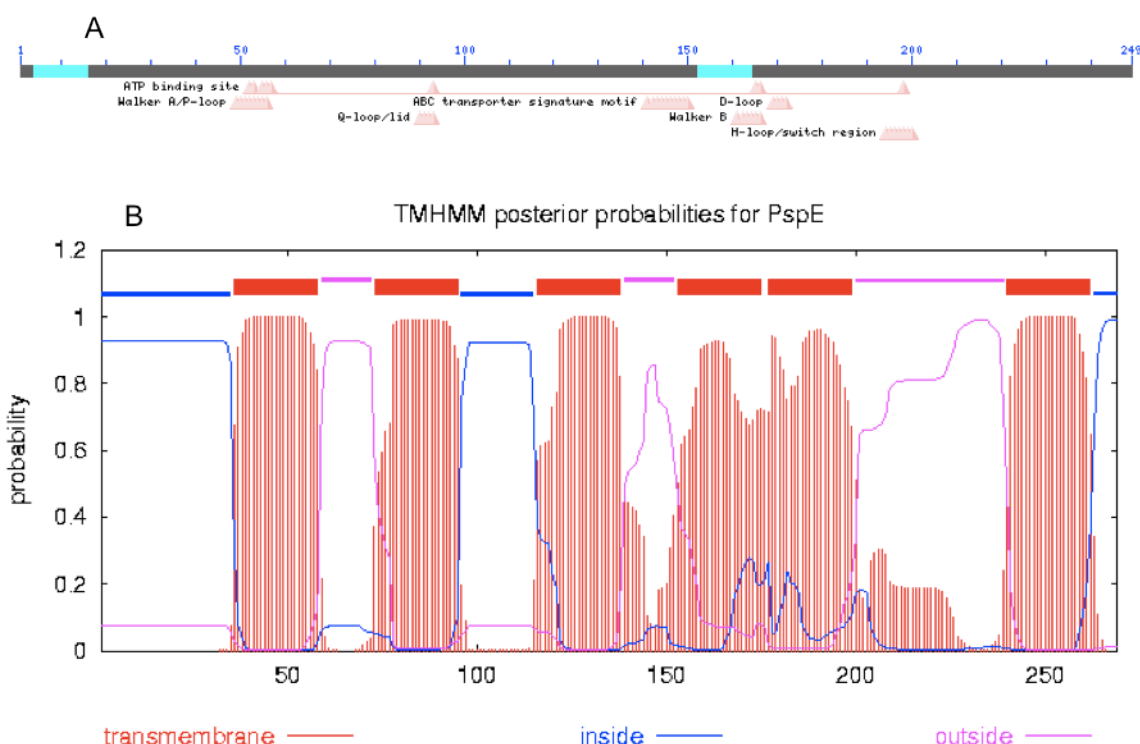


Figure 5.6 : Features of the ABC transporter PspEF.

A; Putative conserved domains detected in PspF (using BLASTP (Altschul *et al.* 2001)). Areas in blue were omitted from the domain search due to compositional bias. **B**; Predicted transmembrane helices in PspE (using TMHMM (Krogh *et al.* 2001)).

In lantibiotic producers, immunity is commonly provided (in part) through an ABC transporter with the generic name LanFEG (as described in Chapter 1). The LanF protein forms a homodimer of ATP-binding proteins that associate with the LanEG heterodimer embedded in the membrane. There is some sequence similarity between PspEF and proteins conferring self-protection on other producers of type AI lantibiotics, the highest of

which is 25 % identity over 240 amino acids to SpaE in the subtilin gene cluster of *B. subtilis*.

Nisin is the best-studied lantibiotic and the immunity mechanism is well understood. The nisin biosynthetic gene cluster encodes two immunity mechanisms which both contribute to producer immunity. NisFEG exports nisin that has already entered the cytoplasmic membrane to reduce the number of cell-attached nisin molecules (Siegers and Entian 1995). In parallel, the lipoprotein NisI is orientated to the outside of the cytoplasmic membrane where it sequesters nisin at the surface of the membrane, thus preventing it from binding lipid II and forming pores (Kuipers *et al.* 1993a). In addition approximately half of the NisI produced escapes the lipid modification and is instead secreted into the medium where it is likely to play a role in nisin immunity (Koponen *et al.* 2004).

In contrast, several lantibiotic immunity systems have been described which lack a lipoprotein LanI entirely. Instead the LanFEG ABC transporter alone confers self-protection on the producing strain. For example *S. epidermidis*, the producer of epidermin, uses the ABC transporter EpiFEG to confer immunity (Peschel and Gotz 1996). Likewise MrsFEG is the only requirement for self-resistance to mersacidin (Guder *et al.* 2002).

However the lantibiotic producer to which PspEF has highest similarity is MibEF from *M. corallina* (Table 5.1). MibEF is an ABC transporter proposed to confer self-protection to microbisporicin (Foulston and Bibb 2010). In this cluster, the LanE functions without a LanG to presumably form a homodimeric permease. This is also thought to be the case in the lactacin 3147 and the bovicin HJ50 gene clusters in which no *lanG* genes have been identified (Draper *et al.* 2009; Liu *et al.* 2009). In accordance with this, the six TM helices of PspE are proposed to form a dimeric protease unit with another PspE to traverse the membrane 12 times in total.

Sequence alignments reveal that in contrast to the Q loop, LanF proteins contain the E loop, where instead of a glutamine residue, a glutamic acid residue is present (Figure 5.7). The E loop in the NukF protein was found to play a significant role in resistance to nukacin ISK-1 in *Staphylococcus warneri* ISK-1. The amino acid replacements E85Q and E85A in the E loop were found to impair the level of immunity and transport activity whilst maintaining ATPase activity at WT levels (Okuda *et al.* 2010). The reciprocal experiment in the daunorubicin transporter of *Streptomyces peuceticus*, DrrA, revealed that the amino acid replacement Q88E reduced transport activity (Rao and Kaur 2008). Thus it would appear that LanF and DrrA form two separate groups of ABC transporters. Indeed in the alignment in Figure 5.6, PspF is more similar to NukF and other LanF proteins, all of which display a conserved E-loop (in purple), while PspT and PspZ are more similar to DrrA, LanT and LanZ proteins with a conserved Q-loop (in orange).

```

MibF      -----MAALEIRDLHKHYD---DFHALDGANLTVPDGSLYGLLGPNGAGKTT 44
PspF      MSSPFPPTQPAPALVIRRLHKHYG---AHRALDGVLDTAGRGEVYGLLGPNGAGKTT 57
SpaF      --MKRIKGESDMSSVLETKGLTKKFG---KETAVADVSLTVQKNSIYGLLGPNGAGKST 55
MutF      -----MDYMLETKNLTQKQFG---KQTAVNQLNLKVERHSIYGLLGPNGSGKST 45
EpiF      -----MTYYKIETRNLNKRFK---KETILKDINLKVYPNEVYGLLGINGAGKST 46
NukF      -----MINNIVQTQNLTKKFS---DSYSVDNLSLNIKSKEIYGFLPNGAGKST 46
NisF      -----MQVKIQNLSKTYK---EKQVLQDISFDIKSGTVCGLLGVNGAGKST 43
MibZ      -----MNTSISITDLTKRYRRGGDRPALNGVSLTVDGG---MTALLGANGAGKTT 47
PspZ      -----MTITTERLTKVYS---GGRRALDGVLDLDIGGG---TVGLLGANGAGKTT 43
MibT      ----MTVP-----AFELSDLTVRYG---PVTAVDGVVSAGSAPGLVTALLGPNGAGKSS 46
PspT      ----MTGPGSTETPAAFEITGLVVRHG---KVTAVRDVSIASPPGRVTALLGPNGAGKSS 53
CinT      -----MRGIDLTVAQGETFGFLGPNGAGKTT 26
DrrA      ---MNTQPTR-----AIETSGLVKYYN---GTRAVDGLDLNVPAGLVYGILGPNGAGKST 49
          : . : : * * * * * : :

MibF      LMKAVTGLRHPTSGHISLFRPYYERL----LTQVGALLESPGLWTQLDVAVSHLRIHAR 99
PspF      LMKALLGLQRPSTSGDIEIFGRAVGRES----LSEVGALIEVPGLWPALTGDETLRVHAR 112
SpaF      TLKILTGILRKTSGEIFFDGHKWRSD----LKNIGSLIESPPLYDNLTAFENVKVHAT 110
MutF      TLKMITGMLRKTSGHILIDGHDWSRKD----LENIGALIESPPLYENLTARENLKVRTL 100
EpiF      LMKIICGILPQTSGLDIYLDGSPLTRKE----LTNIGSLIESPPTYNHLSAQDNLKVVAL 101
NukF      TMKMLLGLMQPTKGNIKIFNQDISKNRDE--ILMHVGALIEEPSYKNLTSAEQNLQVQR 104
NisF      LMKILFGLISADTGKIFFDQEKTNNO-----LGALIEAPAIYMNLSAFDNLKTKAL 95
MibZ      LMRI CVGVL RPDGGRVVVGGHDLGTAAGRRVAVKRI LGYLPQELS MYD DLTGREFLDYIAL 107
PspZ      LLRILTVGLRPTSGRVVAGHDLSTGAGLTALKHTLGYLPQELSLY PDLTAREFLDYIGV 103
MibT      LLRVLSTIAPPSSGTARVFGHDVRAEPL--AARTRIGLVFQERALD TDL SAEQNLRFHAR 104
PspT      LLSVLSTASRPDEGTVRVFVGHVRAEPT--AARRRLGLV FQERTLDKELSVERN LWFHAR 111
CinT      TISMLTTLATPTTGRIEIAGHDTRTAPQ--QVRRNLGLV FQETTLDP E LTAVENLRFHAD 84
DrrA      TIRMLATLLRPDGGTARVFGHDVTSEPD--TVRRRISVTFQYASVDEGLTGTENLVMMGR 107
          : * . . : : * :

MibF      LRGVPETR-----IGEVL SLMNLT E VSTRK VAKYSLGMRWRLGIAIALLGRPRLVVLDPE 154
PspF      LRCVPEGW-----IDPALRLVGMTGARTRKVGTYSLGMRWRLGIAVALLARPRLVVLDPE 167
SpaF      LLGLSKER-----INSVLETVDMLHAGKKRAGQFSMGMRQRLGIAIALLNHPKLLILDEP 165
MutF      MLGLPDSR-----IDEVLKIVDLTNTGKKRAGQFSMGMKQRLGIAIALLNSPQLLILDEP 155
EpiF      NENIDFNE-----ISKVLELVNLNVEPQKKVKNFSLGMKQRLGIAMALIKKPKLLVLDPE 156
NukF      LLNLPSKN-----VKEALKIVRLTEHKDKLVKNYSLGMKQRLGIALAIVKFPKLLILDEP 159
NisF      LFGISDKR-----IHETLEVIGLAETGKKRAGKFSLGMKQRLGIGMAILTEPQFLILDEP 150
MibZ      LKGVDKRVRRDQIEQMLELTGLSEHAGRRLLGGYSGMKRRLGIAQALLAEPELIVVDEP 167
PspZ      LKCIDRRRERRRQVDERIAEVLGQERADERLAGFSGGMKRRVGI AQALMGDPSVLIVDEP 163
MibT      LFGVGRAR-AREDILVLLERFGLAGRDRVETLSGGLARRLEIARALLHRPGLLILDEP 163
PspT      LFGMARED-ARARIDMLDLFGLADRRGQAI EELSGGLARRVEIARALLHRPGLLIVLDEP 170
CinT      LYALPRAG-LAGRIAEMLELVGLSARGDSLGRTFSGGMQRRL E IARGLLHRPRLFLDEP 143
DrrA      LQGYSWAR-ARERAAELIDGFGLDARDRLLKTYSGGMRRRLDIAASIVVTPDLLFLDEP 166
          : : * * : * : * . . : * . . : * * *

MibF      MNGLDPVGI RDMRATLRALT-AAGTTVMVSSHQLAEIAHICDHVGVLVAGRTAYEGPLHG 213
PspF      TNGLDPVGI REMRGIIRSLA-GEGVTVLIASHQLAEVAQVCDRVQVLVAGRTRYEGSLAG 226
SpaF      TNGLDPIGIQELRDLIRSF-LEGITVILSSHILSEVEQIADHIGIISDGHLYQGEIN- 223
MutF      TNGLDPIGIQELRNLRISFP-TQGITV ISSHILSEIQMTADHIGI IANGVLGYQDRIH- 213
EpiF      SNGLDPYGIQELRELLKSLT-NLNTSVIISSHILSEIQQLADHVGIIHNKLEYQEENK- 214
NukF      TNGLDPSGIQEIRELIKSFPKTYGMTVLISHLLSEIEHMANTVGIINRGKLEFGKLT- 218
NisF      TNGLDPDGI AELLNLILKLK-AGVTVILSSHQLHEISKVASQIIILNKGKIRYNRANN- 208
MibZ      TAGLDPSEMRFRSLLAGLG-GARRTVVLSTHILDDAAQTCPNITVLHGQRVAYQGSTAG 226
PspZ      TTGLDPHERMFRFTLLAGLG-RGR-TVLLSTHILDDVAQTCPEVAVLAAGRLLFHDATTA 221
MibT      TNGLDPEARQTVWDDLIRLRSLELVTVLYSTHYMDEA-ELADQIIILSEGRVAGFGSPGR 222
PspT      TNGLDPSARAAVWADLLLLRRELGVTTILYATHYMDEA-EYADEIVILAKGRIVRQGS 229
CinT      TIGLDPQTRAQVWAHLAEVREREATTIFLTHYLDEA-EQCRIAIIIDGRIVAQGS 202
DrrA      TTGLDPRSRNQVWD-IVRALVDAGTTVLLTQYLDEADQLADRIAVIDHGRVIAEGTTGE 225
          * * * . : : : : : . : : *

```

Figure 5.7 : Sequence alignment of the three ATP-binding proteins in the planosporicin cluster (PspF, PspZ, PspT) with ten highly similar ATP-binding proteins.

Four are from high-GC lantibiotic clusters; MibF, MibZ, MibT (microbisporicin) and CinT (cinnamycin). Five are from low-GC lantibiotic clusters SpaF (subtilin), EpiF (epidermin), NisF (nisin), NukF (nukacin ISK-1) and MutF (mutacin III). The last is DrrA, the daunorubicin export ATP-binding protein from *Streptomyces peuceticus*. The alignment was constructed using ClustalW2 and is truncated at the C-terminus (Chenna *et al.* 2003)). Common motifs are highlighted, namely the Walker A and Walker B motifs (red), E/Q loop (purple/orange), Signature motif (blue), and H loop (green) (Okuda *et al.*, 2010).

On the basis of the observed similarity of PspEF to ABC transporters involved in lantibiotic immunity, PspEF is proposed to function as the mechanism of immunity by diminishing the concentration of planosporicin molecules near the lipid II target by actively transporting them away from the membrane.

5.4 Regulatory proteins

5.4.1 PspX

Sigma factors are key regulators of bacterial transcription, able to recruit the RNA polymerase core enzyme to relevant promoter sequences. Over twenty years ago, a new sigma factor was purified from *E. coli* RNA polymerase and named σ^E or σ^{24} (Erickson and Gross 1989). The subsequent discovery of homologues in other bacteria including the σ^E of *S. coelicolor* revealed that these sigma factors form a distinctive subfamily within the σ^{70} family (Lonetto *et al.* 1994). These sigma factors were found to be involved in the response to cell envelope or oxidative stress as well as the regulation of iron uptake, development and virulence, so were named the extracytoplasmic function (ECF) sigma factors (Missiakas and Raina 1998). Members of the σ^{70} family have four regions that are generally conserved, of which ECF sigma factors share similarity across three. Region 1 ensures that the sigma factor will only bind the promoter when it is complexed with core RNA polymerase, but is only found in some ECF sigma factors (Lonetto *et al.* 1994). *pspX* encodes a 185 amino acid protein containing the conserved domain pfam04542, part of the cl08419 superfamily. Pfam04542 corresponds to the most conserved region of the σ^{70} class, region 2. This contains the primary core RNA polymerase binding determinant (region 2.1), which interacts with the clamp domain of beta prime, the largest polymerase subunit (Malhotra *et al.* 1996). Region 2.3 includes four aromatic residues involved in promoter melting, while region 2.4 contains the -10 sequence recognition determinants (Lonetto *et al.* 1994). Region 3 is not present in ECF sigma factors, with the consequence that these proteins are often shorter than others in the σ^{70} family (Lonetto *et al.* 1994). Region 4 contains a helix-turn-helix DNA binding motif that recognises the -35 element of promoters (Campbell *et al.* 2002). In accordance with this, all sequences with a significant alignment to PspX are annotated as RNA polymerase sigma factors of either the σ^{70} family or more specifically the σ^{24}/σ^E extracytoplasmic function (ECF) subfamily.

The two protein sequences bearing most similarity to PspX are ECF sigma factors MibX from *M. corallina* (58 % identity end-to-end) and Tbis_3432 from *T. bispora* DSM 43833 (56 % identity end-to-end) (Table 5.1). Five further proteins from actinomycetes make up the top seven sequences with most similarity to PspX. These are ECF sigma factors Franean1_3991 from *Frankia* sp. EAN1pec (52 % over 163 residues), Sros_1172 from *S.*

roseum DSM 43021 (49 % over 168 residues), ZP_06162157.1 from *Actinomyces* sp. oral taxon 848 str. F0332 (43 % over 176 residues), Blon_1401 from *B. longum* subsp. *infantis* ATCC 15697 (42 % over 170 residues) and Bbr_1068 from *B. breve* UCC2003 (40 % over 170 residues). Slightly less conserved was Elen_1454 from *E. lenta* DSM 2243 (42 % over 161 residues) and CAI_78578 from an uncultured candidate division OP8 bacterium (34 % over 183 residues). All nine of these PspX-like proteins are encoded in gene clusters containing minimally *lanXWJYZ* (Figure 5.5).

A couple of years ago, Hidden-Markov Model analysis grouped 1735 ECF proteins into 43 major phylogenetically distinct groups on the basis of sequence similarity (Staron *et al.* 2009). The associated web tool, *ECFfinder*, was used to classify PspX into one of the groups. The best match had a score of 184.1 to group ECF01, a group of RpoE-like ECF sigma factors from nine different phyla. In contrast, it had previously been noted that group ECF33 contained three sigma factors with sequence similarity to MibX (Foulston 2010). These three proteins also shared a significant percentage identity with PspX. These were BBta_3011 from *Bradyrhizobium* sp. BTAi1 (37 % over 184 residues), M115118 from *Mesorhizobium loti* MAFF303099 (38 % over 174 residues) and Rpal_2022 from *Rhodopseudomonas palustris* TIE-1 (35 % across 175 residues).

5.4.2 PspW

Very few sequences in the current NCBI database produce a significant alignment with the 314 residue protein PspW. The two most similar sequences are the hypothetical protein Tbis_3433 from *T. bispora* DSM 43833 (43 % over 201 residues) and putative anti-sigma factor MibW from *M. corallina* (38 % over 201 residues) (Table 5.1). Of the total of nine hits, four more are of particular interest. These are the hypothetical proteins Sros_1171 from *S. roseum* DSM 43021 (31 % over 210 residues), Krac_2613 and Krac_2809 from *Ktedonobacter racemifer* DSM 44963 (30 % over 142 residues and 21 % over 156 residues, respectively) and CAI78580 from uncultured candidate division OP8 bacterium (26 % over 170 residues). In all six cases the PspW-like protein is encoded downstream of a gene previously identified to encode a PspX-like protein (Figure 5.5). Subsequently, these PspW-like proteins were also used as a query in a BLASTP search against the NCBI database to reveal more PspW-like proteins. Sros_1171 from *S. roseum* DSM 43021, ZP_06162158 from *Actinomyces* sp. oral taxon 848 and Franean1_3992 from *Frankia* sp. EAN1pec were shown to have significant similarity to MibW. Proteins sharing a high percentage identity to ZP_06162159 from *Actinomyces* sp. oral taxon 848 include *E. lenta* DSM 2243 Elen_1456, *B. breve* UCC2003 Bbr_1070 and *B. longum*

subsp. *infantis* ATCC 15697 Blon_1403. These results have been incorporated into the schematic in Figure 5.5.

The location of *pspW* within the *psp* cluster gives a clue as to its possible function. *pspW* appears translationally coupled to the ECF sigma factor *pspX*. It has been noted that both groups linked to PspX, ECF01 and ECF33, contain characteristic membrane-anchored anti- σ factors (Staron *et al.* 2009). In accordance with this, TMHMM v.2 was used to predict the number of TM helices in each of these proteins (Krogh *et al.* 2001). PspW, Sros_1171, Krac_2613 and CAI78580.1 were predicted to have five TM helices, whereas Tbis_3433, MibW, Krac_2809, ZP_06162158 and Franean1_3992 were all predicted to have six (Figure 5.8). All nine proteins had at least 60 amino acids at the N-terminus that did not form part of a helix. The C-terminus of all nine proteins is cytoplasmic (22 residues in PspW) and could serve to bind the cognate sigma factor encoded by the neighbouring gene, sequestering it away from the RNA polymerase sigma factor. The proteins with six TM helices also have a cytoplasmic N-terminus, whereas those with only five TM helices have an extracellular N-terminus. However this interpretation relies on all TM helices being predicted successfully. The six TM helices of MibW are predicted to occur within the region from 74-244 residues, which coincides with the region showing significant similarity to PspW, implying PspW may instead form six TM helices. A comparison of 17 programs that can identify transmembrane regions concluded that TMHMM v2.0 was the most effective (Moller *et al.* 2001). However many other programs are available. The preferred model from TMpred predicted six TM helices in PspW with a cytoplasmic N-terminus (Hofmann and Stoffel 1993). Five helices were the same as predicted by TMHMM, but an additional helix was also predicted between residues 106 to 133 with the consequence that the N-terminus becomes cytoplasmic. Therefore it is more likely that the longer cytoplasmic N-termini serve to sequester the cognate sigma factor, but in any case it is likely that the sigma factor binding domain will be positioned at the same location in all PspW-like anti-sigma factors.

An alignment of all nine amino acid sequences did not indicate any conserved motifs that could be involved in interacting with the ECF sigma factors. The anti-sigma factor domain (ASD) was initially identified in the N-terminus of *Rhodobacter sphaeroides* ChrR and *E. coli* RseA (Campbell *et al.* 2007). The ASD consensus represents a structural fold conserved in an estimated 33 % of anti-sigma factors. The ASD is involved in contacting the cognate ECF sigma factor in an interaction which is mediated in response to specific environmental signals by a C-terminal domain (Campbell *et al.* 2007). However no such motif was identified in PspW-like proteins.

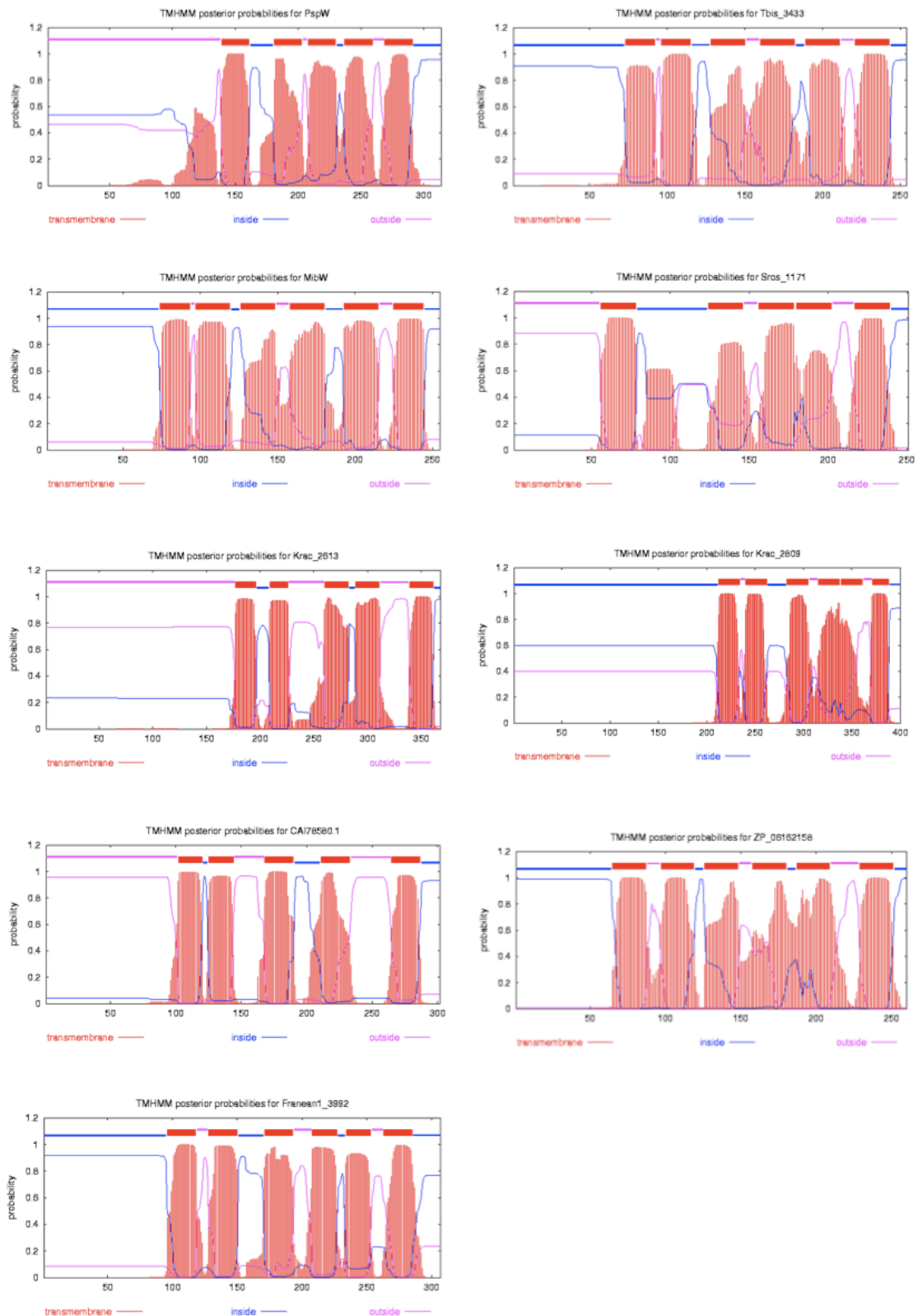


Figure 5.8 : A schematic showing the transmembrane helices predicted to occur in PspW and eight PspW-like proteins.

Namely Tbis_3433, MibW, Sros_1171, Krac_2613, Krac_2809, CAI78580.1, ZP_06162158 and Franean1_3992. Predictions were made using the full length protein sequences submitted to TMHMM (Krogh *et al.* 2001). Transmembrane helices are illustrated in red, cytoplasmic loops in blue and external loops in pink.

Although the membrane topology of PspW is not fully defined, it is likely that PspW does function to regulate the activity of PspX. The membrane location of PspW means that on binding PspX, it could sequester the sigma factor away from its promoter binding sites thus preventing transcription initiation. In accordance with characterised systems, an extracellular signal (such as cell wall stress) would then be detected either directly by the anti-sigma factor or be mediated through an additional protein, allowing the release of the sigma factor and expression of the relevant regulon. In group ECF01, which PspX bears most similarity to, the best characterised example is σ^W from *B. subtilis* (Staron *et al.* 2009). σ^W is regulated by the anti-sigma factor RsiW which is in turn regulated through proteolytic degradation (Schobel *et al.* 2004; Zellmeier *et al.* 2006). Indeed proteolytic degradation is a common mechanism for the release of the sigma factor upon receipt of the appropriate signal (Heinrich and Wiegert 2009). Alternatively, a conformational change can regulate the binding of the anti-sigma factor to its cognate sigma factor, such as with *S. coelicolor* σ^R . Here regulation occurs through a conformational change in the anti-sigma factor, the zinc metalloprotein RsrA, due to a thiol-disulphide redox switch (Paget *et al.* 2001). In addition, PspX activity could be regulated at the level of transcription or by the proteolytic processing of a pro- σ -factor. Examples of both of these mechanisms have been characterised in *S. coelicolor*. Transcription of *sigE* is regulated by the CseB/CseC two-component system (Hutchings *et al.* 2006). σ^{BldN} is synthesised as an inactive pro- σ factor which is subject to proteolytic processing to release active BldN (Bibb *et al.* 2000).

5.4.3 PspR

The 260 amino acid PspR contains the conserved domain pfam cd06170 (part of the superfamily cl10457) between residues 190 to 250. This corresponds to a helix-turn-helix motif responsible for DNA binding. Proteins belonging to this group include the LuxR-like transcriptional regulators. Within this family, there are both transcriptional activators and repressors. In *Vibrio fischeri*, N-acyl derivatives of homoserine lactone act as signalling molecules which bind to the N-terminal domain of LuxR, resulting in dimerisation and activation of gene expression through quorum-sensing (Stevens *et al.* 1994). In contrast, in *Sinorhizobium meliloti* it is phosphorylation of the N-terminal domain of FixJ which leads to multimerisation and the subsequent transcriptional activation of nitrogen fixation genes (Da Re *et al.* 1999). In PspR, there is an N-terminal domain which may be modified to modulate the DNA binding property of the C-terminal LuxR-like DNA binding domain.

PspR bears most similarity to MibR, the putative DNA-binding protein from *M. corallina* (33 % identity over 205 residues) (Table 5.1). Other proteins with similarity to PspR come from a range of organisms. The two-component response regulator YP_075969 from

Symbiobacterium thermophilum IAM 14863 shares 38 % identity over 82 residues. *S. thermophilum* depends on microbial commensalism so is currently uncultivable under laboratory conditions. The 3.57 Mb genome has been sequenced, revealing a GC content of 68.7 %, yet the genome bears most similarity to those in the Firmicutes phylum (Ueda *et al.* 2004). Aside from these two proteins, a majority of other alignments show similarity only within the putative helix-turn-helix domain of the protein. Thus PspR is a putative transcriptional regulator of the *psp* gene cluster.

5.5 Unknown proteins

5.5.1 PspJ

The PspJ protein has a significant alignment with just six proteins in the current NCBI database, all of which are annotated as hypothetical proteins. Four are from high-GC actinomycetes, namely Tbis_3437 from *T. bispora* DSM 43833, MibJ from *M. corallina*, Franean1_3993 from *Frankia* sp. EAN1pec and Sros_1170 from *S. roseum* DSM 43021. Two further hits are krac_2614 and krac_2810 from *K. racemifer* DSM 44963 in the Chloroflexi phylum. As only five PspJ-like proteins were identified, a Position-Specific Iterated BLAST (PSI-BLAST) was used as a variation on BLASTP in which the significant alignments are combined into a multiple alignment from which a position-specific score matrix is constructed and used to search the database for additional significant alignments. This whole process is then iterated as necessary to find other members of the same protein family (Altschul *et al.* 1990). Within three iterations, a hypothetical protein CAI78582.1 from uncultured candidate division OP8 bacterium (22 % over 125 residues) and a hypothetical protein Blon_1403 *B. longum* subsp. *infantis* ATCC 15697 (24 % over 194 residues) were also identified as part of the PspJ family. In addition, candidate PspJ proteins from other clusters subject to BLASTP searches against the NCBI database revealed similarity to other putative PspJ proteins. These were identified in *Actinomyces* sp. oral taxon 848 str. F0332, *B. breve* UCC2003, *E. lenta* DSM 2243 and *Eggerthella* sp. 1_3_56FAA. Figure 5.5 depicts the arrangement of these *pspJ*-like genes in species in which enough genome sequence is available to investigate the flanking regions. In nearly all cases the *lanJ* is adjacent to *lanY* in a cluster similar to *pspXWJYZ*. A schematic for the cluster in *Eggerthella* sp. is not shown. At the time of writing the cluster is spread over two contigs. Contig1.16 contains *lanJYZ* on the right hand border while Contig1.17 contains *lanX* on the left hand border so it is likely that a *lanW* occurs in between the two contigs.

Although BLASTP did not reveal any conserved domains, the server TMHMM predicted that PspJ contains four TM helices (Krogh *et al.* 2001). The same prediction was made for MibJ, which is proposed to reside as a membrane protein (Foulston and Bibb 2010). Other

PspJ-like proteins have varying numbers of predicted TM helices; Franean1_3993 and Sros_1170 are both predicted to form five TM helices, but Tbis_3437, Bbr_1070, ZP_06162159, Blon_1403 and Elen_1456 are all predicted to form six TM helices (Figure 5.9 A). While the number of helices varies, all proteins are predicted to contain a loop of approximately 20 residues in the second quarter of the protein. For PspJ this loop is predicted to be extracellular, however for all other PspJ-like proteins it is predicted to be cytoplasmic. An alignment of all nine proteins does not reveal a distinct motif, however L. Foulston observed the loop included a number of conserved charged and polar residues (aspartate, glutamate, glutamine, threonine and arginine) indicating that they could form a catalytic site or be involved in protein-protein interaction (Figure 5.9 B) (Foulston 2010). However, TMpred identified six possible TM helices in PspJ (Hofmann and Stoffel 1993). Five were identical to those predicted by TMHMM but there was an additional helix between residues 46 to 65. If PspJ is a six TM helix protein, then it will have a cytoplasmic loop between residues 65 to 87 like the other PspJ-like proteins.

The striking syntenous conservation of *lanX*, *lanW*, *lanY*, *lanZ* and *lanJ* in a wide range of actinomycete genomes, primarily from the orders Actinomycetales, Bifidobacteriales and Coriobacteriales, implies that these proteins function together in a common role. LanX is known to function as a sigma factor with LanW as its cognate anti-sigma factor (as described in Section 5.4). It is plausible that LanJ also functions in the mechanism of ECF factor regulation.

5.5.2 PspQ

pspQ encodes a 213 residue protein with highest percent identity to Tbis_3434 of *T. bispora* DSM 43833 (56 % over 129 residues) and MibQ of *M. corallina* (62 % over 103 residues) (Table 5.1). Other sequences producing significant alignments included Bbr_1073 of *B. breve* UCC2003 (35 % over 119 residues), Blon_1406 of *B. longum* subsp. *infantis* ATCC 15697 (40 % over 75 residues) and ZP_06162162.1 *Actinomyces* sp. oral taxon 848 str. F0332 38/110 (35 % over 110 residues). Thus in Figure 5.5 half of the clusters encode a PspQ-like protein in addition to *pspXWJYZ*. Out of the many proteins identified in BLASTP to have a significant alignment with PspQ, several were annotated as lipoproteins.

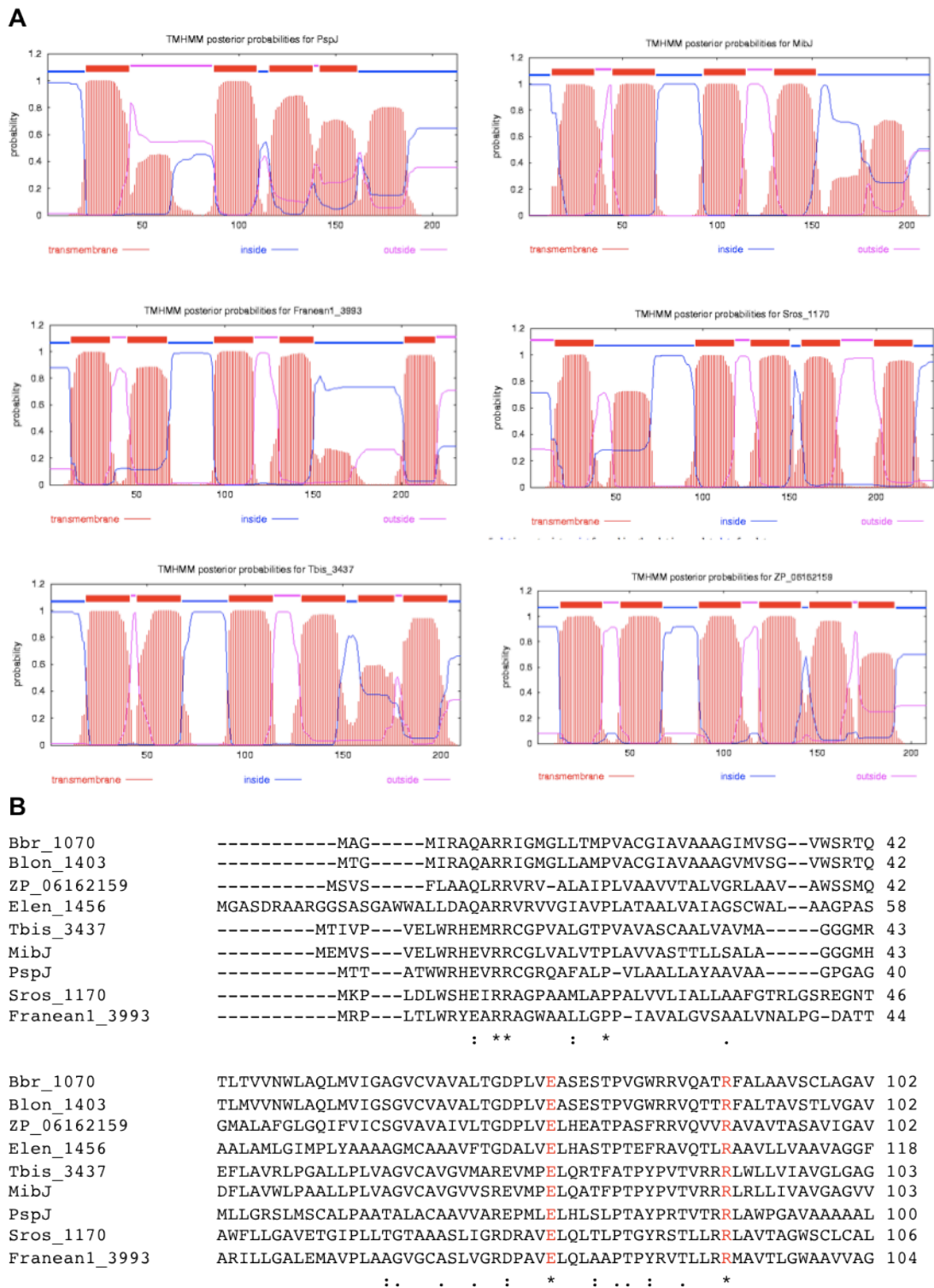


Figure 5.9 : Features of PspJ-like proteins.

A; Predicted transmembrane helices in PspJ-like proteins (using TMHMM (Krogh *et al.* 2001)). Proteins were predicted to contain either 4 (top; PspJ and MibJ), 5 (middle Franean1_3993 and Sros_1170) or 6 (bottom; Tbis_3437 and ZP_06162159) TM helices.

B; Sequence alignment (constructed using ClustalW2 (Chenna *et al.* 2003)) of the N-termini of PspJ-like proteins. The loop between residues 60-100 contains a conserved glutamate and arginine (in red).

Alignment of PspQ with the five PspQ-like proteins revealed a conserved lipobox motif in the putative signal peptide (Figure 5.10). Lipoproteins are modified at the N-terminus with an N-acyl diacyl glyceryl group (derived from phospholipids), which serves to anchor these proteins at the membrane-aqueous interface (Sutcliffe and Harrington 2002; Hutchings *et al.* 2009). Lipoproteins contain a lipobox motif which consists of an invariant cysteine residue in the +1 position) that is lipid-modified. The residues in positions -3, -2 and -1 fall into a consensus sequence of L/V/I, A/S/T/V/I, G/A/S. In addition, the N-terminal residues commonly contain positively charged residues while the region between the N-terminus and the lipobox consists of hydrophobic and uncharged residues. These features were used to identify likely lipoproteins using the DOLOP protein database (Madan Babu and Sankaran 2002). Figure 5.11 depicts how PspQ was identified as a putative lipoprotein by having 3 positively charged residues within the first 14 residues of the signal sequence, a hydrophobic stretch of 14 residues, followed by the lipobox LTAC. Tbis_3434 was identified by having 4 positively charged residues within the first 10 residues of the signal sequence, a hydrophobic stretch of 13 residues, followed by the lipobox VSGC. MibQ was identified by having 2 positively charged residues within the first 8 residues of the signal sequence, a hydrophobic stretch of 12 residues, followed by the lipobox LAGC. Bbr_1073, Blon_1406 and ZP_06162162.1* did not meet the criteria for a lipoprotein.

The SignalP 3.0 server was used to predict the likelihood that PspQ-like proteins have a signal peptide sequence (Emanuelsson *et al.* 2007). PspQ, Tbis_3434 and MibQ were all predicted to contain signal peptides with a probability ranging from 0.998-1.000, while Bbr_1073, Blon_1406 and ZP_06162162.1 were predicted to be non-secreted proteins (signal peptide probability <0.08 for all three proteins). The alignment in Figure 5.10 implies that ZP_06162162 may be annotated incorrectly, as the N-terminus appears truncated with the start codon only two bases before the lipobox motif. L. Foulston reannotated this ORF to include an additional 36 amino acids at the N-terminus and named the resulting protein sequence ZP_06162162* (Foulston 2010). However the SignalP server still predicted ZP_06162162* to be a non-secreted protein, with a signal peptide probability of just 0.201.

The LipoP 1.0 server was used to determine if the signal peptides detected above were likely to be lipoprotein signal peptides (Juncker *et al.* 2003). The server uses a Hidden Markov Model to report a score for the query sequence that corresponds to lipoprotein probability. For MibQ, the best prediction (with a score of 17.6393) is a lipoprotein with a cleavage site between amino acid 23 and 24 (LLLAG|CTGGG) as shown in Figure 5.10. For Tbis_3434 the best prediction (with a score of 10.4199) predicts cleavage between amino acids 30 and 31. But another prediction with a score just 0.65449 lower predicts a

signal peptide with cleavage between amino acids 26 and 27 (LLVSG|CAGAG). Although a lipoprotein signal peptide was predicted between amino acids 31 and 32 (LLLTA|CGGSG) in PspQ, the prediction had a score of just 2.87306. In contrast, Bbr_1073, Blon_1406 and ZP_06162162* were not predicted to be lipoproteins according to the algorithm used in LipoP.



Figure 5.10 : An alignment of PspQ with five PspQ-like proteins.

Namely Tbis_3434 (*Thermobispora bispora* DSM 43833), MibQ (*M. corallina*), Bbr_1073 (*Bifidobacterium breve* UCC2003), Blon_1406 (*Bifidobacterium longum* subsp. *infantis* ATCC 15697) and ZP_06162162.1 (*Actinomyces* sp. oral taxon 848 str. F0332). Alignment constructed using ClustalW (Chenna *et al.* 2003) with the conserved lipobox motif L₃-[A/S/T]₂-[G/A]₁-C₊₁ illustrated in red (Sutcliffe and Harrington 2002). The cleavage site predicted in Tbis_3434, MibQ and PspQ by LipoP is indicated (Juncker *et al.* 2003).

So although a multiple sequence alignment revealed a putative lipobox in the N-terminus of all six proteins (Figure 5.10), a more thorough investigation into other features commonly found in lipoproteins revealed that only PspQ, Tbis_3434 and MibQ are likely to be true lipoproteins (Figure 5.11). The C-terminus of PspQ contained the conserved domain pfam10709, which is a member of the superfamily cl11305 (Finn *et al.* 2008). Although this family is conserved in bacteria, its function is not known. Thus the role of the probable lipoprotein PspQ in planosporicin biosynthesis is unknown. However there is a precedent for the presence of a lipoprotein in lantibiotic gene clusters. The nisin and

subtilin gene clusters encode the lipoproteins Nisl and Spal, respectively (Figure 5.11). These proteins are anchored to the membrane through covalent N-terminal lipidation at the conserved cysteine of the lipobox motif after removal of the signal peptide. Once in place, they act as a mechanism of immunity by sequestering the respective lantibiotic to prevent interaction with lipid II (Qiao *et al.* 1995; Stein *et al.* 2003). The striking similarity between PspQ and MibQ may reflect the distinctive similarity in the structure of the target lantibiotics planosporicin and microbisporicin. In contrast the noted lack of strong sequence similarity among other LanI proteins is thought to reflect the different lantibiotic structures that each LanI protein must recognise (Kuipers *et al.* 1993a). It has been postulated that MibQ could function as a self-resistance mechanism for *M. corallina*, working independently to MibEF to intercept microbisporicin, preventing interaction with its cellular target (Foulston and Bibb 2010). Prominent examples of the function of other lipoproteins includes that of CseA, a lipoprotein which negatively regulates the ECF sigma factor σ^E in *S. coelicolor* (Hutchings *et al.* 2006). The location of *pspQ* near *pspXW* in the planosporicin gene cluster and a similar arrangement in the respective gene clusters in *M. corallina* and *T. bispora* implies that LanQ may instead play a role in the activation of the ECF sigma factor LanX.

	N-terminus	Hydrophobic region	Lipobox
PspQ:	MTDEHRRPATASAR	AAAAAPILAGAALL	LTAC
Tbis_3434:	MRTKSTACRR	GGVLAGLLVTALL	VSGC
MibQ:	MTNTTRAR	LSGAGLLAAALL	LAGC
Nisl:	MRR	YLILIVALIGITG	LSGC
Spal:	MFLKR	CVIVFGCFIVLFM	LSAC

Figure 5.11 : The tripartite nature of the signal sequence of lipoproteins.

Three PspQ-like proteins predicted to be lipoproteins (PspQ, Tbis_3434 and MibQ) and two characterised lipoproteins from lantibiotic clusters are shown. The individual parts in the signal sequence are separated into the positively-charged N-terminus, hydrophobic region and lipobox. Charged amino acids in the N-terminus are in red. In all three examples 12-14 residues separate the lipobox and the charged residues. The lipobox is in blue with the consensus sequence L/V/I, A/S/T/V/I, G/A/S.

5.5.3 PspV

PspV is a 438 amino acid protein which bears closest resemblance to two hypothetical proteins from the Actinomycetales; MibV from *M. corallina* and Ndas_4690 from *Nocardiopsis dassonvillei* subsp. *dassonvillei* DSM 43111 (Table 5.1). In both of these

organisms, the *lanV* appears to be located within a lantibiotic gene cluster (figure 5.12). MibV is encoded within the operon *mibABCDTUV* containing genes essential for microbisporicin biosynthesis, however the function of MibV is unknown (Foulston and Bibb 2010). Ndas_4690 is encoded in an operon of eight genes that correspond to an as-yet uncharacterised lantibiotic gene cluster (Figure 5.12). This operon contains a putative prepropeptide Ndas_4696, which contains the motif 'INLD', a variation on the FNLD motif found in the leader peptides of type AI lantibiotics. However the putative propeptide does not have a significant alignment with any characterised lantibiotics, indicating a novel structure. Downstream of the prepropeptide gene are genes encoding two typical type AI modification enzymes; a lantibiotic dehydratase (Ndas_4694) and a lantibiotic cyclase (Ndas_4693), followed by genes encoding an ABC transporter (Ndas_4692), a LuxR-type regulator (Ndas_4688) and several other possible lantibiotic modification enzymes (Figure 5.12). Unlike other ABC transporters from characterised lantibiotic clusters from high-GC organisms, Ndas_4692 is a single protein encoding two permease and two ATPase domains in an alternating pattern. Both permease domains are predicted to contain six TM helices in the TMHMM server v2.0. Three other proteins are annotated as an O-methyltransferase (Ndas_4691), a cytochrome P450 (Ndas_4689) and a luciferase-like monooxygenase (Ndas_4695). There is a precedent for each of these enzymes in lantibiotic clusters. In the *cyp* cluster, *cypM* encodes an S-adenosyl methionine (SAM)-dependent methyltransferase presumed to be required for the methylation of the N-terminal alanine of cypemycin (Claesen and Bibb 2010). Cytochrome p450 enzymes typically function as mono-oxygenases, but in the *mib* cluster, MibO is proposed to be responsible for the dihydroxylation of proline (Foulston and Bibb 2010). The monooxygenase, GarO, from the *act* gene cluster of *A. garbadinensis* is proposed to oxidize the sulphur of the C terminal MeLan bridge, forming the sulfoxide residue of actagardine (Boakes *et al.* 2009).

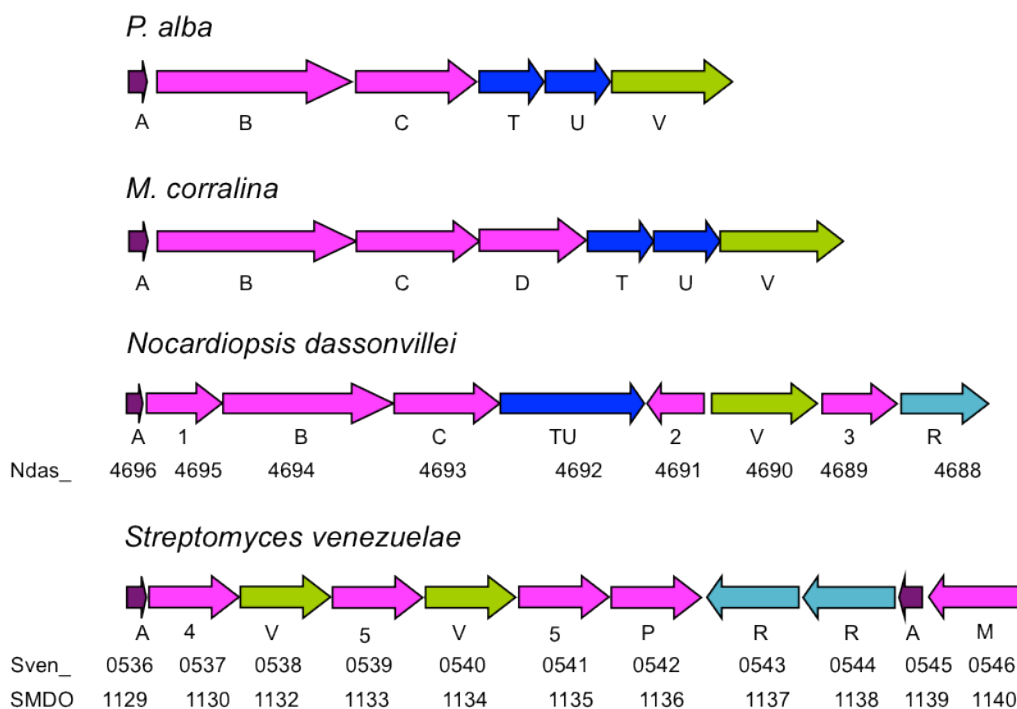


Figure 5.12 : A schematic showing the context of genes encoding proteins with significant similarity to PspV in the genomes listed.

Genes are coloured and labelled according to the respective genes from the *psp* cluster, with the locus tag given below. Genes labelled according to nomenclature from other clusters are D; flavin-dependent decarboxylase, P; serine protease and M; dehydratase/cyclase. Genes labelled with numbers are 1; monooxygenase, 2; O-methyltransferase, 3; cytochrome P450, 4; glycosyltransferase, 5; 'radical SAM' protein, 6; 'major facilitator' protein.

Along with MibV and Ndas_4690, proteins from three strains of *Bacillus amyloliquefaciens* make up the top five proteins with the most significant alignment to PspV. Further PspV-like proteins were predominantly from the *Streptomyces* genus. Namely *S. venezuelae*, *Streptomyces ambofaciens*, *Streptomyces viridochromogenes*, *Streptomyces roseosporus* and *S. griseus*. These are all annotated as hypothetical proteins with no conserved domains, however several are encoded in clusters that include a small hypothetical peptide. One of these clusters also contains a typical lantibiotic modification enzyme. *S. venezuelae* contains two proteins with similarity to PspV; SMD01132/Sven_0540 (34 % over 196 residues) and SMD01134/Sven_0538 (27 % over 187 residues). These occur in an operon along with SMD01129/Sven_0536, a putative 39 residue peptide (Figure 5.12). The adjacent operon encodes a putative 66 residue peptide (SMD01139/Sven_0545), along with a putative LanM bifunctional dehydratase/cyclase (SMD01140/Sven_0546), indicating this to be a type AII or B lantibiotic cluster. This LanM had previously

been identified as a cryptic lantipeptide cluster through bioinformatic analysis of the *S. venezuelae* genome, sequenced by Diversa Corp. (Claesen 2010). Analysis of the LanM (SMD01140/Sven_0546) revealed the Zinc-binding site residues were conserved, but the catalytic His residues were missing. The adjacent LanA (SMD01139/ Sven_0545) contains a serine and several threonines that may be dehydrated by the LanM. Notably, the LanA contains only one cysteine which is located at the C-terminus so may form an AviCys (although there is no LanD homologue) as observed in linaridins, removing the need for an active LanM cyclase domain. The function of this LanM has previously been investigated through deletion analysis in *S. venezuelae* and heterologous expression in *S. lividans*. The deleted strain showed no difference in antibacterial activity against *M. luteus* compared to WT *S. venezuelae* (likely producing chloramphenicol or jadomycin) (Claesen 2010). Chemical induction with GlcNAc prompted a zone of inhibition from *S. lividans* heterologously expressing cosmid 1-B5 from the *S. venezuelae* cosmid library (insert contains the LanM cluster and the PspV-like proteins), however this phenotype was not reproducible (Claesen 2010). *S. venezuelae* microarray data revealed that neither of these operons are induced during development or under stress simulated through sub-inhibitory levels of daptomycin treatment (data not shown).

The association of PspV-like proteins MibV, Ndas_4690 and Sven_0538/0540 with lantibiotic gene clusters suggests they may have a novel role in lantibiotic biosynthesis. The fact that both type AI and type AII/ B clusters contain putative *lanV* genes implies they play a generic role perhaps as a scaffold or chaperone to aid the binding of modification or processing enzymes to the prepropeptide.

5.5.4 ORF-10 and ORF-11

The ORF located at position -10 from the prepropeptide *pspA* encodes a 102 amino acid protein. This protein has very little significant similarity to proteins in the current NCBI database. Interestingly, two of the only proteins to bear resemblance to the protein encoded by *orf-10* are annotated as protease Do from *Ochrobactrum intermedium* LMG 3301 and *Ochrobactrum anthropi* ATCC 49188 (32 % and 33 % over 79 residues, respectively) (Table 5.1). *orf-10* is in a likely operon of two genes. The ORF located at position -11 from the prepropeptide *pspA* encodes a 78 amino acid protein with no conserved domains and bears similarity only to hypothetical proteins.

Protease Do, or DegP, is a periplasmic serine protease required for *E. coli* survival at high temperatures. However, all LanP enzymes demonstrated to cleave the leader peptide to release the mature lantibiotic so far share homology with the serine protease subtilisin. LanP proteases are encoded by several low-GC lantibiotic gene clusters (e.g. NisP for

nisin; EpiP for epidermin) as described in Chapter 1, but so far no actinomycete lantibiotic gene cluster is known to contain a separate *lanP*. It is proposed that a protease encoded elsewhere in the genome fulfils the role of removing the leader peptide. This has been demonstrated for subtilin, which uses an extracellular serine protease to cleave the leader peptide, releasing the mature lantibiotic. In contrast, type AI LanT enzymes contain an extra N-terminal peptidase domain suggesting a role in processing of the leader peptide concomitant with export. Section 5.4 suggested that the release of the ECF sigma factor PspX from the cognate anti-sigma factor PspW could be mediated through proteolysis. If so, it is possible that the putative protease ORF-10 together with ORF-11 may encode proteins which fulfill these roles.

5.6 Synteny

Section 5.3, 5.4 and 5.5 identified proteins from the NCBI database that had significant similarity to proteins encoded by the planosporicin biosynthetic gene cluster. Many of these proteins occurred in similar gene clusters. Most notably, there was remarkable similarity to the microbisporicin gene cluster from *M. corallina*. However the lantibiotic biosynthetic enzymes LanB and LanC found in *B. clausii* KSM-K16 and *Actinomyces* oral taxon 848 also bore a resemblance to those of the *psp* cluster. This implies that planosporicin may be one of several microbisporicin-like lantibiotics forming a distinct subfamily in the type AI group. The peptide sequences producing significant alignments to the planosporicin prepropeptide, PspA, were discussed in detail in Chapter 3. In brief, most similarity was seen with the microbisporicin prepropeptide (MibA) from *M. corallina* (67 % over 46 residues) and the conserved domain protein, actoracin (ZP_06162152.1), from *Actinomyces* sp. oral taxon 848 str. F0332 (55 % over 42 residues) (Table 5.1). The third annotated protein with significant similarity was the putative antibiotic protein megateracin (BMQ_pBM70151) from *B. megaterium* QM B1551 (45 % over 47 residues). A search of the literature revealed planosporicin-like peptides in *B. clausii* (clausin) and *Kineosporia* sp. (kineosporiacin). To investigate this further, the context of the genes encoding these planosporicin-like prepropeptides was examined in order to determine the extent of similarity between the gene clusters.

5.6.1 *mib* cluster

Chapter 3 described the similarity between the lantibiotics planosporicin and microbisporicin. The annotation of the planosporicin biosynthetic gene cluster revealed it bears a striking resemblance to the microbisporicin cluster. Nearly every gene of the *psp*

cluster has a counterpart in the *mib* cluster (Table 5.1). The similarity extends from a sharing a high percentage identity through to a striking level of synteny. This is best demonstrated through use of the Artemis Comparison Tool (Figure 5.13). The *lanABCTUV* operon is almost identical in both clusters, with the addition of *mibD* between *mibC* and *mibT*. Adjacent to *lanA*, both clusters contain divergently transcribed *lanXW*. Both clusters also contain the operons *lanJYZQ* and *lanR*. However the *mib* cluster also contains a *mibO* gene between *mibZ* and *mibQ*. Additionally there is a reorientation of these two operons between the clusters. The immunity genes *lanEF* have the biggest change in location, being downstream of *pspR* and upstream of *mibV*. The 20 genes of the *mib* cluster include five not observed in the *psp* cluster, namely *mibD*, *mibO*, *mibH*, *mibS* and *mibN*. These genes encode enzymes catalysing post translational modifications to the microbipsoricin prepropeptide which are not observed in planosporicin. MibD catalyses the incorporation of the AviCys residue at the C-terminus, MibO is a candidate for the catalysis of the dihydroxylation of proline and MibHSN are proposed to catalyse the chlorination of tryptophan (Foulston and Bibb 2010).

5.6.2 *Actinomyces* sp. cluster

Chapter 3 described an as-yet uncharacterised putative lantibiotic from an *Actinomyces* sp. that was named actoracin (Foulston 2010). The genome of *Actinomyces* oral taxon 848 strain F0332 was isolated from the human mouth and sequenced as part of the human microbiome project. L. Foulston identified the gene cluster corresponding to actoracin biosynthesis in nucleotide region 154000-170000 of the genome sequence (Figure 5.14) (Foulston 2010). As expected for a lantibiotic cluster, the prepropeptide was encoded adjacent to the biosynthetic genes *lanB*, *lanC* and *lanD* (described in Section 5.3). Additional similarity to the *psp* and *mib* clusters was evident through a neighbouring cluster containing conserved *lanXWJYZ* (described in Sections 5.3, 5.4 and 5.5). Unusually for a lantibiotic cluster, there is just one ABC transporter, LanYZ. This could function to both export the lantibiotic and confer immunity on the producing strain. Alternatively this function could be provided through genes encoded elsewhere in the genome. If this is not the case, then it may be that the gene cluster encodes a lantipeptide without anti-microbial function that may instead act as a signalling molecule.

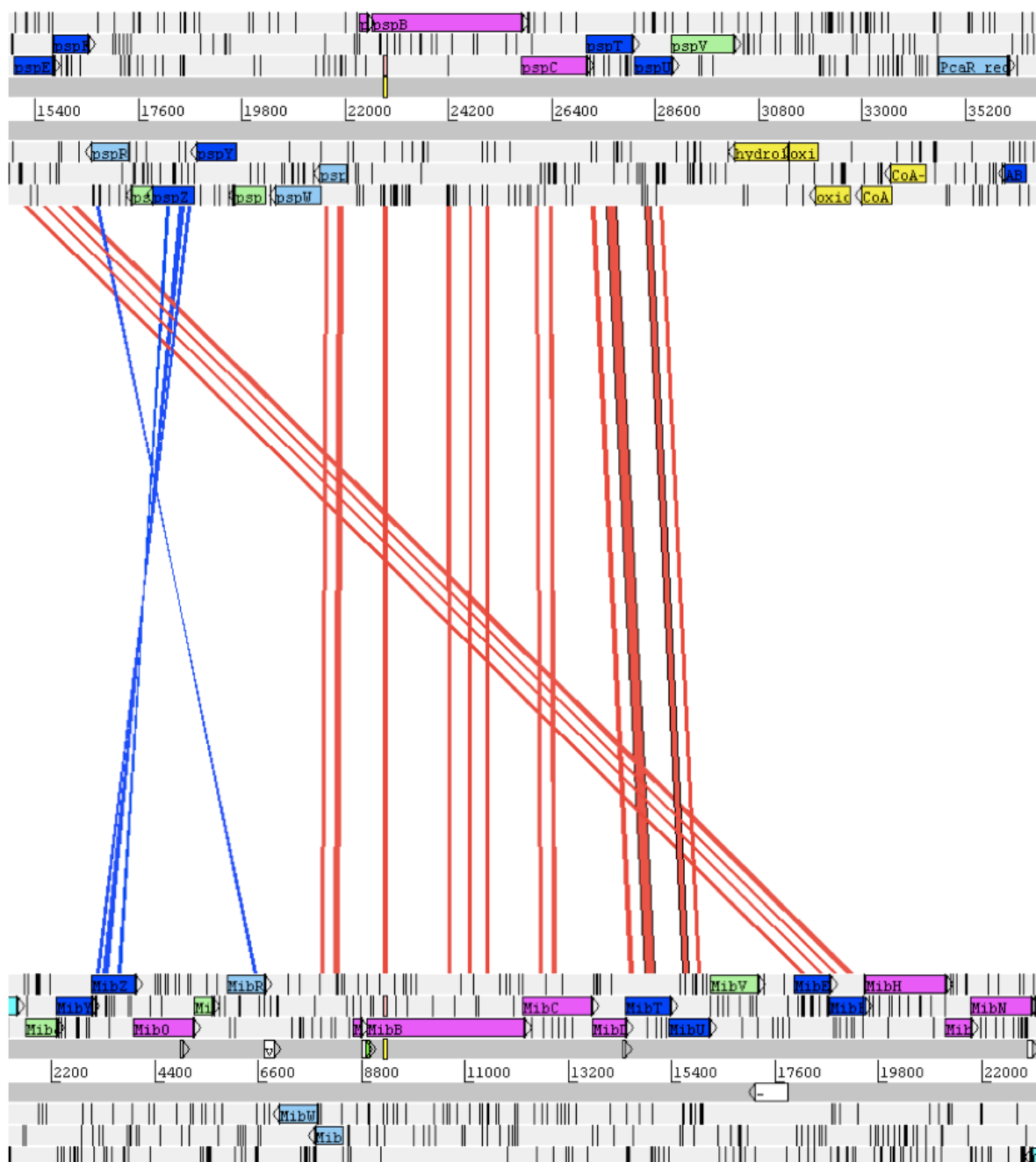


Figure 5.13 : A comparison of the *psp* and *mib* clusters using the Artemis Comparison Tool (ACT; Sanger).

The alignment used the insert DNA sequence from pIJ12321 (*P. alba*) and pIJ12125 (*M. corallina*) with minimum cut-off 40 and maximum cut-off 210 and is cropped to show only the biosynthetic gene clusters. The insert DNA sequence from pIJ12125 was kindly provided by L. Foulston.

5.6.3 *Bacillus* sp. clusters

Chapter 3 also identified planosporicin-like prepropeptides in several species of *Bacillus*. *B. megaterium* QM B1551 encodes a prepropeptide likely to be expressed as a lantibiotic termed megateracin. Interestingly, two *lanA* genes encoding the same 46 amino acid peptide (BMQ_pBM70152 and BMQ_pBM70151) are indicated in the current annotation of *B. megaterium* QM B1551 (Figure 5.14). Immediately downstream is a lantibiotic dehydratase, an ABC transporter, lantibiotic cyclase and a flavoprotein. The ABC transporter BMQ_pBM70149 contains both an ATPase and a permease domain, typical of a low-GC Gram-positive LanT. Further downstream are two further ATP-binding components of ABC transporters and three hypothetic proteins which TMHMM reveals are likely to be the permease component, none of which have significant similarity to the ABC transporters of the *psp* and *mib* clusters. Other features typical of a low-GC gene cluster include a two-component system comprising a DNA-binding response regulator (BMQ_pBM70141) and a sensory histidine kinase (BMQ_pBM70140).

B. clausii has been described as the producing strain of the lantibiotic clausin (Bouhss *et al.* 2009). Although this prepropeptide has not been annotated, the lantibiotic modification enzymes downstream have been. The *lanBCD* operon (ABC_3559-3557) is typical of a lantibiotic gene cluster. The LanBC enzymes have a high percentage identity with PspBC (as described in Section 5.3). Likewise, all three enzymes are similar to MibBCD (Foulston 2010). However further downstream this similarity ends. Regulation is likely mediated through a two-component histidine kinase-response regulator pair. Thus the *B. clausii* cluster resembles a typical low-GC type AI cluster.

This theme is continued through two more *Bacillus* clusters identified due to sharing a significant alignment with the enzymes PspB and PspC. Gene clusters found in *B. mycoides* and *B. cereus* are similar to the *psp* and *mib* gene clusters only with respect to the lantibiotic modification enzymes (Figure 5.14). The method of export, immunity and regulation all appear very similar to low-GC Gram-positive type AI clusters.

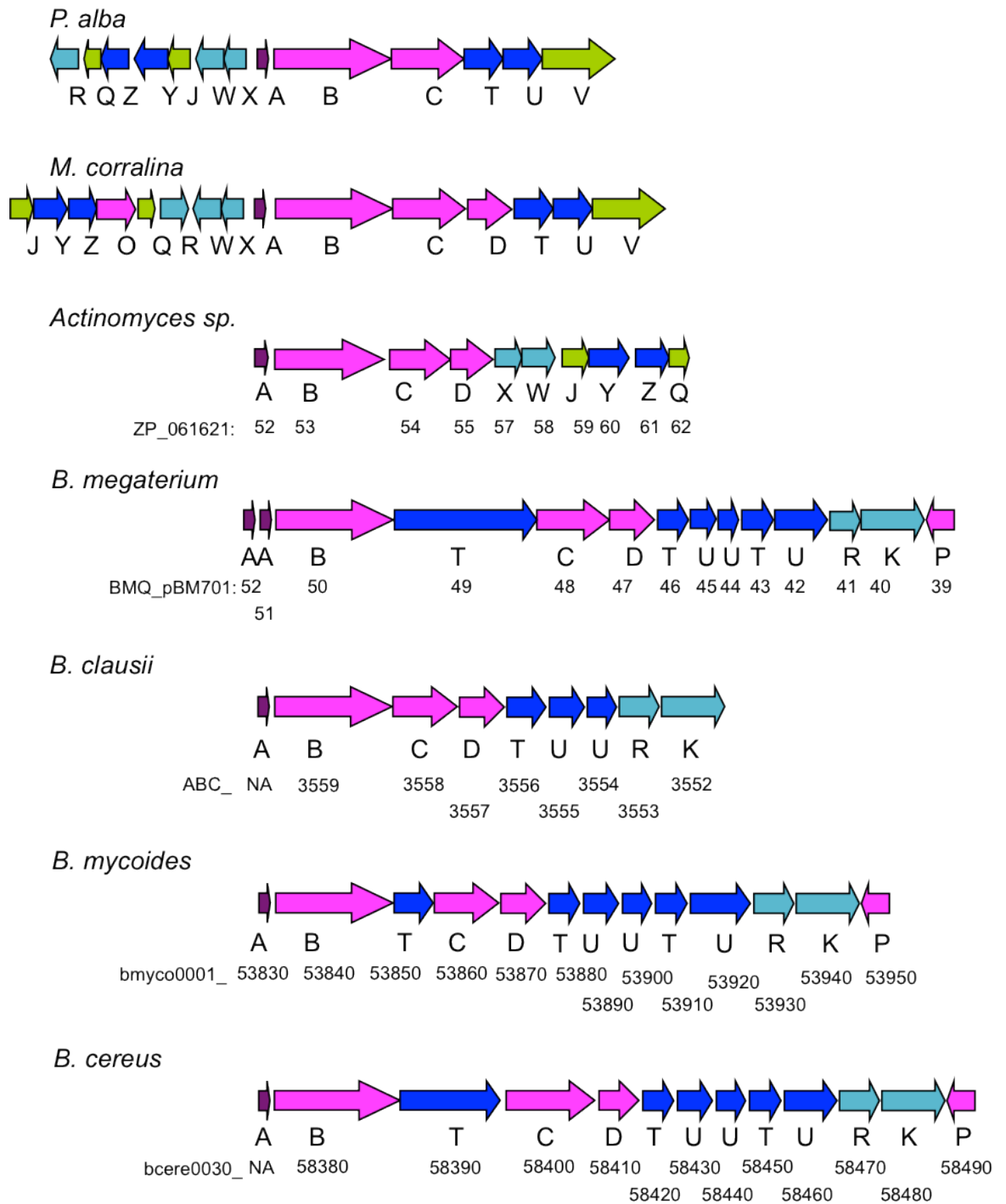


Figure 5.14 : A schematic showing the arrangement of clusters encoding proteins identified to have significant similarity to PspABC in the genomes listed.

Genes are coloured and labelled according to the respective genes from the *psp* cluster, with the locus tag given below. Genes labelled according to nomenclature from other clusters are O; cytochrome P450, D; flavin-dependent decarboxylase, P; serine protease and K; sensor histidine kinase.

5.7 Discussion

Chapter 3 characterised planosporicin as a type A lantibiotic due to its elongated structure. Chapter 4 identified the gene cluster responsible for planosporicin biosynthesis and this Chapter completes the annotation and bioinformatic analysis of this cluster. The *psp* cluster contains the modification enzymes *pspB* and *pspC* downstream of *pspA*, and so can now be characterised as a member of class I lantibiotics due to the use of two individual enzymes to incorporate (Me)Lan bridges instead of one bifunctional *lanM*. The cluster contains three ABC transporters; *pspTU*, *pspYZ* and *pspEF*. All consist of one permease component (predicted to contain six TM helices) and are likely to form a homodimer in the membrane. The ATPase domains are predicted to form a homodimer that interacts with the cytoplasmic interface of the permease dimer. PspF is distinctive in containing an E-loop instead of the canonical H-loop seen in PspT and PspZ. The E-loop implies that PspEF forms the immunity transporter conferring self-protection on *P. alba* from its own lantibiotic. The location of *pspTU* downstream of *pspC* implies this transporter functions to export the modified lantibiotic out of the cell. The presence of a third ABC transporter, PspYZ, is unusual in lantibiotic clusters and a function is not yet known. There are a couple of putative regulatory mechanisms encoded by the cluster. *pspX* encodes a sigma factor and is translationally coupled to its proposed cognate anti-sigma factor gene *pspW*. Additionally there is a transcriptional regulator encoded by *pspR*, which may function as an activator or repressor of *psp* transcription. Unusually for an actinomycete lantibiotic cluster, there is a candidate protease encoded by *orf-10* which may be responsible for cleavage of the leader peptide to release the mature lantibiotic. *orf-11* appears to be located in the same operon, but its function is unknown. Three further genes within the *psp* cluster also have as-yet unknown functions. PspV aligns only to hypothetical proteins, PspJ is likely to be a membrane protein and PspQ is a putative lipoprotein.

This Chapter detailed the annotation of the planosporicin biosynthetic gene cluster on the basis of percent identity to known lantibiotic biosynthetic genes. Overall, the planosporicin gene cluster is most similar to the cluster for microbisporicin, both of which originate from genera in the same family, *Streptosporangiaceae*. However the six genes, *pspXWJYZQ*, also bear a resemblance to several smaller clusters in the order Actinomycetales; *Thermobispora*, *Streptosporangium*, *Frankia* and *Actinomyces*. Two other orders also contain clusters that encode proteins with a significant alignment; Bifidobacteriales (*Bifidobacterium*) and Coriobacteriales (*Eggerthella*). Likewise, when looking at clusters containing a protein with a high level of similarity to PspV, two more clusters from the Actinomycetales were observed, one from the genus *Nocardiopsis* and another from *Streptomyces*.

Percent identity is a factual measurement, whereas homology is a hypothesis supported by evidence. The fact that all of these proteins sharing a high percentage of identity occur within the same phylum provides some evidence to suggest that they may be homologous. However it cannot be proven that the sequence similarity is due to common ancestry. It is perhaps more likely that these proteins resulted from horizontal gene transfer between two organisms, so should be termed xenologues. The function of xenologues tends to be similar although this can depend on how substantial the change in context was for the transferred gene.

The synteny observed between the planosporicin and microbisporicin gene clusters implies they may be related through horizontal gene transfer. The propagation of this lantibiotic biosynthetic gene cluster is possible due to the selective advantage conferred on the new host. Natural selection drives chemical innovation, as subsequent to transfer, diversification is common, resulting in a multiplication of the members of that natural product family (Fischbach *et al.* 2008). The differences observed between the *psp* and *mib* clusters must have occurred reasonably quickly and produced intermediates that justified keeping the cluster. The changes observed in the *psp* and *mib* clusters occurred within individual genes through mutation so that the amino acid percent identity ranges from 31 % to 67 % (Table 5.1). More drastic changes to the number of genes resulted in those encoding enzymes catalysing additional modifications of microbisporicin such as the amino-vinyl cysteine, chlorination of tryptophan and dihydroxylation of proline being subsequently deleted by the *psp* cluster or acquired by the *mib* cluster.

When investigating proteins with a significant alignment to the planosporicin prepropeptide, the species identified are more diverse. Some are Actinobacteria (*Microbispora* and *Actinomyces*) while others are Firmicutes (*Bacillus*). Similarity through vertical descent is unlikely, while horizontal transfer is possible. However convergent evolution could instead select for a structure which has a particularly effective mode of action (Fischbach 2009).

Although function based on percent identity gives a good indication of likely roles in planosporicin biosynthesis, further confirmation is needed through genetic manipulation of the cluster. Chapter 6 uses heterologous expression of the *psp* cluster to define the minimal gene set within pIJ12321 for planosporicin biosynthesis. Chapters 7 and 8 describe numerous genetic knockouts in *P. alba* with the aim of elucidating the function of individual *psp* genes.

5.8 Summary

PspA is the prepropeptide of planosporicin.

PspB is a putative lantibiotic dehydratase proposed to dehydrate serine residues to form Dha and threonine residues to form Dhb.

PspC is a putative lantibiotic cyclase proposed to activate the thiol moiety of cysteine residues to perform a nucleophilic attack on Dha/Dhb residues in the nascent peptide chain to form (Me)Lan bridges.

PspT (ATPase) and PspU (permease) together form an ABC transporter proposed to be responsible for secretion of planosporicin.

PspY (permease) and PspZ (ATPase) together form an ABC transporter of unknown function.

PspE (permease) and PspF (ATPase) together form an ABC transporter proposed to confer immunity to planosporicin on the producer, *P. alba*.

PspX is likely to be an ECF σ -factor involved in the regulation of planosporicin expression.

PspW is translationally coupled to PspX and likely acts as its cognate anti- σ factor.

PspR is a LuxR-like protein due to a helix-turn-helix domain at the C-terminus so is proposed to be involved in regulation of the *psp* cluster.

PspJ is a hypothetical protein predicted to contain several transmembrane helices.

PspQ has a predicted signal peptide sequence containing a lipobox motif LTAC corresponding to the conserved motif of lipoproteins.

PspV is a hypothetical protein with similarity to proteins in clusters containing either lanB/C or lanM biosynthetic enzymes.

pspXWJYZQ encode proteins with most homology to the products of gene clusters with the same synteny. Some of these gene clusters also encode lantibiotic biosynthetic genes.

The *psp* gene cluster bears a striking resemblance to the *mib* gene cluster of *M. corallina* encoding the related lantibiotic, microbisporicin.

Chapter 6 : Heterologous production of planosporicin

6.1 Introduction

Chapter 5 provided a detailed bioinformatic analysis of the putative planosporicin biosynthetic gene cluster. The next step in the analysis of planosporicin biosynthesis was the expression of the *psp* gene cluster in a heterologous host. The reasons for this were two-fold. On a theoretical level, heterologous production would provide demonstrable proof that the cloned fragment contained all of the genes required for planosporicin production. On a practical level, expression of the *psp* cluster in a genetically tractable host would provide a platform for later deletion analysis. Often the development of efficient protocols for the genetic manipulation of a relatively uncharacterised natural producer is time consuming and may give unsatisfactory results. Thus the *in vivo* manipulation of the biosynthetic gene cluster in a heterologous host could provide an alternative route to determine the function of the constituent genes.

At the time this work was undertaken, there was no published method for the genetic manipulation of the genus *Planomonospora*. Indeed only three genera of the *Streptosporangiaceae* had been shown to be genetically manipulable, namely *Nonomuraea* (Stinchi *et al.* 2003), *Planobispora* (Beltrametti *et al.* 2006) and *Microbispora* (Foulston and Bibb 2010).

In contrast the genus *Streptomyces* is well studied and several species are widely used laboratory strains that are easy to manipulate genetically. Furthermore, there are several examples of the successful expression of secondary metabolite biosynthetic gene clusters from non-streptomycete actinomycetes heterologously in *Streptomyces*. For example, the capreomycin NRPS gene cluster from *Saccharohrix mutabilis* subsp. *capreolus* was integrated into the chromosome of *S. lividans* yielding of 50 mg/ml capreomycin without modifying the regulation of the cluster (Felnagle *et al.* 2007).

Moreover, there were four examples of the expression of high-GC lantibiotic pathways in heterologous hosts at the time when this work was attempted. Three moved a gene cluster between different *Streptomyces* species. Transfer of the cinnamycin cluster from *S. cinnamoneus cinnamoneus* DSM 40005 to *S. lividans* resulted in heterologous production of cinnamycin (Widdick *et al.* 2003). Transfer of the cypemycin lantipeptide

gene cluster from *Streptomyces* sp. OH-4156 to *S. venezuelae* and *S. coelicolor* M1146 resulted in heterologous production of cypemycin (Claesen and Bibb 2010). And transfer of the grisemycin lantipeptide gene cluster from *Streptomyces griseus* to *S. coelicolor* M1146 resulted in heterologous production of grisemycin (Claesen and Bibb 2011).

There was also one example where a high-GC lantibiotic gene cluster had been expressed in a different genus. The actagardine gene cluster was transferred from *Actinoplanes garbadinensis* into *S. lividans* 1326, resulting in heterologous production of actagardine (Boakes *et al.* 2009). The same approach was used to produce actagardine in *S. lividans* TK24 and *S. coelicolor* M1146 (Bell 2010). Under the current classification system, the genus *Actinoplanes* is in the family *Micromonosporaceae* while the genus *Streptomyces* in the *Streptomycetaceae* (Goodfellow *et al.* 1990). Thus this latter work set a precedent for the successful transfer of a high-GC lantibiotic gene cluster between families. This Chapter describes the mobilisation of the planosporicin biosynthetic gene cluster into various heterologous hosts and the attempts made to detect production of the mature, exported lantibiotic.

6.2 Expression of the planosporicin gene cluster in *Streptomyces* species

6.2.1 Mobilisation of cosmids

pIJ12321 (cosmid B4-1), identified in Chapter 4 and annotated in Chapter 5, appeared to contain the entire *psp* gene cluster. In contrast, pIJ12322 (cosmid F13-1) appeared to lack the 3' 233 bp of the 1317 bp *pspV* coding sequence at the right hand border of the proposed cluster. To assess whether *pspV* was required for planosporicin production, each cosmid was integrated at the phage attachment site (*attB*) of a heterologous host. The stable integration of a construct into the chromosome of the host posed several advantages, not least that it obviated the need for continued antibiotic selection, which is of obvious benefit when assaying for the production of an antimicrobial compound.

pIJ10702 (also known as pMJCOS1) is a derivative of the SuperCosI vector used to construct the *P. alba* cosmid library in Chapter 4. In pIJ10702, the kanamycin resistance gene (*aph*) of SuperCosI has been replaced by an *oriT-attP-int-aac(3)IV* cassette (Yanai *et al.* 2006). This cassette was isolated by digesting pIJ10702 with *SspI* and recovering the 5.2 kb fragment from an agarose gel. The SuperCos backbone of both pIJ12321 and pIJ12322 was then PCR-targeted with the *SspI* fragment from pIJ10702. Double-crossover recombination replaced the *aph* gene of the SuperCos backbone with the *oriT-attP-int-aac(3)IV* cassette through sequence homology in the regions flanking both the resident resistance gene and the incoming cassette. The *aac(3)IV* gene conferred

resistance to apramycin, which allowed selection of the replacement event in *E. coli*, and which could also be used subsequently to select for transfer of the cosmid to actinomycete hosts. The *oriT* is an *E. coli* origin of transfer that would subsequently allow the cosmid to be conjugated into an actinomycete host (Kieser *et al.* 2000). The attachment site (*attP*) and integrase gene (*int*) of phage Φ C31 would permit stable integration of the cosmids into the chromosome of the heterologous host (Bierman *et al.* 1992). The recombination event to generate the targeted pIJ12321 and pIJ12322 derivatives was carried out in *E. coli* BW25113 harboring pIJ790 as described in Chapter 2. Transformation of bacteria with linear DNA is normally prohibited due to the bacterial intracellular exonucleases that degrade linear DNA. However, pIJ790 encodes three genes (*gam*, *bet* and *exo*) that confer the bacteriophage λ Red recombination functions. Exo is an 5' to 3' double-strand (ds)DNA-dependent exonuclease. It degrades linear dsDNA to leave long, 3' single strand (ss)DNA overhangs (Court *et al.* 2002). Bet is a ssDNA binding protein that protects the overhangs from single-strand nuclease attack and promotes annealing between complementary ssDNA targets (Karakousis *et al.* 1998). Gam functions in parallel to inhibit the host nucleases, RecBCD and SbcCD. Thus linear dsDNA introduced into the cell is protected from degradation, promoting a greatly enhanced rate of recombination (Gust *et al.* 2003). The constructs were subsequently moved into *E. coli* DH5 α to verify the absence of gross rearrangements within the *P. alba* insert DNA via restriction analysis. Targeting of pIJ12321 created pIJ12323 (Figure 6.1), while targeting of pIJ12322 yielded pIJ12324.

Many *Streptomyces* species, such as *S. coelicolor*, carry a methyl-specific restriction system as a mechanism for the destruction of heterologous DNA. Thus the successful introduction of constructs requires prior passaging through the non-methylating *E. coli* ET12567 containing the non-transmissible RP4 derivative pUZ8002 (Gust *et al.* 2003). This enables RP4 *tra*-mediated mobilisation of the cosmid *in trans* into *Streptomyces* by intergeneric conjugation. *S. lividans* does not contain a methyl-sensing restriction system, so alternative, more robust *E. coli* strains harboring pUZ8002 can be used instead as the donor strain. However, for simplicity, pIJ12323 and pIJ12324, as well as a pIJ10702 vector-only control, were introduced into *E. coli* ET12567/pUZ8002, which was then used to conjugate the constructs into both *S. lividans* TK24 and *S. coelicolor* M1146.

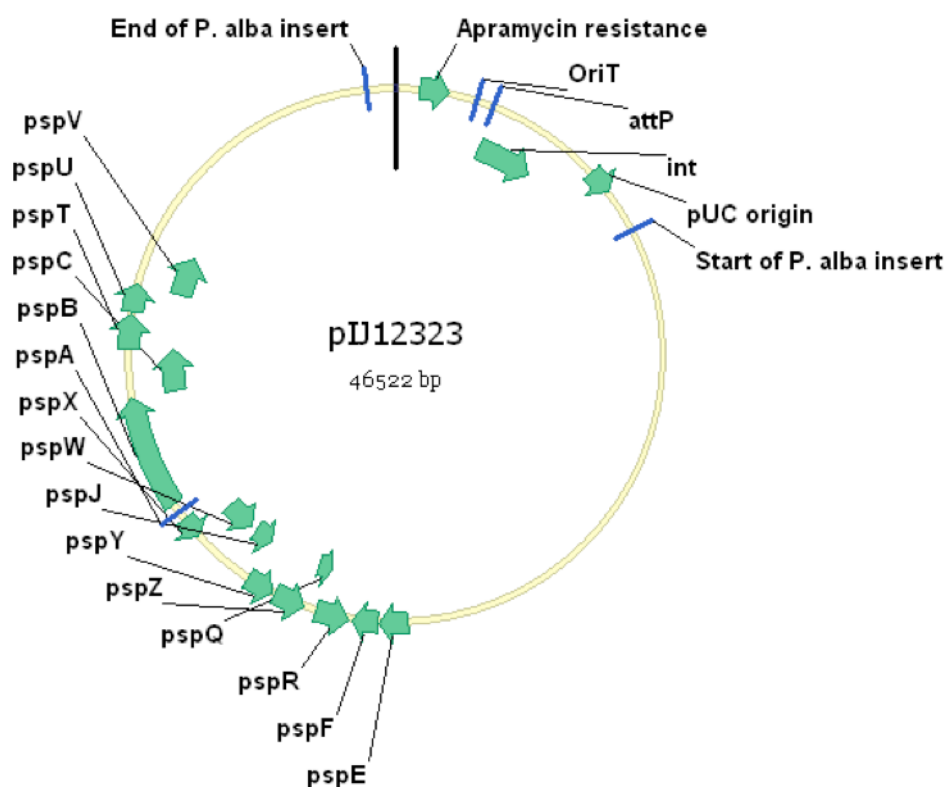


Figure 6.1 : Construct pIJ12323; pIJ12321 (cosmid B4-1) with the backbone targeted with the *oriT-attP-int-aac(3)IV* cassette from pIJ10702.

S. lividans has been used successfully as a heterologous host for expression of lantibiotic biosynthetic gene clusters from *S. cinnamoneus cinnamoneus* and *A. garbadiensis* (Widdick *et al.* 2003; Boakes *et al.* 2009). *S. coelicolor* M1146 is a derivative of *S. coelicolor* M145 with four endogenous secondary metabolic gene clusters (those responsible for actinorhodin, prodiginine, CPK and CDA biosynthesis) deleted (Gomez-Escribano and Bibb 2011). With these potentially competitive sinks for carbon and nitrogen removed, and with no endogenous antibiotic activity remaining, *S. coelicolor* M1146 is a more effective host for the heterologous expression of secondary metabolic gene clusters. This strain has been used to heterologously produce cypemycin from *S. sp.* OH-4156 and the novel lantipeptide grisemycin from *S. griseus* (Claesen and Bibb 2010; Claesen and Bibb 2011). Further derivatives based on *S. coelicolor* M1146 were subsequently constructed through the introduction of point mutations into *rpoB* and *rpsL* to pleiotropically increase the level of secondary metabolite production (Gomez-Escribano and Bibb 2011). *S. coelicolor* M1154 contains the additional mutations *rpoB*[C1298T] and *rpsL*[A262G]; however, this strain was not available at the time of this work.

Exconjugants were selected using apramycin resistance conferred by the *aac(3)IV* gene introduced into the backbone of SuperCosI by PCR-targeting (see above). Four exconjugants from each mating were streaked three times on selective medium for clone

purification. Genomic DNA was prepared from one exconjugant from each mating. *S. lividans* TK24 clones containing the empty vector pIJ10702, pIJ12323 and pIJ12324 were assigned strain numbers M1446, M1286 and M1287, respectively (Table 6.1). Likewise, *S. coelicolor* M1146 containing pIJ10702, pIJ12323 and pIJ12324 were assigned strain numbers M1410, M1288 and M1289, respectively (Table 6.1). PCR analysis using primers annealing within the *psp* gene cluster confirmed the successful introduction of pIJ12323 and pIJ12324 into *S. lividans* TK24 (creating M1286 and M1287) and *S. coelicolor* M1146 (creating M1288 and M1289) (Figure 6.2).

Plasmid	Description	<i>S. lividans</i> TK24	<i>S. coelicolor</i> M1146	<i>Nonomuraea</i> ATCC 39727
pIJ10702	pMJCosI	M1446	M1410	M1294
pIJ12323	pIJ10702 B4-1	M1286	M1288	M1295
pIJ12324	pIJ10702 F13-1	M1287	M1289	M1296
pIJ12325	pIJ10702 4B8	M1290	M1292	M1297
pIJ12326	pIJ10702 9A7	M1291	M1293	M1298
pIJ12327	pIJ10702 B4-1 Δ orf+7-orf+14	NA	NA	M1299
pIJ12328	pIJ10702 B4-1 Δ orf+7-orf+14 and Δ orf-12-orf-19	NA	NA	M1300
pIJ12329	pIJ10702 B4-1 Δ orf+7-orf+14 and Δ orf-10-orf-19	NA	NA	M1301

Table 6.1 : Strains made in this Chapter through integration of various plasmids into the Φ C31 phage attachment site.

6.2.2 Bioassays and MALDI-ToF analysis

pIJ10706, pIJ12323 and pIJ12324 integrated into *S. lividans* TK24 and *S. coelicolor* M1146 were assessed for the production of antibiotic activity. Four exconjugants from each conjugation were grown as 1 x 1 cm patches on a range of agar media. After 5 days incubation, each plate was overlaid with the target organism, *M. luteus*, as described in Chapter 2. However no halo of inhibition was observed (data not shown). The agar media used included those that gave apparently high levels of planosporicin production by *P. alba*; ISP4, AF/MS, D/Seed and M8.

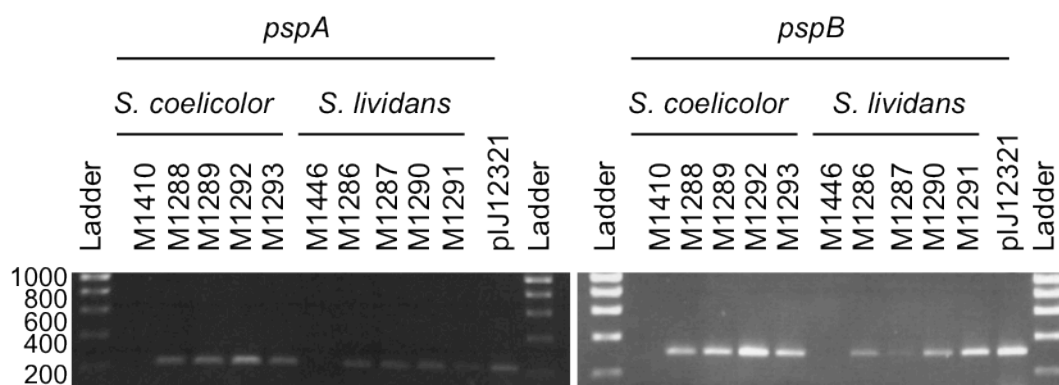


Figure 6.2 : PCR confirmation of the integration of *Planomonospora* cosmids pJ12323, pJ12324, pJ12325 and pJ12326 in *S. coelicolor* M1146 and *S. lividans* TK24.

Genomic DNAs extracted from *S. coelicolor* M1410, M1288, M1289, M1292 and M1293 and from *S. lividans* M1446, M1286, M1287, M1290 and M1291 were used as templates to amplify *pspA* (1289FlanA and 1289RlanA; 230 bp) and *pspB* (3088F and 3088R; 303 bp) from the *psp* gene cluster. pJ12321 cosmid DNA was used as a positive control. PCR products were run on a 1 % agarose gel by electrophoresis. The ladder is Hyperladder I (Bioline) with band sizes annotated in bp.

Two exconjugants from each conjugation were cultured in liquid media. After 4, 6 and 8 days of growth, culture supernatants were assayed for bioactivity against *M. luteus*. No halo of inhibition was observed when the strains were cultured in ISP4, AF/MS or D/seed liquid media (data not shown). The use of R5 medium prompted the production of actinorhodin by the *S. lividans* strains, affording a small halo of inhibition. However when sent for MALDI-ToF analysis, the largest peaks were in the 1500-1600 Da range and were also observed in the vector only control. No peaks with masses corresponding to planosporicin were produced by *Streptomyces* strains harbouring the *psp* gene cluster (data not shown).

The lack of compounds of the expected mass of planosporicin upon MALDI-ToF analysis implied that *Streptomyces* species are not capable of expressing the *psp* gene cluster. There are many potential explanations for this lack of expression, providing a multitude of avenues to pursue, however time restrictions prevented further investigation (discussed in Section 6.6). Instead the remainder of this Chapter focused on finding a host where the cluster was expressed.

6.3 Expression of the planosporicin gene cluster in *Nonomuraea*

6.3.1 Mobilisation of cosmids

Subsequently, an alternative heterologous host was sought. The *Streptomyces* genus is only distantly related to *Planomonospora*, whereas *Nonomuraea* is located with *Planomonospora* in the family *Streptosporangiaceae*. Although not as genetically tractable as *Streptomyces*, several tools are now available for the genetic manipulation of *Nonomuraea* species. *Nonomuraea* ATCC 39727 produces A40926, the precursor of the glycopeptide dalbavancin. The *dbv* gene cluster, responsible for A40926 biosynthesis, was isolated and characterised from *Nonomuraea* ATCC 39727 (Sosio *et al.* 2003). The first instance of genetic manipulation of *Nonomuraea* sp. ATCC 39727 was the generation of a deletion mutation in the *dbv* gene cluster (Sosio *et al.* 2003; Stinchi *et al.* 2003). More recently, methods for introducing larger constructs into *Nonomuraea* through conjugation with *E. coli* were reported. This enabled the specific deletion of genes within the A40926 gene cluster (Marcone *et al.* 2010a). Most recently, the *mib* gene cluster was integrated into the *Nonomuraea* chromosome, resulting in the heterologous production of microbisporicin (Foulston and Bibb 2010). Thus *Nonomuraea* appeared to be a good candidate for the heterologous expression of the *psp* gene cluster.

pIJ10702, pIJ12323 and pIJ12324 (described in Section 6.2.1) were integrated into the chromosome of *Nonomuraea* sp. ATCC 39727 through selection for resistance to apramycin (although there is no experimental evidence that this integration was site-specific). The protocol was based on that used for conjugation into *Streptomyces*. However, as *Nonomuraea* does not sporulate under lab conditions, cosmids were transferred directly into freshly harvested mycelium as described previously (Foulston and Bibb 2010; Marcone *et al.* 2010c). Putative exconjugants were streaked at least three times on V0.1 agar medium with nalidixic acid and apramycin selection. Although common practise when conjugating into spores, this is especially important when conjugating into mycelial fragments. It is possible that within a mycelial compartment that contains multiple genomes, the presence of the integrated resistance gene in just one chromosome may provide sufficient resistance to enable the survival of non-targeted genomes, resulting in a mixed population. Thus successive rounds of growth on selective agar medium are necessary to remove non-targeted nucleoids, leaving a genetically homogenous strain.

The efficiency of conjugating cosmids into *Nonomuraea* was very low. In total, two exconjugants with integrated pIJ12323 and just one exconjugant with integrated pIJ12324 were identified. These were grown under selection in liquid culture from which genomic DNA was purified to be used as a template in PCR analysis. Fragments corresponding to *pspA* and *pspE* were amplified from all three exconjugants to confirm the presence of the

planosporicin biosynthetic gene cluster (Figure 6.3). In addition, amplification of *dbv5* from the A40926 gene cluster confirmed the host as *Nonomuraea* (Figure 6.3). *Nonomuraea* with integrated pIJ10702 was assigned strain M1294, while *Nonomuraea* with integrated pIJ12323 was assigned strain M1295, and *Nonomuraea* with integrated pIJ12324 was assigned strain M1296 (Table 6.1).

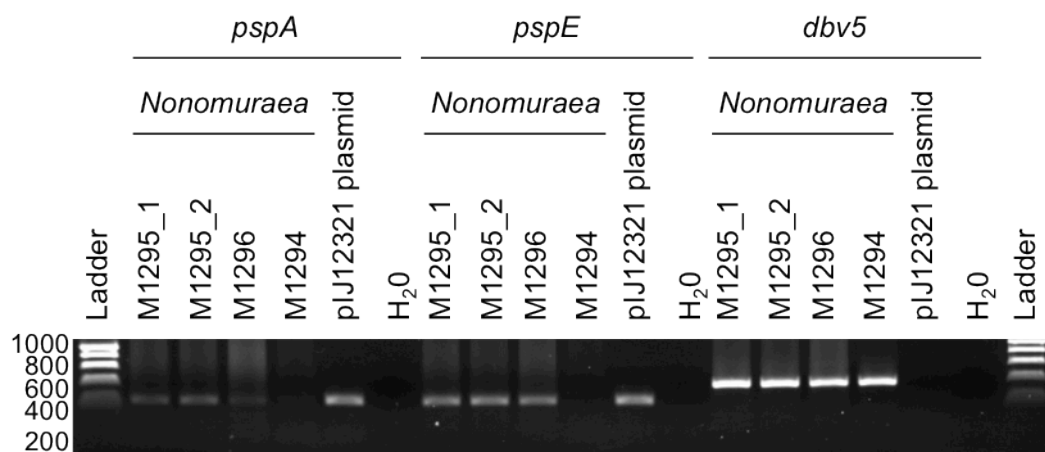


Figure 6.3 : PCR confirmation of the integration of *Planomonospora* cosmids pIJ12323 and pIJ12324 in *Nonomuraea* ATCC 39727.

Genomic DNAs extracted from *Nonomuraea* M1295 and M1296 were used as templates to amplify *pspA* (1289FlanA and 1289RlanA; 230 bp) and *pspE* (FlanEF and RlanEF; 221 bp) from the *psp* gene cluster and *dbv5* from the A40926 gene cluster (Dbv5F and Dbv5R; 375 bp). Genomic DNA from *Nonomuraea* M1294, which contains the empty pIJ10702 vector, was used as a negative control. pIJ12321 cosmid DNA was used as a positive control. PCR products were run on a 1 % agarose gel by electrophoresis. The ladder is Hyperladder I (Bioline) with band sizes annotated in bp.

6.3.2 Bioassay and MALDI-ToF analysis

6.3.2.1 *Nonomuraea* M1295

Nonomuraea M1294 (harbouring the vector only control pIJ10702) and *Nonomuraea* M1295 (harbouring pIJ12323) were assessed for the production of a compound with antibiotic activity. *Nonomuraea* is not an ideal host for the production of heterologous antibacterial compounds since it produces the glycopeptide antibiotic A40926. Thus bioassays of *Nonomuraea* exconjugants grown on agar medium would likely be futile since even the vector only control would produce a halo of inhibition, although it is conceivable that a larger zone of inhibition could be indicative of expression of a heterologous gene cluster.

Consequently the strains were cultured in liquid medium, and culture supernatants assayed for bioactivity against *M. luteus* and sent for MALDI-ToF analysis to confirm planosporicin production. It was anticipated that VSP medium, which proved effective for growth of *Nonomuraea*, would also support planosporicin production. Moreover, expression of the *mib* gene cluster of *Microbispora corallina* in the same strain of *Nonomuraea* had been achieved in this medium (Foulston and Bibb 2010). However, initial results were disappointing. While the majority of culture supernatants produced halos of inhibition of *M. luteus* potentially attributable to A40926, MALDI-ToF analysis failed to reveal peaks with masses corresponding to planosporicin (data not shown).

Subsequently, additional media were assessed for their ability to support planosporicin production in *Nonomuraea*. It was assumed that as a member of the *Streptosporangiaceae*, *Nonomuraea* would grow best in media optimised for genera within this family. Six such media were tested; ISP4, AF/MS, D/Seed, M8, medium 266 (YS) and *Streptosporangium* medium (SM). A compound with anti-microbial activity (presumably A40926) was observed in supernatants of cultures of *Nonomuraea* M1294 (vector only control) grown in AF/MS, D/seed, YS and SM media, producing a small, often diffuse halo of growth inhibition of *M. luteus* (Figure 6.4). A significantly larger halo with a more defined boundary was observed with culture supernatants from *Nonomuraea* M1295 (containing the *psp* gene cluster) grown in ISP4 and SM media (Figure 6.4). The lack of antimicrobial production by *Nonomuraea* M1294 in ISP4 media gave a clean background for the assessment of planosporicin production (Figure 6.4). However, repeat cultures in ISP4 medium often failed to reveal any antibiotic production. Due to these reproducibility issues, the best medium for studying the production of planosporicin in *Nonomuraea* was deemed to be SM.

Both exconjugants of *Nonomuraea* with integrated pIJ12323 secreted peptides into the culture supernatant with masses corresponding to planosporicin. Figure 6.5 demonstrates the sharp zone of inhibition observed in *Nonomuraea* M1295 compared to *Nonomuraea* M1294. MALDI-ToF analysis confirmed that this sharp zone was due to planosporicin production (Figure 6.6). This proved beyond reasonable doubt that the cloned DNA insert of pIJ12323 (derived from pIJ12321) contained the entire planosporicin biosynthetic gene cluster.

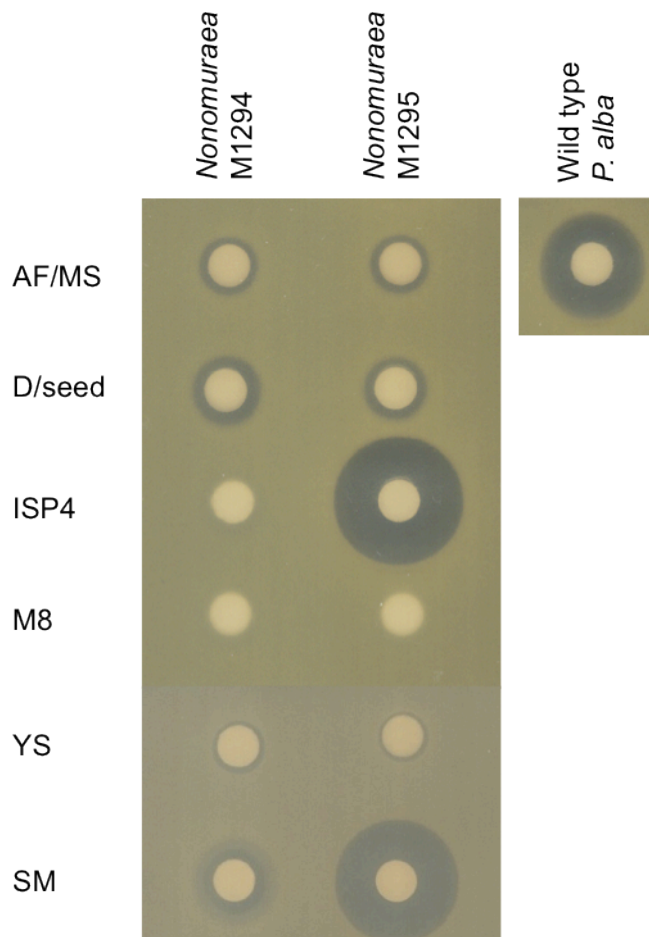


Figure 6.4 : Heterologous expression of the planosporicin gene cluster integrated into the *Nonomuraea* ATCC 39727 chromosome.

Nonomuraea M1294 and M1295 were grown in 6 different media; AF/MS, D/seed, ISP4, M8, medium 266 (YS) and streptosporangium medium (SM). A positive control of *P. alba* in AF/MS was grown in parallel. Supernatant samples were taken after 6 days of growth and tested for antibiotic activity. 40 μ l supernatant was applied to antibiotic assay discs which were laid onto a lawn of *M. luteus*. The plate was incubated for 36 hours at 30 °C before zones of inhibition were photographed.

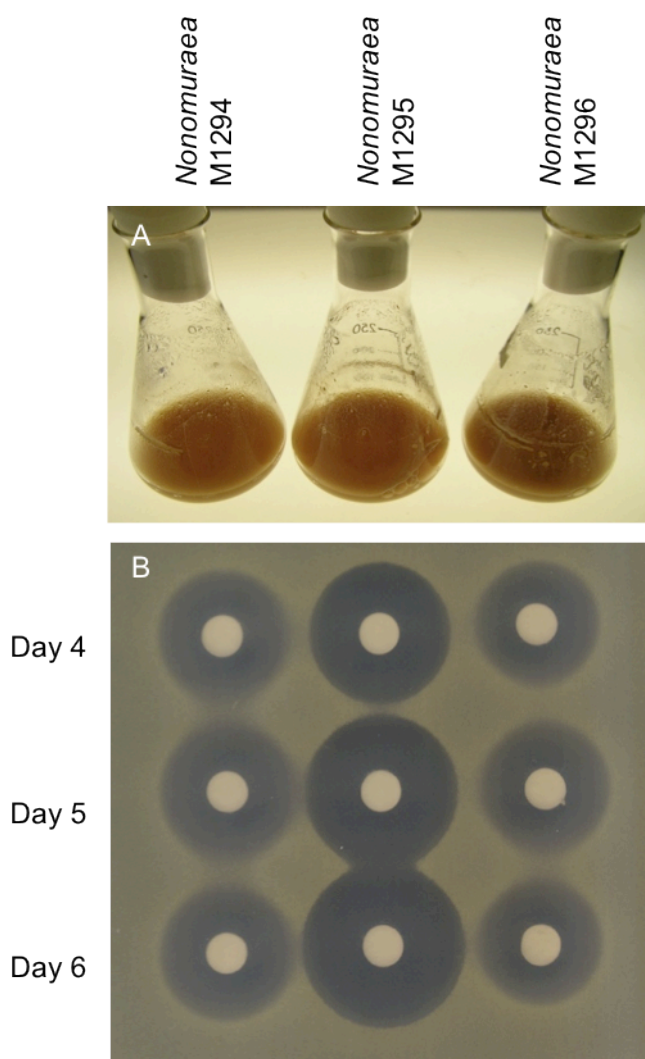


Figure 6.5 : Heterologous expression of the planosporicin gene cluster integrated into the *Nonomuraea* ATCC 39727 chromosome.

A; *Nonomuraea* M1294, M1295 and M1296 were cultured in SM medium. **B**; Supernatant samples were taken at 4, 5 and 6 days of growth. 40 μ l of each culture supernatant were applied to antibiotic assay discs which were laid onto a lawn of *M. luteus*. The plate was incubated for 36 hours at 30 °C before zones of inhibition were photographed.

6.3.2.2 *Nonomuraea* M1296

Section 6.3.2.1 demonstrated that the heterologous production of planosporicin by *Nonomuraea* M1295 was most consistent when cultured in SM. This media was originally developed to culture novel *Nonomuraea* species and subsequently used to prompt secondary metabolite production from the *Streptosporangium* genus (Pfefferle *et al.* 2000; Roes and Meyers 2008).

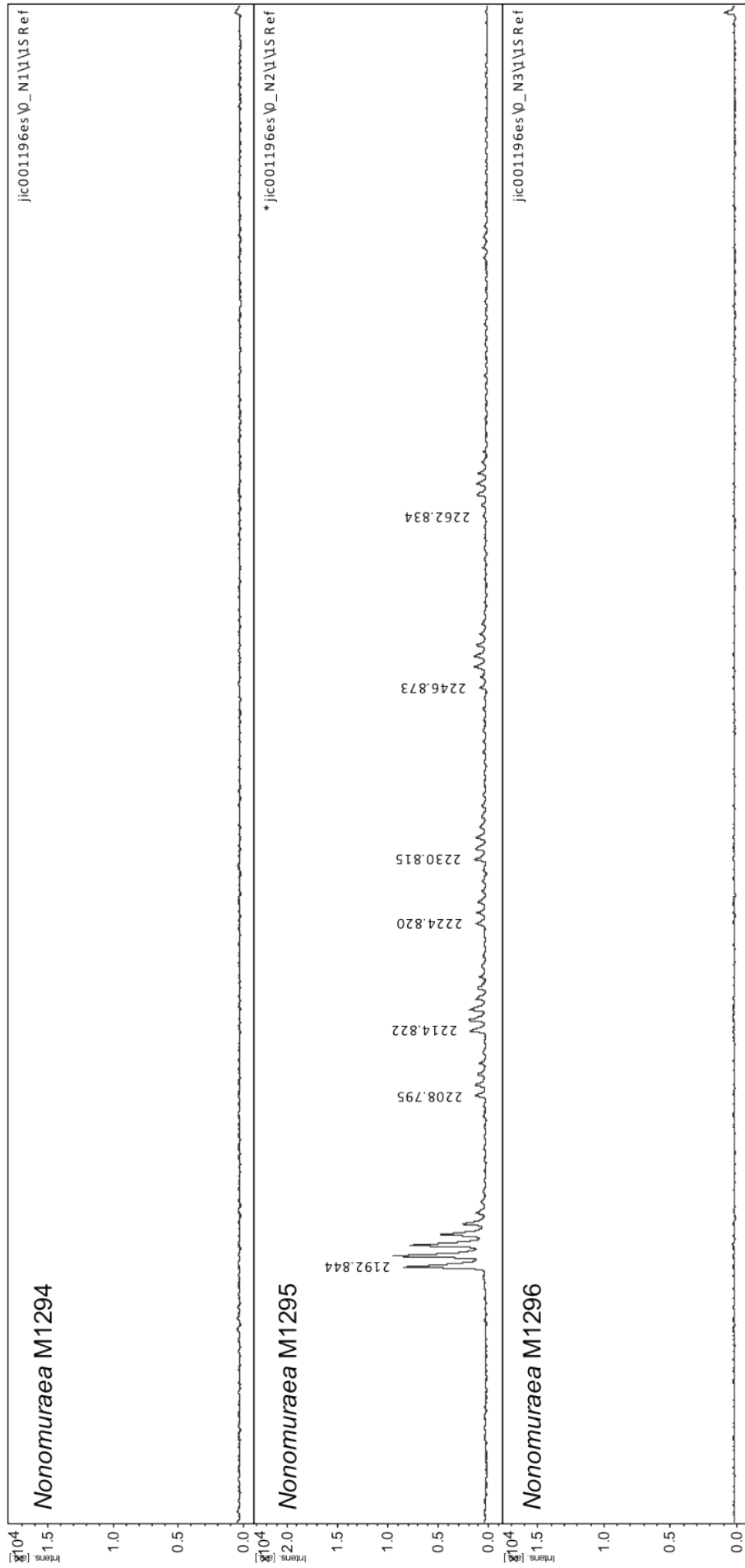


Figure 6.6 : Heterologous production of planosporicin in *Nonomuraea* ATCC 39727.

Nonomuraea M1294, M1295 and M1296 were grown in SM medium. Supernatant samples were taken after 8 days of growth and analysed by MALDI-ToF mass spectrometry. Intensity is given on the y-axis in arbitrary units (au) and the mass/charge ratio (m/z) on the x-axis. The monoisotopic mass of each m/z peak is labelled.

Nonomuraea M1296 (lacking part of *pspV*), *Nonomuraea* M1294 (vector only control), and the positive control strain *Nonomuraea* M1295 were cultured in parallel in SM media. After 4, 5 and 6 days, samples of culture supernatant were assessed for the production of antibiotic activity. *Nonomuraea* M1296 supernatants gave a very diffuse halo of *M. luteus* inhibition (Figure 6.5). This bioassay phenotype closely resembles the vector only control. MALDI-ToF analysis failed to reveal peptides corresponding to the mass of planosporicin (Figure 6.6). A40926 has a molecular mass $[M+H]^+$ of 1731 Da (Sosio *et al.* 2003). However a peak with this mass was not detected, indeed no significant peptide masses in the range 700-4000 Da were routinely observed in M1294 supernatants. Thus the diffuse halo observed in bioassays with M1294 was likely due to a non-peptide antimicrobial produced by *Nonomuraea* ATCC 39727 in SM media. The lack of a full-length *pspV* in M1296 appears to prevent planosporicin production. MALDI-ToF analysis of supernatant samples in a broad spectrum scan did not reveal the presence of early intermediates such as the 5484.616 Da theoretical mass of planosporicin with the leader sequence still attached. This implies that *pspV* has an essential role in biosynthesis, at least in a heterologous host. However, despite many conjugation attempts, only one exconjugant of *Nonomuraea* M1296 was obtained, so the lack of planosporicin production may be due to a problem with this particular isolate.

6.4 Construction of a minimal gene set for planosporicin production

6.4.1 Creation of constructs

The likely boundaries of the planosporicin biosynthetic gene cluster within pIJ12321 could be inferred from the putative functions of the constituent genes (as discussed in Chapter 5). Furthermore, the marked synteny of the planosporicin and microbisporicin biosynthetic gene cluster provided further clues. To confirm the boundaries of the planosporicin biosynthetic gene cluster, potentially nonessential DNA regions were removed from the cosmid inserts in an attempt to define the minimal set of genes required for planosporicin production.

The boundary downstream of the cluster was relatively easy to predict based on the putative function of proteins assigned from their similarity to other proteins in the NCBI database. It is likely that *pspV* denotes the downstream boundary. Further downstream, ORFs are predicted to have a putative role in benzoate metabolism, and thus not to be part of the *psp* cluster. In contrast, the likely upstream boundary was more difficult to predict. *pspEF* are almost certainly part of the cluster due to the high percentage identity of their products to the immunity ABC transporters of other lantibiotic clusters.

Immediately upstream of *pspEF* there are two possibly co-transcribed genes, one of which may encode a protease function (as described in Chapter 5). Lantibiotic gene clusters often encode a protease to cleave the leader peptide, releasing the mature lantibiotic, but so far, no lantibiotic gene cluster from an actinomycete has been found to encode such a function. Further upstream there is a putative terpenoid biosynthetic cluster. Therefore two potential minimal gene sets were made, both with *pspV* delineating the right-hand border and with either *pspE* or the putative protease gene delineating the left-hand border.

Sometimes a restriction fragment can be isolated that contains a minimal gene set. For example a 12.9 kb *SacI* fragment from *Streptomyces chartreusis* conferred tunicamycin production on *S. coelicolor* M1146, delineating the boundaries of the *tun* gene cluster (Wyszynski *et al.* 2010). Due to the lack of convenient restriction sites to excise the putative minimal set of *psp* genes, an alternative strategy was sought. A similar problem was faced when generating the minimal gene set for cypemycin (Claesen and Bibb 2010). In this work the flanking regions were targeted as above but the resistance cassette was removed through the action of FLP recombinase, leaving unique restriction sites *XbaI* and *SspI* either side of the cypemycin minimal gene set. These sites were utilised to excise a 12.2 kb *XbaI-SspI* fragment that was then ligated into pSET152 and the resulting construct conjugated into *S. coelicolor* M1146. However this method can be problematic, as the removal of the resistance cassette by FLP recombinase leaves an 81 bp 'scar' sequence that is a functional FRT site. This represents a potential target for homologous recombination during subsequent targeting with constructs containing FRT sites.

A variation on this method was first used in this work as described in Chapter 2 and outlined in Figure 6.7. Primers with either an *XbaI* or *SpeI* recognition site between the resistance marker and the 39 bp flanking sequence used for recombination in *E. coli* were used to amplify a resistance cassette. PCR targeting replaced the DNA regions outside of the proposed minimal gene set with a resistance cassette flanked by unique restriction sites. Subsequently, the resistance cassette was digested out and the minimal gene set ligated back together. pIJ12321 was first PCR-targeted to remove genes downstream of the cluster, to the right of *pspV*. This cosmid was targeted a second time using two alternative constructs to remove genes upstream of the cluster. All three of these reduced gene sets were PCR-targeted to replace the kanamycin resistance gene (*aph*) in SuperCosI with the *oriT-attP-int-aac(3)IV* cassette (as detailed in Section 6.2.1). Removal of all genes downstream of *pspV* and subsequent introduction of the *oriT-attP-int-aac(3)IV* cassette into the vector backbone created pIJ12327 (Figure 6.8). Removal of all genes downstream of *pspV* and upstream of the putative protease created pIJ12328 (Figure 6.8). Removal of two further genes upstream of *pspE* (including the putative *lanP*) created pIJ12329 (Figure 6.8).

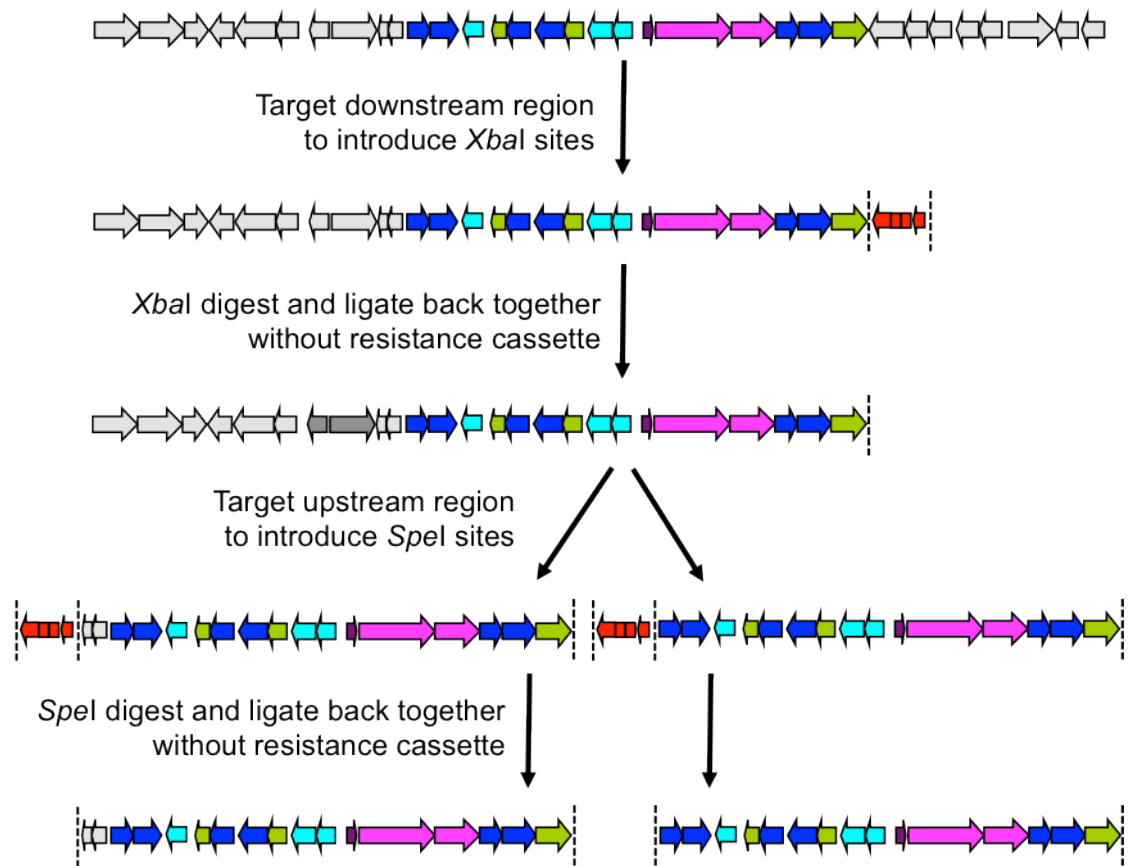


Figure 6.7 : Strategy for the construction of a minimal *psp* gene set.

Schematic displays cosmid insert only (backbone not shown) with orfs as annotated in Chapter 4. The DNA insert of pIJ12321 was PCR-targeted to replace all ORFs downstream of *pspV* with a resistance cassette (in red) and unique *Xba*I restriction sites (dashed lines). The resistance gene was removed by *Xba*I digestion and the resulting construct subjected to a second targeting event upstream of *pspE* or *orf-10* with a resistance cassette (in red) and unique *Spe*I restriction sites (dashed lines). The resistance gene was subsequently removed through *Spe*I digestion. In total, three reduced *psp* cosmids were created.

These three constructs were conjugated into *Nonomuraea* using the method described in Section 6.3.1. *Nonomuraea* exconjugants were obtained containing pIJ12327, pIJ12328 or pIJ12329, with each construct integrated into the Φ C31 phage attachment site. Genomic DNA was purified from four exconjugants with integrated pIJ12327, five exconjugants with pIJ12328 and three exconjugants with pIJ12329. PCR analysis using primers annealing within *pspA* and *pspE* confirmed the presence of the *psp* gene cluster (Figure 6.9).

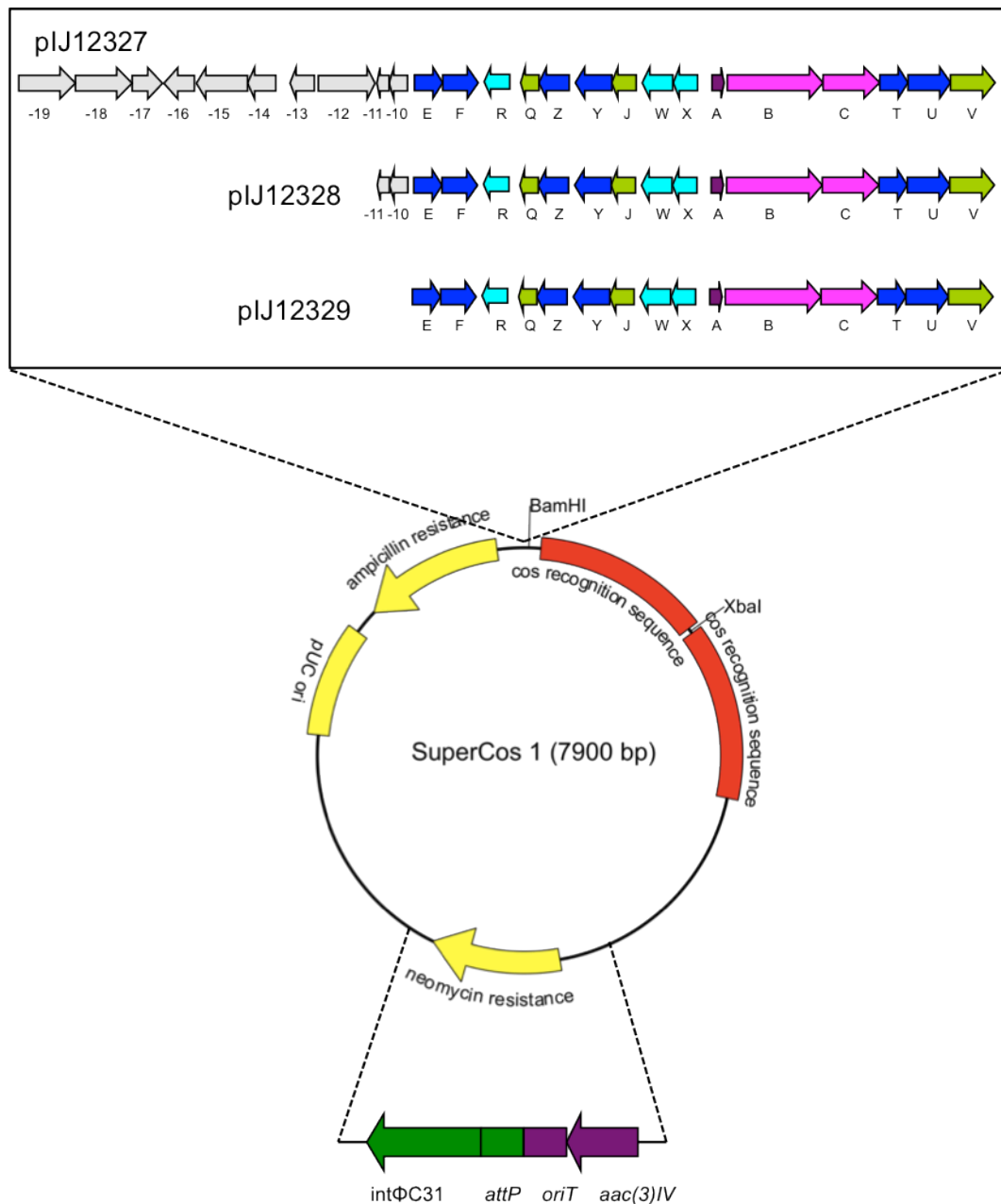


Figure 6.8 : Schematic illustrating the reduced *psp* gene clusters.

The insert DNA of cosmid pIJ12321 was PCR-targeted to create three reduced versions. The SuperCos1 backbone of each of these constructs was PCR-targeted to introduce an integrase (*intΦC31*), phage attachment site (*attP*) and origin of transfer (*oriT*), creating pIJ12327, pIJ12328 and pIJ12329. The diagram is not drawn to scale.

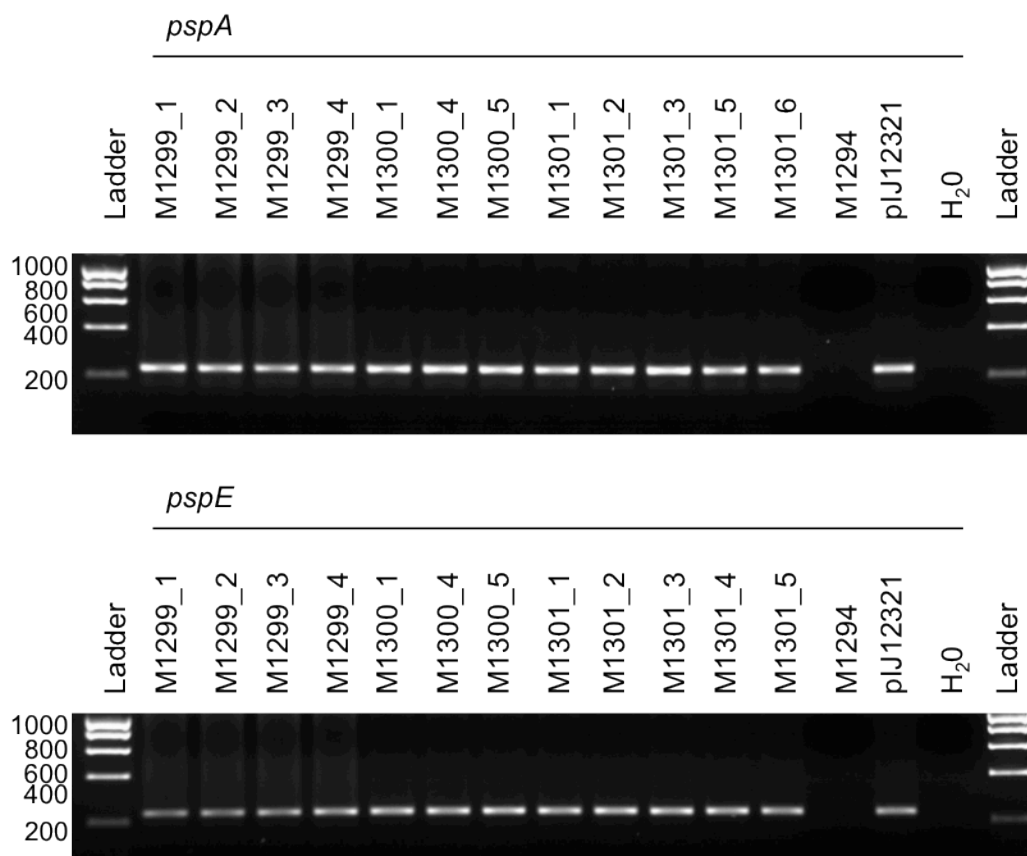


Figure 6.9 : PCR confirmation of the integration of *Planomonospora* cosmids pIJ12327, pIJ12328 and pIJ12329 in *Nonomuraea* ATCC 39727.

Genomic DNAs extracted from *Nonomuraea* M1299 (integrated pIJ12327), M1300 (integrated pIJ12328) and M1301 (integrated pIJ12329) were used as templates to amplify *pspA* (1289FlanA and 1289RlanA; 230 bp) and *pspE* (FlanEF and RlanEF; 221 bp) from the *psp* gene cluster. Genomic DNA from *Nonomuraea* M1294, which contains the empty pIJ10702 vector, was used as a negative control. pIJ12321 cosmid DNA was used as a positive control. PCR products were run on a 2 % agarose gel by electrophoresis. The ladder is Hyperladder I (Bioline) with band sizes annotated in bp.

6.4.2 Bioassay and MALDI-ToF analysis

To define the right-hand boundary, four exconjugants of *Nonomuraea* with integrated pIJ12327 were tested for the heterologous expression of planosporicin. All four exconjugants produced a compound with antibiotic activity that MALDI-ToF revealed to be due to planosporicin (data not shown). To define the left-hand boundary, three exconjugants of *Nonomuraea* with integrated pIJ12328 and five exconjugants of *Nonomuraea* with integrated pIJ12329 were grown in SM and samples of supernatant were assayed for antibiotic production (Figure 6.10). It is interesting to note that the halo size varies between different exconjugants carrying the same integrated construct. This

might reflect the presence of both WT and integrant genomes in the mycelium of the *Nonomuraea* exconjugants. *Nonomuraea* pIJ12328_5 and *Nonomuraea* pIJ12329 exconjugants 1, 3 and 5 gave larger halos of inhibition than the *Nonomuraea* pIJ10702 empty vector control (M1294). Of the three *Nonomuraea* pIJ12328 exconjugants, only pIJ12328_5 gave peaks corresponding to planosporicin upon MALDI-ToF analysis, whereas all five *Nonomuraea* pIJ12329 exconjugants produced the lantibiotic (data not shown). *Nonomuraea* pIJ12327_1 was assigned strain M1299, *Nonomuraea* pIJ12328_5 was assigned strain M1300 and *Nonomuraea* pIJ12329_3 was assigned strain M1301 (Table 6.1). These three strains were used in subsequent assays.

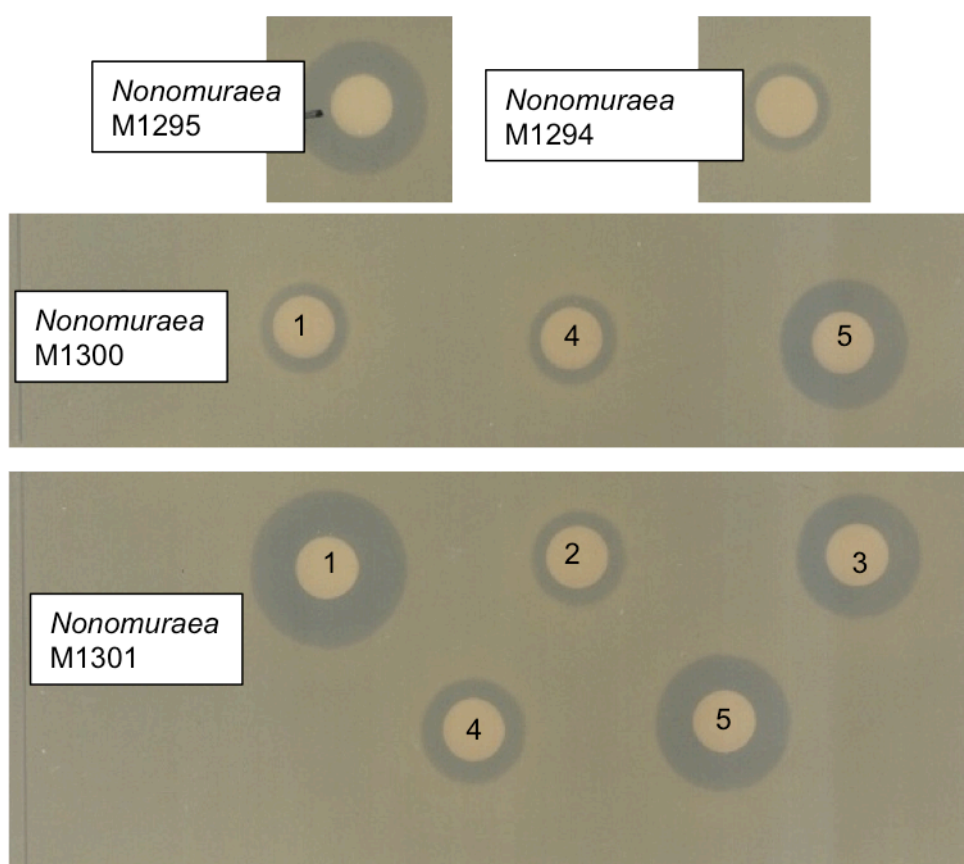


Figure 6.10 : Heterologous production of planosporicin from reduced gene sets integrated into the *Nonomuraea* ATCC 39727 chromosome.

Nonomuraea M1300 clones 1, 4 and 5 and M1301 clones 1 to 5 were cultured in SM medium for 6 days. A positive control of *Nonomuraea* M1295 and a negative control of *Nonomuraea* M1294 were cultured in parallel. Samples of culture supernatants were tested for antibiotic activity. 40 μ l of supernatant were applied to antibiotic assay discs which were laid onto a lawn of *M. luteus*. The plate was incubated for 36 hours at 30 °C before zones of inhibition were photographed.

Thus the heterologous expression of pIJ12327, pIJ12328 and pIJ12329 in *Nonomuraea* resulted in planosporicin production. Bioassays with *M. luteus* revealed antibiotic production, while MALDI-ToF analysis confirmed that the compound was planosporicin (Figure 6.11 and 6.12). Consequently, the gene encoding a putative protease is not required for cleavage of the planosporicin leader peptide in *Nonomuraea*. The halo produced through bioassay with *Nonomuraea* M1299, M1300 and M1301 was comparable in size to that produced by *Nonomuraea* M1295. Thus the 15 genes from *pspE* to *pspV* are sufficient for planosporicin production in *Nonomuraea*, and are predicted to constitute a minimal gene set.

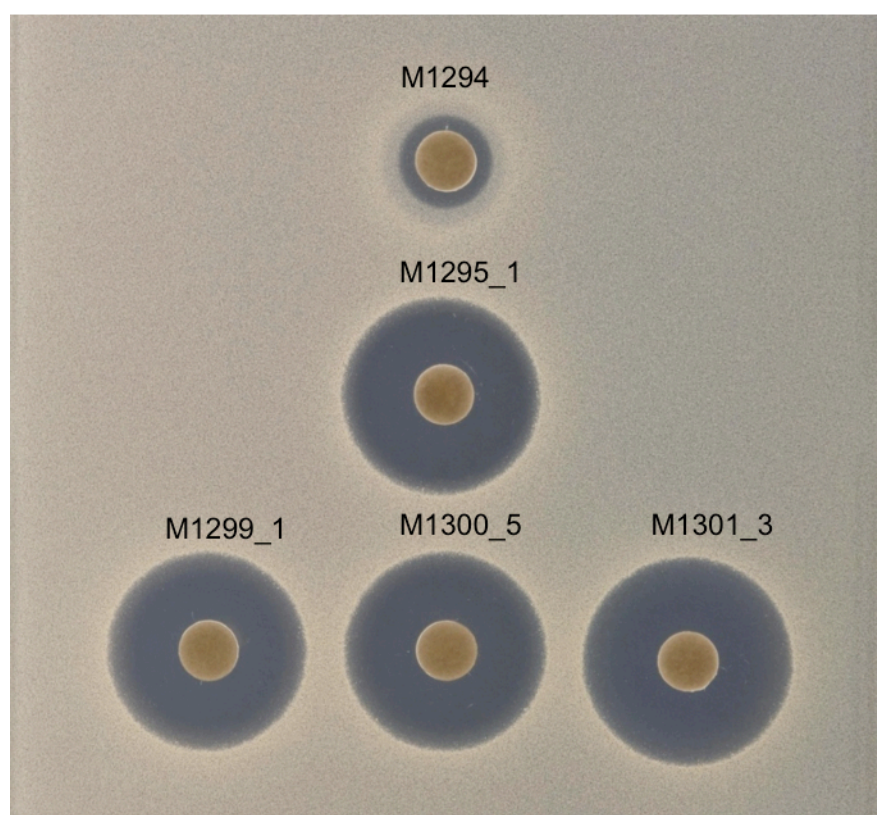


Figure 6.11 : Heterologous expression of the planosporicin gene cluster integrated into the *Nonomuraea* ATCC 39727 chromosome.

Nonomuraea M1294, M1295_1, M1299_1, M1300_5 and M1301_3 were cultured in SM medium. Supernatant samples were taken after 8 days of growth and tested for antibiotic activity. 40 μ l of supernatant were applied to antibiotic assay discs that were laid upon a lawn of *M. luteus*. The plate was incubated for 36 hours at 30°C before zones of inhibition were photographed.

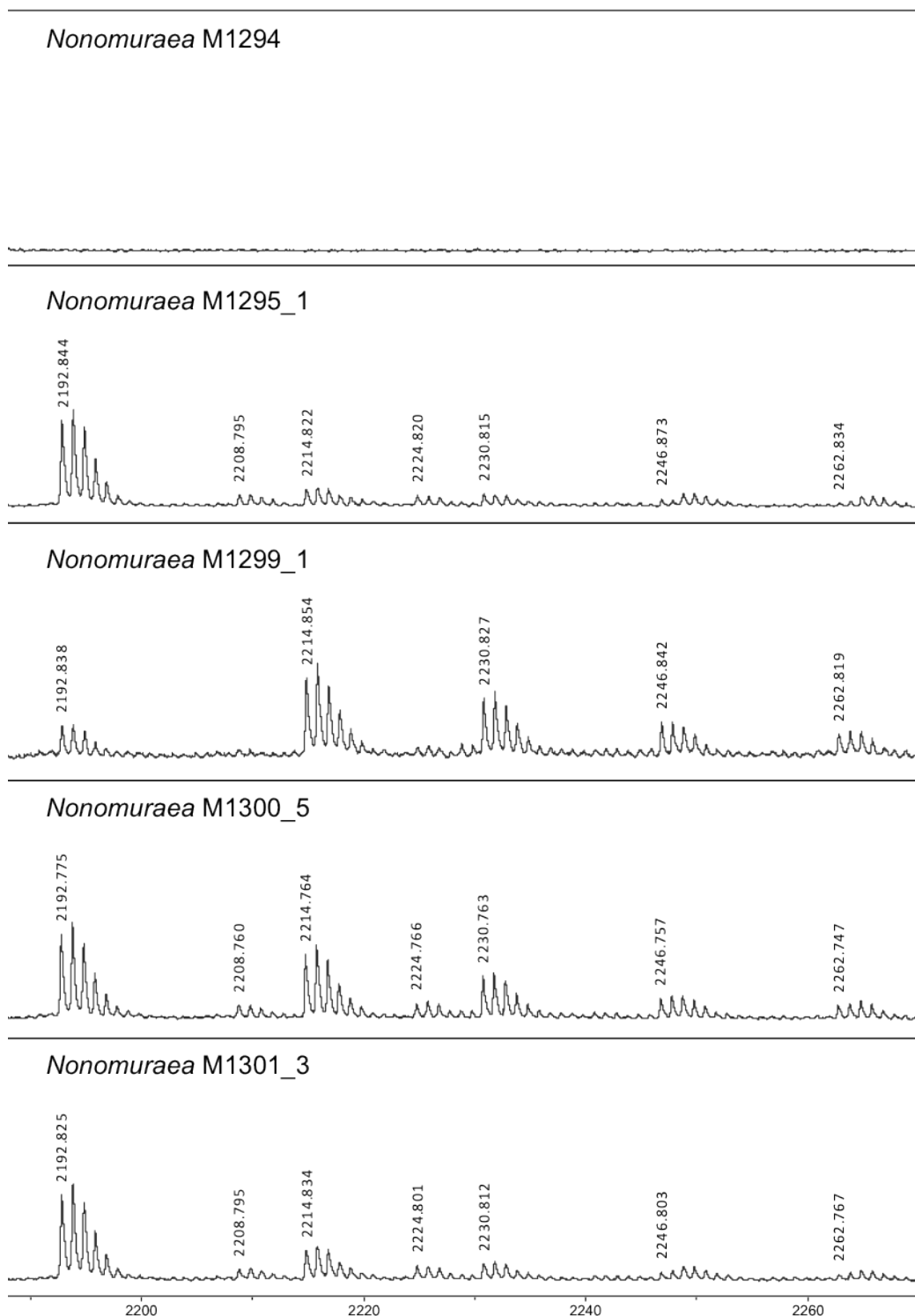


Figure 6.12 : Heterologous production of planosporicin from reduced gene clusters derived from pIJ12321 integrated into the *Nonomuraea* ATCC 39727 chromosome. *Nonomuraea* M1294, M1295_1, M1299_4, M1300_5 and M1301_3 were cultured in SM medium. Supernatant samples were taken after 8 days of growth and analysed by MALDI-ToF mass spectrometry. Intensity is given on the y-axis in arbitrary units (au) and the mass/charge ratio (m/z) on the x-axis. The monoisotopic mass of each m/z peak is labelled.

6.5 Heterologous expression of *P. sp.* DSM 14920 cosmids

6.5.1 Creation of constructs

As described in Section 4.7 of Chapter 4, the NAICONS consortium constructed a cosmid library from *P. sp.* DSM 14920 genomic DNA which had also been reported to produce planosporicin. Four partially overlapping cosmids each containing all or part of the planosporicin biosynthetic gene cluster and flanking regions were received from NAICONS. Chapter 4 confirmed that pIJ12325 and pIJ12326 contained a number of planosporicin biosynthetic genes.

The vector used by NAICONS to create the cosmid library was a SuperCos1 derivative carrying the *oriT-attP-int-aac(3)IV* cassette. Thus the cosmids were introduced into the donor strain *E. coli* ET12567/pUZ8002 by transformation and mobilised into three different hosts; *S. lividans* TK24, *S. coelicolor* M1146 and *Nonomuraea* ATCC 39727 according to the method described in Section 6.2.1 and 6.3.1.

Integration of pIJ12325 and pIJ12326 into *S. lividans* TK24 created strains M1290 and M1291, respectively, while their integration into *S. coelicolor* M1146 created strains M1292 and M1293, respectively (Table 6.1). PCR analyses using gDNAs isolated from these strains as templates revealed that the cosmids had integrated into the host's genome (Figure 6.2). Integration of pIJ12325 and pIJ12326 into the Φ C31 phage attachment site of *Nonomuraea* created strains M1297 and M1298 (Table 6.1). One exconjugant from each mating was obtained and the amplification of *dbv5* from the A40926 cluster confirmed these strains as *Nonomuraea* derivatives. Both *pspA* and *pspE* were successfully amplified from M1297 and M1298, confirming of the presence of the planosporicin biosynthetic gene cluster (Figure 6.13).

6.5.2 Bioassay and MALDI-ToF analysis

The lack of a complete sequence for the insert DNA of the two cosmids meant it was not certain that the entire biosynthetic gene cluster was present on either construct. Planosporicin production was not detected from *S. lividans* M1290 or M1291, nor from *S. coelicolor* M1292 or M1293 when cultured in liquid AF/MS and R5 media. No antibiotic compounds were detected by bioassay and no compounds corresponding to planosporicin were detected in MALDI-ToF (data not shown).

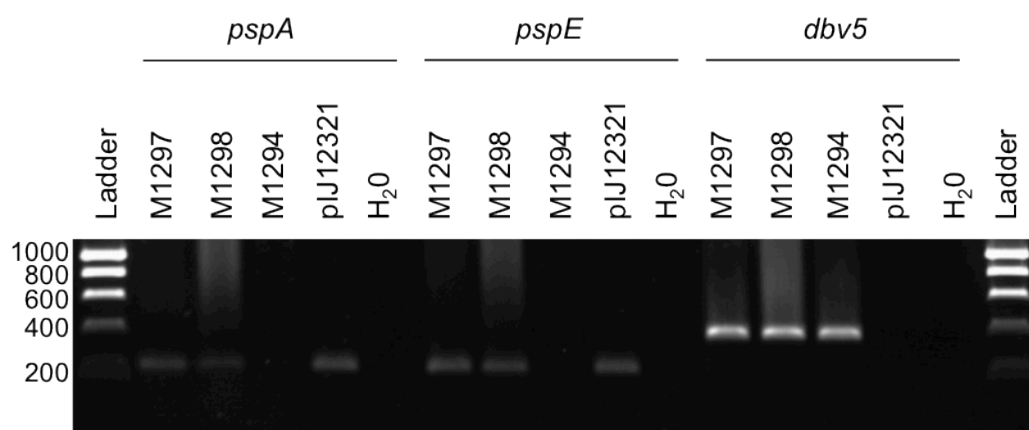


Figure 6.13 : PCR confirmation of the integration of *Planomonospora* cosmids pIJ12325 and pIJ12326 in *Nonomuraea* ATCC 39727.

Genomic DNAs extracted from *Nonomuraea* M1297 and M1298 were used as templates to amplify *pspA* (1289FlanA and 1289RlanA; 230 bp) and *pspE* (FlanEF and RlanEF; 221 bp) from the *psp* gene cluster and *dbv5* from the A40926 gene cluster of *Nonomuraea* (Dbv5F and Dbv5R; 375 bp). Genomic DNA from *Nonomuraea* M1294, which contains the empty pIJ10702 vector, was used as a negative control. pIJ12321 cosmid DNA was used as a positive control. PCR products were run on a 1 % agarose gel by electrophoresis. The ladder is Hyperladder I (Bioline) with band sizes annotated in bp.

Prior success with the heterologous expression of the *psp* gene cluster from *P. alba* in *Nonomuraea* made this host the main candidate for expression of the cluster from *P. sp.* DSM 14920. Both *Nonomuraea* M1297 and M1298 produced planosporicin. Bioassay with *M. luteus* revealed antibiotic activity, and MALDI-ToF analysis identified a compound with a mass of 2192 Da, the mass of planosporicin (Figure 6.14 and 6.15).

Culture supernatants from *Nonomuraea* M1297 and M1298 carrying the planosporicin gene cluster from *P. sp.* DSM 14920 gave the same spectrum of peaks in MALDI-ToF as *Nonomuraea* M1295 carrying the *psp* cluster from *P. alba* (Figure 6.15), confirming that both *Planomonospora* strains do indeed produce planosporicin.

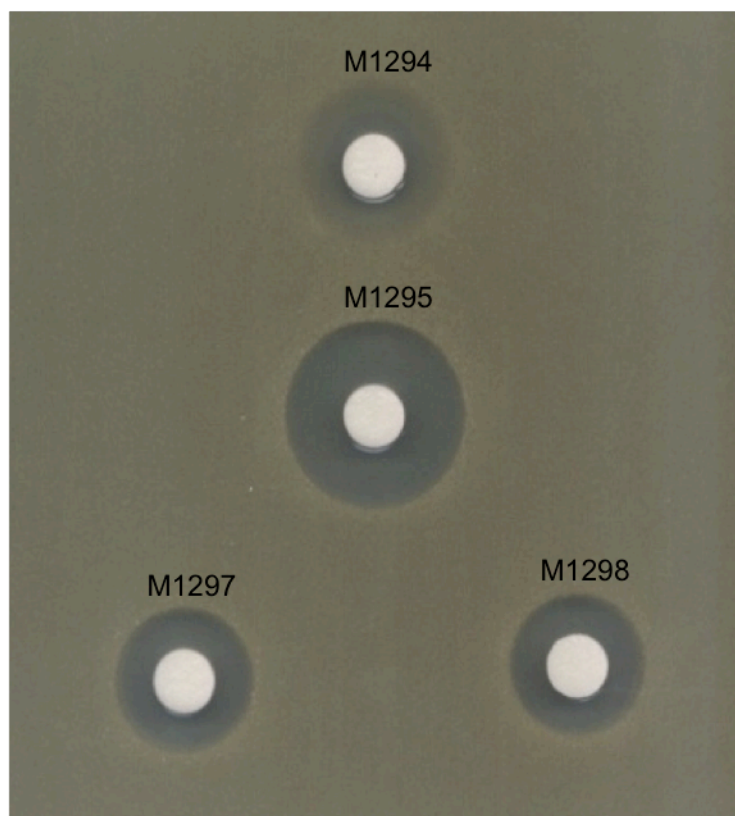


Figure 6.14 : Heterologous expression of the planosporicin gene cluster integrated into the *Nonomuraea* ATCC 39727 chromosome.

Nonomuraea M1294, M1295, M1296, M1297 and M1298 were cultured in SM medium. Supernatant samples were taken after 8 days of growth and tested for antibiotic activity. 40 μ l of supernatant were applied to antibiotic assay discs that were laid upon a lawn of *M. luteus*. The plate was incubated for 36 hours at 30 °C before zones of inhibition were photographed.

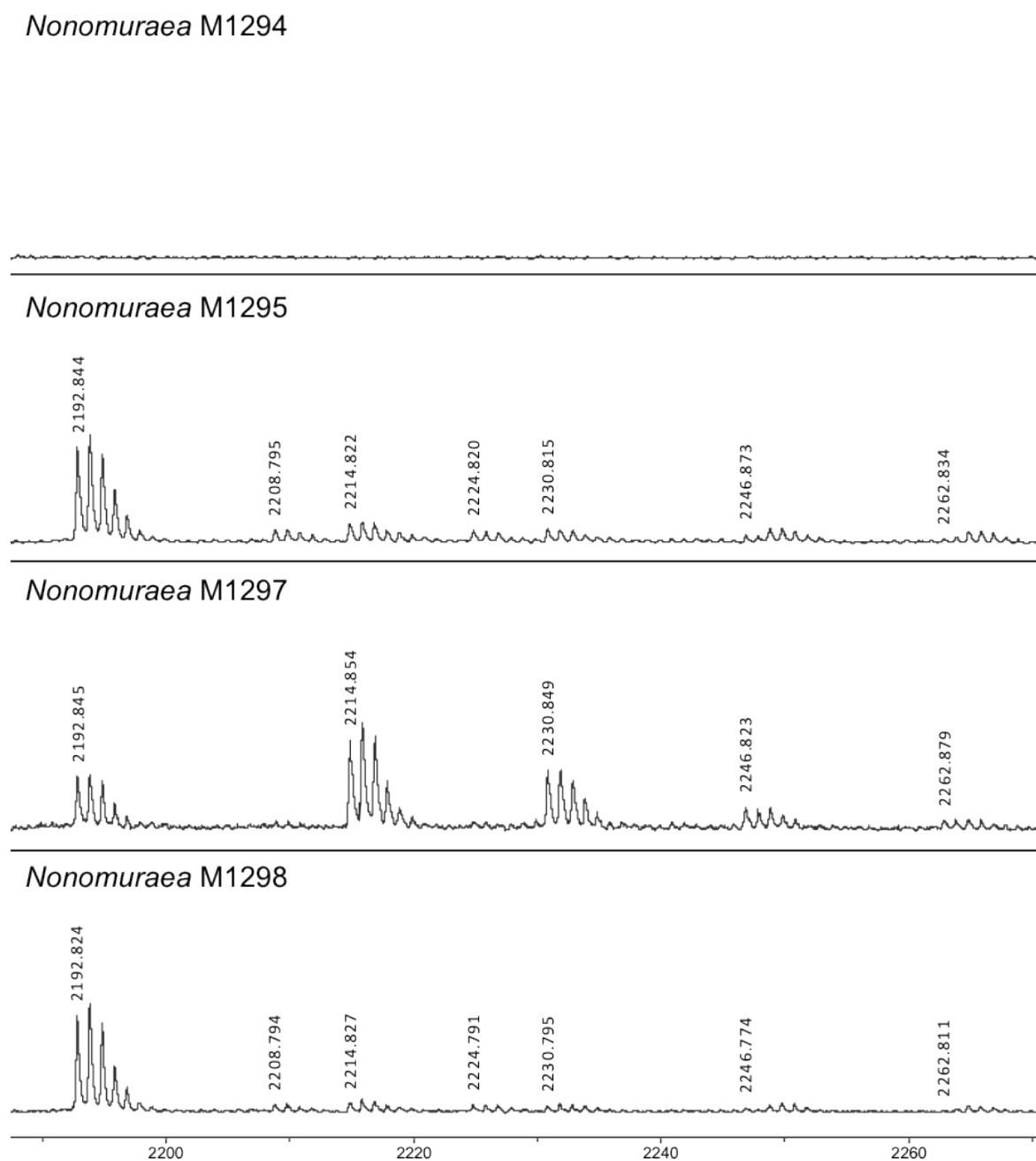


Figure 6.15 : Heterologous production of planosporicin from *P. sp.* DSM 14920 cosmids integrated into the *Nonomuraea* ATCC 39727 chromosome.

Nonomuraea M1294, M1295, M1297 and M1298 were cultured in SM medium. Supernatant samples were taken after 8 days of growth and investigated by MALDI-ToF mass spectrometry. Intensity is given on the y-axis in arbitrary units (au) and the mass/charge ratio (m/z) on the x-axis. The monoisotopic mass of each m/z peak is labelled.

6.6 Discussion

Attempts to express the *psp* gene cluster of *P. alba* in two heterologous *Streptomyces* species were not successful. Similar failures in heterologous production from gene clusters moved between genera have been observed for both low-GC and high-GC lantibiotic clusters. For example, it is not possible to heterologously produce the lantibiotic nisin in *B. subtilis* 168 after transfer of the nisin biosynthetic gene cluster from *L. lactis* ATCC 11454 (Yuksel and Hansen 2007). Likewise, no microbisporicin was detected when the complete *mib* gene cluster was mobilised into several *Streptomyces* species (Foulston and Bibb 2010). A number of strategies were previously utilised in an attempt to prompt microbisporicin production by *S. lividans* TK24. Growth conditions were altered and external inducers such as ATP and GlcNAc were added, as well as attempts at auto-induction through the addition of microbisporicin (Foulston 2010). The constitutive promoter *ermE**p was introduced to drive expression of *mibX*, *mibA* and *mibE* in the event that the heterologous host did not transcribe the cluster at a high level (Foulston 2010). *mibW*, encoding an anti-sigma factor, was deleted in an attempt to relieve possible repression of transcription (although this resulted in an apparent suppressor mutation in *mibX* that is likely to abolish sigma factor function) (Foulston and Bibb 2011). A shotgun library of *M. corallina* DNA was introduced into an *S. lividans* derivative containing the *mib* gene cluster in an attempt to provide a potentially missing *Microbispora* regulatory gene (Foulston 2010). Yet all of these attempts failed to promote microbisporicin production in the heterologous streptomycete host. Although transcription of genes of the *mib* cluster was detected in *S. lividans*, a Western blot carried out to detect production of the lantibiotic was inconclusive and no peaks corresponding to microbisporicin were observed in MALDI-ToF (Foulston 2010). The similarity between the planosporicin and microbisporicin gene clusters, particularly with regards the likely mechanism of regulation, implies that inferences made about the failure of heterologous expression of the *mib* cluster are likely true for the *psp* cluster too. Thus a full investigation into the reasons for this failure was not pursued here.

It is likely that physiological differences due to the taxonomic distance between *Streptomyces*, a member of the *Streptomycetaceae*, and *P. alba*, a *Streptosporangeacea*, may prevent heterologous production of planosporicin. This is consistent with the expression of the *psp* gene cluster in *Nonomuraea*, a fellow member of the *Streptosporangeacea* family.

However, *Nonomuraea* is not an optimal host for gene inactivation studies of heterologous gene clusters: it produces at least one other molecule with antibiotic activity, and the transfer of cosmids by conjugation from *E. coli* is inefficient. Furthermore, when used to investigate the microbisporicin gene cluster, *Nonomuraea* exconjugants harboring deleted

versions of the *mib* cluster produced both false-positive and false-negative results compared to the phenotype of *M. corallina* deletion mutants (Foulston 2010). In this work, the expression of planosporicin appeared to vary between different exconjugants containing the same integrated construct. Of the three *Nonomuraea* exconjugants with integrated pIJ12328, only one produced planosporicin. It is possible that despite streaking all three exconjugants several times on selection, the WT genotype was not completely removed from two exconjugants. It is possible the WT genotype in the mixed genome became the predominant form during growth without selection to provide samples for bioassay and MALDI-ToF analysis. In contrast, the five exconjugants of *Nonomuraea* with integrated pIJ12329 all produced planosporicin but exhibited different levels of production. A mixed population of WT and integrant genomes could account for this. However the copy number of the *psp* gene cluster could also differ between exconjugants due to tandem integration of cosmids. Although analysis of production in this work was not quantitative, heterologous production of novobiocin in *S. coelicolor* revealed that derivatives with tandem integrations gave higher production levels than those with a single-copy of the gene cluster (Eustaquio *et al.* 2005).

This work provides some evidence for increased conjugation efficiency with smaller vectors. Only two exconjugants of *Nonomuraea* with integrated pIJ12323 and one exconjugant of *Nonomuraea* with integrated pIJ12324, pIJ12325 or pIJ12326 were identified despite many conjugations with *E. coli* harbouring cosmids of ~46.5 kb. In contrast, the smaller vectors pIJ12327, pIJ12328 and pIJ12329, which varied from 25 to 40 kb, gave many more exconjugants for analysis. The number of *Streptomyces* exconjugants obtained with vectors that integrate site-specifically into the chromosome had previously been found to vary with the size of the construct used. *S. coelicolor* yielded more exconjugants with a 7.8 kb plasmid compared to one which is 11.5 kb, however *S. lividans* showed the opposite trend, yielding more exconjugants with the larger plasmid (Flett *et al.* 1997). It was suggested that an additional restriction site in the larger plasmid was recognised by a *S. coelicolor* endonuclease, reducing the efficiency of conjugal transfer of the larger plasmid. This Chapter suggests that above a certain threshold size, likely ~40 kb, the efficiency of mating decreases dramatically. Thus the single most effective method to increase the number of *Nonomuraea* exconjugants obtained is to reduce the size of the construct.

Chapter 4 described the identification of the planosporicin biosynthetic gene cluster while Chapter 5 gave a full analysis of the constituent genes. This Chapter confirms that this cluster is sufficient for production of planosporicin in a heterologous host. Transfer of a cosmid containing the entire *psp* gene cluster to *Nonomuraea* conferred the ability to

produce planosporicin. Shortening the cosmids' inserts to create reduced gene sets served to identify the minimal set of genes required for planosporicin production.

6.7 Summary

pIJ12321 (cosmid B4-1) and pIJ12322 (cosmid F13-1) carrying insert DNA from *P. alba* were PCR-targeted to enable conjugation and stable integration at the Φ C31 site of actinomycete recipients, creating pIJ12323 and pIJ12324, respectively. Both cosmids were mobilised into *S. lividans* TK24, *S. coelicolor* M1146 and *Nonomuraea* ATCC 39727.

Both *Streptomyces* species failed to produce planosporicin heterologously under any conditions tested, whereas *Nonomuraea* M1295 (harboring pIJ12323) produced planosporicin in an active, exported form as detected by bioassay and MALDI-ToF analysis.

Genes flanking the *psp* cluster were removed from pIJ12321, creating three reduced gene sets pIJ12327, pIJ12328 and pIJ12329.

Nonomuraea strains M1299, M1300 and M1301 (harboring pIJ12327, pIJ12328 and pIJ12329, respectively) all produced planosporicin. Thus the two genes adjacent to *pspE* that were removed from pIJ12329 are not necessary for heterologous production in *Nonomuraea*.

pIJ12325 and pIJ12326 carrying DNA from *P. sp.* DSM 14920 were mobilised into *Nonomuraea* ATCC 39727. Bioassay and MALDI-ToF analysis confirmed the production of planosporicin by the resulting exconjugants.

Chapter 7 : Mutational analysis of planosporicin biosynthesis

7.1 Introduction

The functional analysis of biosynthetic pathways is generally investigated through a combination of heterologous host gene expression and native host gene inactivation. Chapter 6 described the difficulties encountered when striving for planosporicin production in a heterologous host. Although *Nonomuraea* did produce planosporicin, there are several reasons for not pursuing a full analysis of the cluster through individual knockout mutations in this host. Not only does *Nonomuraea* produce its own antibiotic, A40926, it is slow growing and does not sporulate. Consequently, conjugations must be performed with mycelium, which often leads to differences in copy number between clones. Furthermore, if part of the predicted biosynthetic gene cluster is not essential for production in the heterologous host, this may be due to host-encoded genes complementing particular functions. For example, the analysis of the cinnamycin gene cluster in the heterologous host *S. lividans*, revealed several genes that were not essential for biosynthesis in the heterologous host. This prompted subsequent analysis in the natural producer to check if the same phenotype was observed (S. O'Rourke, unpublished).

On balance, although *P. alba* is slow growing, taking up to a week for workable colonies on agar media or sufficient mycelial biomass in liquid culture, it does sporulate. Furthermore, the development of tools for the direct genetic manipulation of *P. alba* presents an opportunity for knowledge-based improvement of planosporicin production in this strain.

Prior to this work, the genetic manipulation of the *Planomonospora* genus had not been demonstrated. Fundamental to the generation of deletion mutants in *P. alba* was the development of methods to mobilise vectors into *P. alba* via conjugation with *E. coli* and the use of resistance markers to select for homologous recombination events. The complementation of these mutants requires the use of a phage integration site in which to integrate the complementation construct. This Chapter first investigates the genetic tractability of *P. alba* and defines the genetic tools available for *P. alba* genetic modification. The remainder of the Chapter focuses on the creation of a number of deletion mutations in planosporicin biosynthetic genes.

7.2 Methods and tools developed for use with *P. alba*

The PCR-targeting technique was originally developed and optimised for use with *Streptomyces*, so modifications were needed to develop a robust conjugation method for *Planomonospora*. It is widely recognised that conjugation is much more efficient than transformation of protoplasts and can be applied to many actinomycetes (Matsushima and Baltz 1996). Although protoplasts from a couple of *Streptosporangiaceae* (*M. corallina* and *Nonomuraea* ATCC 39727) have been made, regenerated and transformed, conjugation into mycelia was subsequently found to be a more efficient method (Foulston and Bibb 2010; Marcone *et al.* 2010b; Marcone *et al.* 2010c). As protoplast transformation has not previously been attempted in *Planomonospora*, conjugation of plasmid constructs from *E. coli* into *P. alba* was the method of choice for the generation of mutant strains. It is not known if *P. alba* possesses a methyl-specific restriction system. In case it does, constructs were routinely passaged through methylation-deficient *E. coli* ET12567/pUZ8002 prior to conjugation into *P. alba* (Macneil *et al.* 1992).

7.2.1 Conjugation methods

The genetic manipulation of *P. alba* requires at least two markers that can be selected both in *E. coli* and *P. alba*. To this end, the antibiotics routinely used as selectable markers in *Streptomyces* genetics were tested to see if *P. alba* is naturally resistant. Individual ISP4 agar plates were made, each containing an antibiotic at the concentration commonly used for selection in other actinomycetes (Kieser *et al.* 2000). *P. alba* was unable to grow on 50 µg/ml apramycin, 40 µg/ml hygromycin, 200 µg/ml spectinomycin with 10 µg/ml streptomycin, 30 µg/ml viomycin, 50 µg/ml thiostrepton or 50 µg/ml kanamycin. These antibiotics were therefore deemed to be suitable for selection of DNA introduced into *P. alba*. Furthermore, *P. alba*, like other actinomycetes, grew in the presence of 50 µg/ml nalidixic acid, so this antibiotic can be used to selectively kill *E. coli* after the conjugation event.

The following text summarises the results obtained in preliminary experiments aimed at developing a workable conjugation protocol for *P. alba*. The conjugative vectors used were generally pIJ10706 and pMS82 with selection for hygromycin resistance.

A number of factors were optimised for the conjugation of *P. alba*. Although preparation of the donor *E. coli* strain was not altered, the preparation method of the recipient was investigated. Constructs were successfully conjugated both into spores and freshly grown *P. alba* mycelium. *P. alba* mycelium was collected at mid to late exponential phase (after 48 hours of growth) and washed in glycerol before mating. This method yielded similar

numbers of exconjugants to conjugation with *P. alba* spores thawed from storage at -20 °C in 20 % glycerol. A marginal increase in the number of exconjugants was observed if the spores were subjected to heat shock (to synchronise germination) prior to conjugation. The number of exconjugants was sometimes observed to increase if the spores were incubated at 37 °C 250 rpm for 2.5 hours after heat shock to encourage the growth of germ tubes. Light microscopy images confirmed the emergence of germ tubes after heat shock. These germ tubes became longer, often with branching and tended to clump together after further incubation. Freshly emerged germ tubes contain few mycelial compartments, so the conjugation event is more likely to produce exconjugants which are homogenous clones. Thus the use of heat shocked spores was favoured for the majority of matings.

The appropriate medium for the preparation of *P. alba* spores and growth of exconjugants was next investigated. The recipient conjugated equally well when mycelium was prepared in ISP4, D/seed, M8 or AF/MS liquid media, but not in the 2xYT broth recommended for *Streptomyces*. The addition of 10 mM MgCl₂ to the agar medium used to plate out conjugations was known to promote conjugation between *Streptomyces* and *E. coli* (Kieser *et al.* 2000). Conjugation mixtures were thus plated out on a range of media all containing 10 mM MgCl₂. The agar media tested were those giving the best growth of *P. alba* (ISP4, M8, D/seed and M8 agar), as well as a range of media used for other actinomycetes (SFM, OBM, MV, R2 and R5 agar). *P. alba* vegetative mycelium was able to grow on all media, although aerial mycelium and sporulation was only observed on ISP4, OBM and MV agar. Exconjugants were routinely observed on ISP4 and often on M8 and D/seed agar. Thus for the majority of conjugations, matings were plated onto ISP4 agar containing 10 mM MgCl₂.

The ratio of the donor *E. coli* to recipient spores was also investigated. In *Streptomyces* conjugations, commonly 10⁸ donor cells and 10⁸ recipient spores are used. This mating is so efficient, serial dilutions are plated to obtain individual exconjugants, or alternatively the number of spores used is reduced. It had been shown previously that the number of exconjugant colonies increased in proportion to the number of spores used (up to a maximum of 10⁸), while a reduction in the initial number of donor cells by a factor of 10 did not affect the exconjugant frequency markedly (Flett *et al.* 1997). In *P. alba* conjugations, no significant difference in conjugation efficiency was observed when the ratio of donor *E. coli* to recipient spores was altered. The incubation time for each mating was also varied. After plating on ISP4 agar containing 10 mM MgCl₂, the conjugation mixture of *E. coli* and *P. alba* was incubated for 18 to 22 hours at 30 °C. Plates were then overlaid with 1 ml water containing 0.5 mg nalidixic acid to kill the *E. coli* donor and the relevant antibiotic to select for exconjugants. With longer incubation times, the haze of *E. coli* growth became

more visible, but the number of exconjugants and their time to emerge did not vary significantly. Thus conjugation plates were routinely overlaid after 20 hours growth at 30 °C. Single colonies of *P. alba* were routinely visible after approximately two weeks growth, although a longer incubation time was often required for sporulation of these colonies (Figure 7.1). Plates overlaid only with nalidixic acid revealed lawns of *P. alba* (Figure 7.1), indicating that viable *P. alba* is present in the conjugation mixture. The final, optimised protocol with revised spore preparation and media usage is detailed in Chapter 2.

7.2.2 Integration sites

The *Streptomyces* temperate phages Φ C31 and Φ BT1 each possess an *attP-int* locus that facilitates integration into conserved *attB* sites present in recipient genomes. In *S. coelicolor*, the Φ C31 *attB* lies within SCO3798, while the Φ BT1 *attB* lies within SCO4848 (Gregory *et al.* 2003). Although both phage integrate intragenically, the genes appear non-essential for *S. coelicolor* growth in the laboratory (Baltz 1998). The Φ C31 and Φ BT1 *attP-int* loci have been incorporated into numerous vectors that are used in a broad range of actinomycetes.

The aim in these studies was to use one or both of these integration sites to introduce a WT copy of a gene to complement various *P. alba* mutants. However initial 454 data coverage (carried out by the University of Liverpool) was not extensive enough to reveal the presence of these *attB* sites in the *P. alba* genome. Instead two constructs were conjugated into *P. alba* to test for their ability to integrate into the host chromosome. pMS82 uses the Φ BT1 *attB* site, while pIJ10706 integrates into the Φ C31 *attB* site, and both vectors contain the hygromycin resistance cassette (Table 7.1). Hygromycin-resistant exconjugants were obtained from both matings. Colony PCR (Figure 7.2) of 12 putative pMS82 exconjugants picked directly from a conjugation plate using primers annealing within pMS82 confirmed the presence of the vector. Subsequent extraction of genomic DNA (gDNA) from *P. alba* exconjugants containing either pMS82 or pIJ10706 followed by PCR with primers internal to the vector, confirmed their presence in each exconjugant (Figure 7.3).

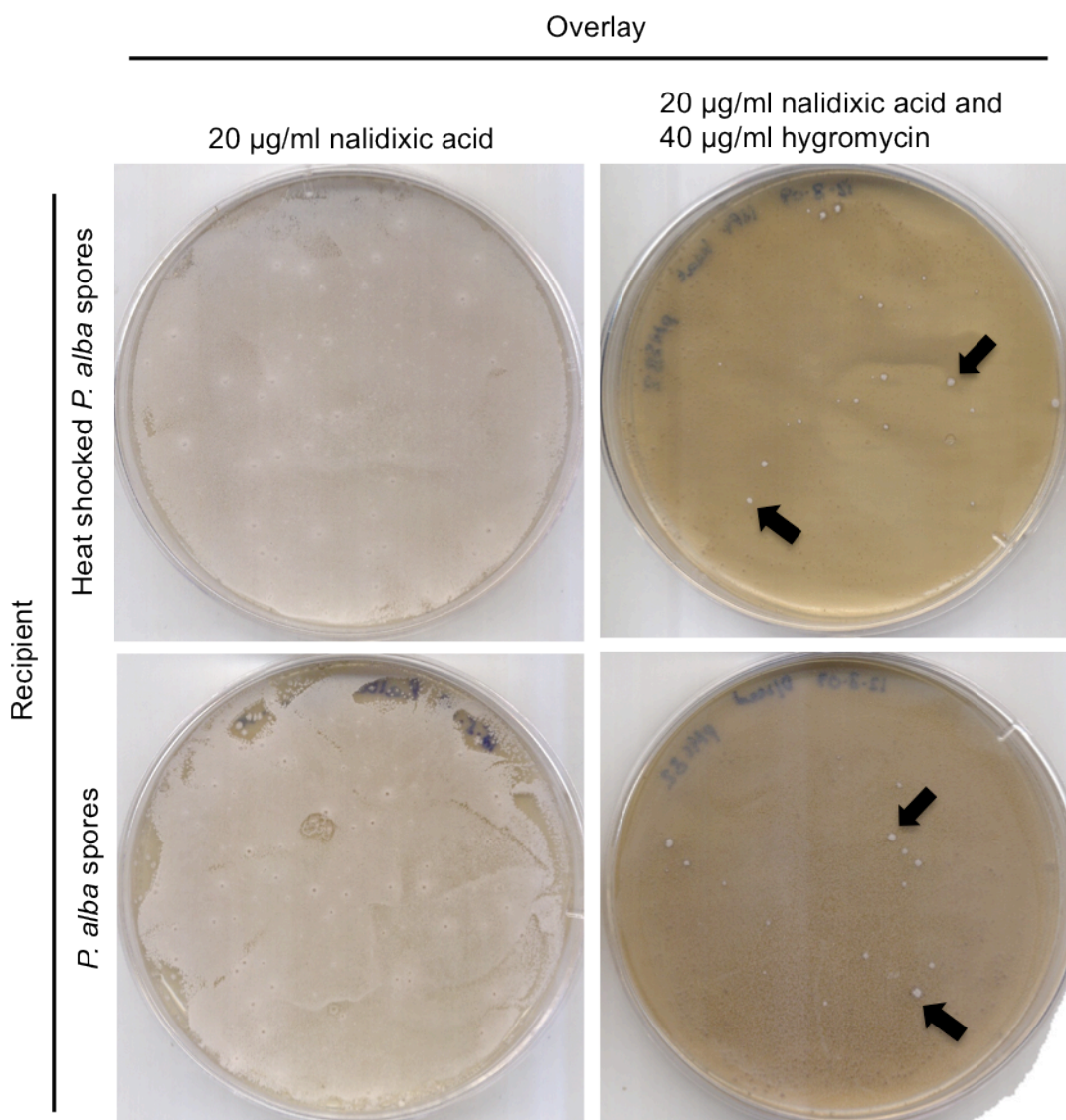


Figure 7.1 : Conjugation with *E. coli* ET12567 pUZ8002 harbouring pMS82.

P. alba spores were used directly in conjugation or subject to 50 °C heat shock for 10 minutes, cooled and mixed with *E. coli*. All conjugations were plated on ISP4 agar containing 10 mM MgCl_2 . After 20 hours at 30 °C, plates were overlaid with 1 ml H_2O containing antibiotics to give a final concentration of either 20 $\mu\text{g/ml}$ nalidixic acid alone (left hand side) or 20 $\mu\text{g/ml}$ nalidixic acid and 40 $\mu\text{g/ml}$ hygromycin (right hand side). After a further 7 days incubation at 30 °C, growth of *P. alba* vegetative mycelium was visible on the plates containing just nalidixic acid. After a further 10 days, a lawn of sporulating *P. alba* was visible on the same plates. After a further 5 days, individual colonies had grown through the hygromycin selection and sporulated (indicated with arrows). Negative control plates were photographed after a total of 17 days at 30 °C. Hygromycin selective plates were photographed after a total of 22 days at 30 °C.

Construct	Integration site	Resistance gene	Size (Kb)	Exconjugants
pSET152	ΦC31	<i>aac(3)IV</i>	5.7	no
pIJ10706	ΦC31	<i>hyg</i>	5.9	yes
pSETΩ	ΦC31	<i>aadA</i>	7.0	no
pMS82	ΦBT1	<i>hyg</i>	6.1	yes
pRT801	ΦBT1	<i>aac(3)IV</i>	5.1	no
pIJ10702	ΦC31	<i>aac(3)IV</i>	9.8	no
pIJ12330	ΦC31	<i>hyg, aac(3)IV</i>	7.0	yes*
pIJ12331	ΦC31	<i>hyg, aadA</i>	7.0	yes*
pIJ12332	ΦC31	<i>hyg, vph</i>	7.1	yes*
pIJ12333	ΦC31	<i>hyg, tsr</i>	7.0	yes*

Table 7.1 : A list of the site-specific integrative vectors conjugated into *P. alba*.

The resistance genes correspond to apramycin resistance (*aac(3)IV*), spectinomycin/streptomycin resistance (*aadA*), hygromycin resistance (*hyg*), viomycin resistance (*vph*) and thiostrepton resistance (*tsr*). *Exconjugants were only obtained if selected with hygromycin.

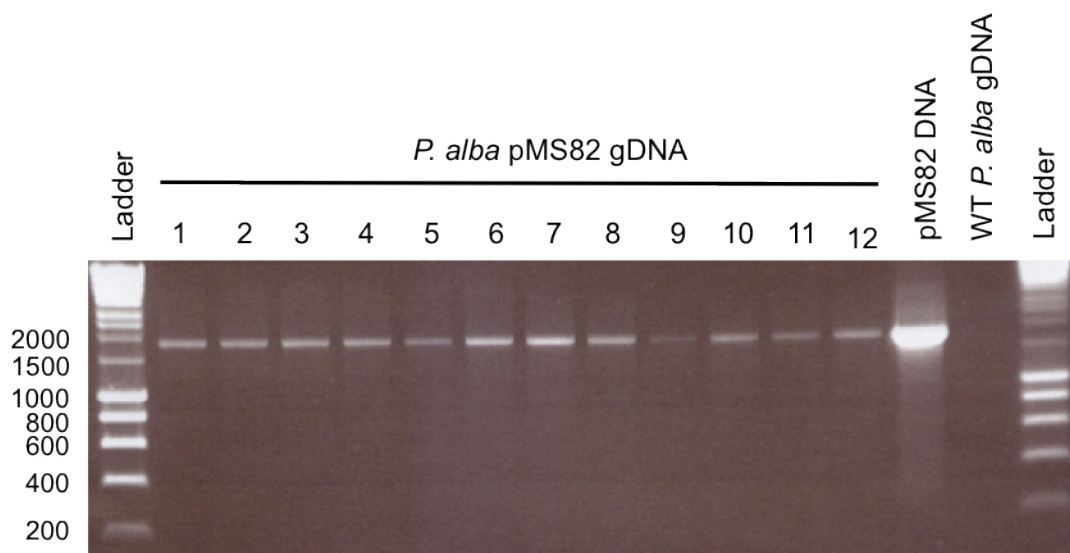


Figure 7.2 : Colony PCR analysis to confirm the integration of pMS82 into the genome of *P. alba*.

Vegetative mycelium from 12 putative *P. alba* exconjugants was suspended in DMSO and used as a template to amplify an 1817 bp fragment of pMS82 (with primers pMS82forB and pMS_Rseq). pMS82 plasmid DNA was used as a positive control and WT *P. alba* gDNA as a negative control. PCR products were run on a 1 % agarose gel by electrophoresis. The ladder is Hyperladder I (Biolone) with band sizes annotated in bp.

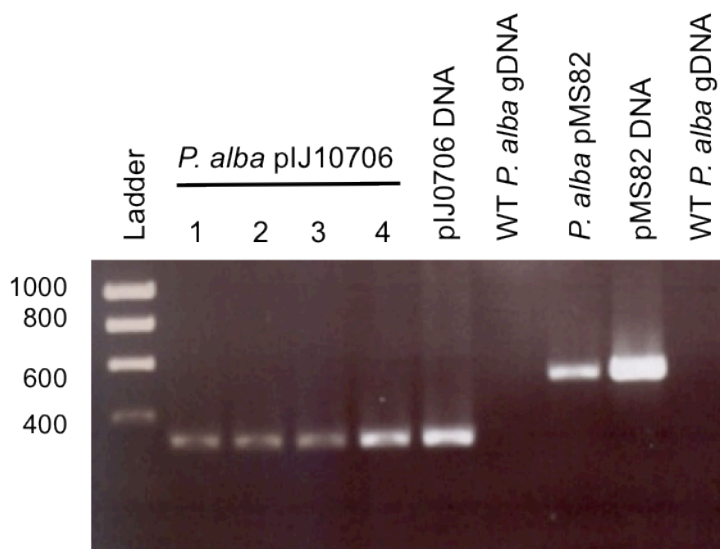


Figure 7.3 : PCR analysis to confirm the integration of pIJ10706 into the *P. alba* genome. Genomic DNA extracted from 4 exconjugants was used as a template to amplify a 323 bp fragment of pIJ10706 (primers pSETF and pSETR). pIJ10706 DNA was used as a positive control and WT *P. alba* gDNA as a negative control. In parallel, genomic DNA extracted from *P. alba* in which colony PCR had indicated that pMS82 had successfully integrated into the host genome was used as a template to amplify a fragment within pMS82 (pMS82forB and pMS_Rseq; 529 bp). pMS82 plasmid DNA was used as a positive control and WT *P. alba* gDNA as a negative control. PCR products were run on a 1 % agarose gel by electrophoresis. The ladder is Hyperladder I (Bioline) with band sizes annotated in bp.

Subsequently, 454 next generation sequencing was repeated at The Genome Analysis Centre (as described in Chapter 4) to provide greater genome coverage. The Φ BT1 *attB* site of *S. coelicolor* is a 73 bp sequence in which the crossover site occurs between bases 35-43. The *P. alba* 454 contig database was searched using this 73 bp sequence as a query. However no significant hits were found. Subsequently the nucleotide sequence of SCO4848, the gene containing the Φ BT1 *attB* site of *S. coelicolor*, was used as a BLASTN input, but again no *P. alba* contigs were identified. In parallel, 43 bp of the Φ C31 *attB* site of *S. coelicolor* was used as a BLASTN input to search the 454 contig database. This analysis identified a 22 bp region within contig00142 which has 95 % identity across 22 bp of the *S. coelicolor* Φ C31 *attB* site (Figure 7.4). Located centrally within the aligned region is a TT pair surrounded by GC nucleotides which are also conserved in the Φ C31 *attB* site of other *Streptomyces* sp. (Combes *et al.* 2002).

```

S. coelicolor 16      GCGTGCCCTTGGGCTCCCCGG 37
                |||||
P. alba       5417    GCGTGCCCTTGGGCACCCCGG 5438

```

Figure 7.4 : An alignment of a section of the *S. coelicolor* Φ C31 insertion site with a region of contig00142 from the *P. alba* 454 contig database.

The TT pair conserved in the Φ C31 *attB* site of other *Streptomyces* sp. is in bold.

Interestingly, when the nucleotide sequence of SCO3798, the gene containing the Φ C31 *attB* site of *S. coelicolor*, was used as a BLASTN input no *P. alba* contigs were identified. It would appear that the *attB* site has a significantly different context in *P. alba* compared to *S. coelicolor*. The presence of only a small conserved region makes database searching difficult. Thus while the more recent 454 data revealed the presence of a likely Φ C31 *attB* site, it was unable to locate a contig containing a Φ BT1 *attB*. Although the putative Φ C31 *attB* site identified through genome scanning of the *P. alba* genome has not been verified experimentally as the site used for integration, the successful integration of both vectors suggests the existence of chromosomal attachment sites for each in the *P. alba* genome. Herein the successful conjugation of integrative plasmids is assumed to occur through site-specific integration of the constructs into the *P. alba* chromosome.

7.2.3 Resistance cassettes

To adapt the procedure of λ RED mediated recombination from *Streptomyces* to *Planomonospora*, cassettes for gene disruptions that can be selected in *E. coli* were assessed for their ability to be selected in *Planomonospora*. It was important to be able to select for at least two different resistance genes; one to select for the targeted replacement of the gene with the cassette by homologous recombination and one to select for the integration of the complementation construct. This work tested a number of constructs containing different resistance markers (Table 7.1).

Section 7.2.2 revealed that conjugations were successful with pIJ10706 and pMS82 vectors, with integration achieved by selecting for resistance to hygromycin (although there is no experimental evidence that this integration was site-specific). To investigate which other resistance cassettes could be used, *P. alba* was conjugated with *E. coli* harbouring one of three further vectors; pRT801, pSET152 and pIJ10702. All vectors contained the apramycin resistance cassette but while pRT801 integrates into the Φ BT1 *attB* site, pSET152 and pIJ10702 integrate into the Φ C31 *attB* site (Table 7.1). No apramycin-resistant exconjugants were obtained from either mating. In addition a fourth

vector, pSET Ω , which integrates into the Φ C31 site was mobilised from *E. coli* into *P. alba* (Table 7.1). pSET Ω contains the *aadA* gene for spectinomycin/streptomycin (spec/strep) resistance; however no exconjugants were obtained after this mating.

To further investigate this, four more constructs based on pIJ10706 were made, each containing an additional resistance gene. Vectors pIJ12330, pIJ12331, pIJ12332 and pIJ12333 contain the additional genes *aac(3)IV*, *aadA*, *vph* and *tsr*, conferring resistance to apramycin, spec/strep, viomycin and thiostrepton, respectively (Table 7.1 and Figure 7.5), and their construction is described below.

Primers BamHIclone and XbaIclone were designed to anneal either side of the selectable markers present in pIJ773 (*aac(3)IV*), pIJ778 (*aadA*) and pIJ780 (*vph*). The primers were designed to include the promoter region of each resistance gene but exclude the *oriT* site, as this is already present in pIJ10706. The LongAmp *Taq* DNA polymerase (NEB) was used to amplify *aac(3)IV*, *aadA* and *vph* and introduce *Bam*HI and *Xba*I. restriction sites 5' and 3', respectively, of each resistance gene. LongAmp consists of a blend of *Taq* and Deep Vent_RTM DNA Polymerases. Deep Vent polymerase exhibits 3' to 5' exonuclease activity to increase the fidelity of *Taq* polymerase which adds a single nucleotide base extension (usually adenine) to the 3' end of amplified DNA fragments. The A-tail of the PCR products was used to ligate the amplified fragment into the commercial vector pGEM-T for blue-white screening (as described in Chapter 4). Four white colonies from each ligation were subjected to colony PCR with the pSET152F and pSET152R primers to successfully amplify across the *lacZ* α gene and confirm that each fragment had indeed been inserted into the vector (data not shown). Plasmid DNA was prepared from one clone of pGEM-T *aac(3)IV*, pGEM-T *aadA* and pGEM-T *vph*. The insert was liberated from the vector backbone using a double digest with *Bam*HI and *Xba*I and isolated from an agarose gel. pIJ10706 was similarly digested with *Bam*HI and *Xba*I, gel extracted and used in a ligation with the three different inserts. pSET152F and pSET152R primers amplified a 295 bp fragment from pIJ10706, whereas the integration of 1045 bp *aac(3)IV*, 1082 bp *aadA* and 1150 bp *vph* would increase the fragment size amplified accordingly. Successful ligation of *aac(3)IV*, *aadA* and *vph* into pIJ10706 to create pIJ12330, pIJ12331 and pIJ12332, respectively (Figure 7.5), was subsequently confirmed by PCR (data not shown). For further confirmation, plasmid DNA was purified from two clones of each construct. A diagnostic digest with *Xba*I and *Bam*HI excised the insert from pIJ10706, revealing two bands of the expected sizes on an agarose gel. High-fidelity PCR with pSET152F and pSET152R amplified a fragment which was sequenced to confirm that the correct fragment had been ligated into each construct. Each plasmid was introduced into *E. coli* ET12567/pUZ8002 by transformation. Two clones from each transformation were confirmed by PCR and one of each was chosen for conjugation into *P. alba*.

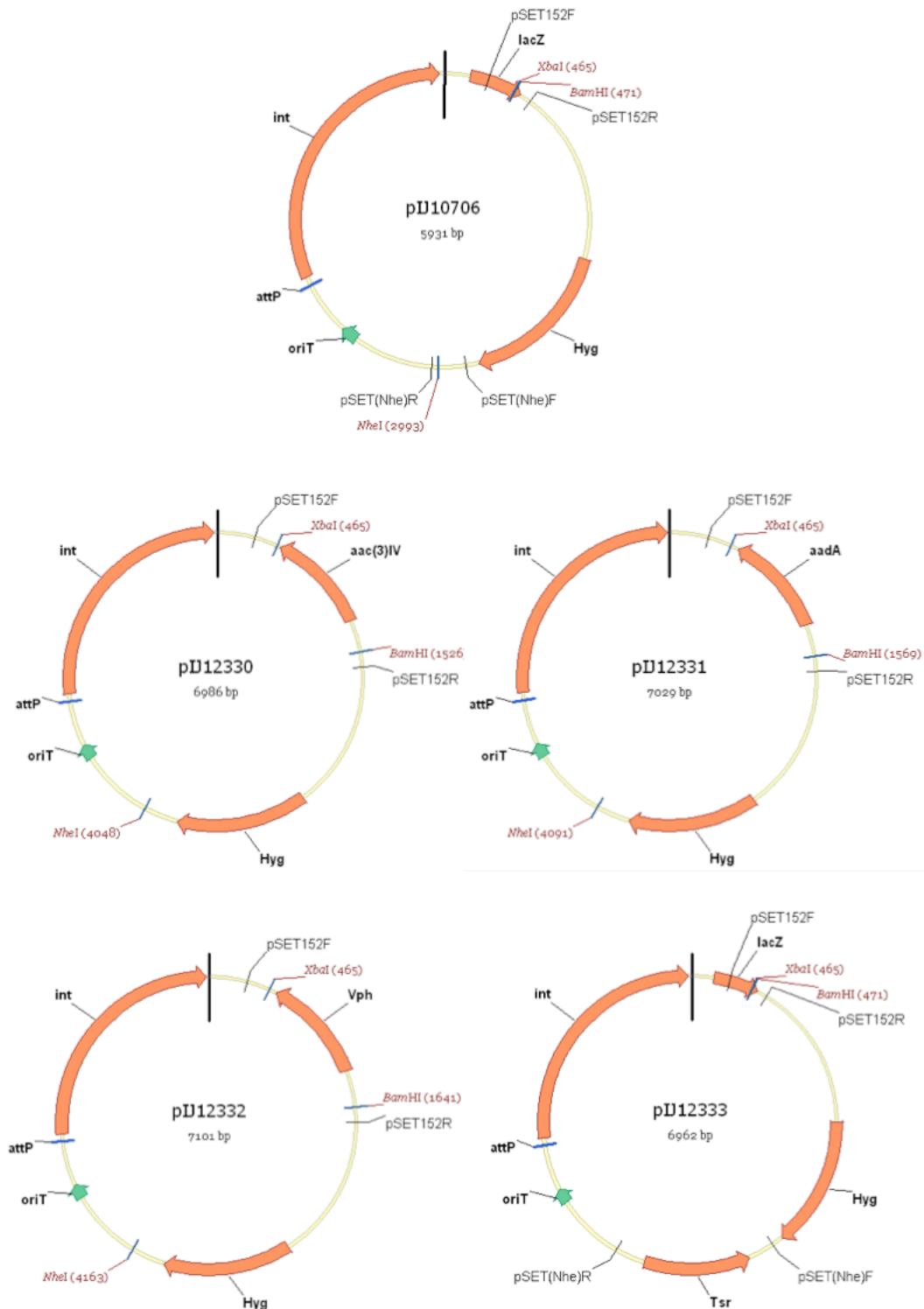


Figure 7.5 : Schematic of four new constructs based on pIJ10706.

pIJ12330, pIJ12331 and pIJ12332 have the *aac(3)IV*, *aadA* and *vph* resistance genes, respectively, cloned in the *BamHI*-*XbaI* sites of pIJ10706. PCR with pSET152F and R primers amplified 302 bp from pIJ10706, 1357 bp of *aac(3)IV* from pIJ12330, 1400 bp of *aadA* from pIJ12331 and 1472 bp of *vph* from pIJ12332 (data not shown). pIJ12333 has the *tsr* resistance gene cloned in the *NheI* site of pIJ10706. Thus PCR with pSET(Nhe)F and R primers amplified 220 bp from pIJ10706 and 1251 bp of *tsr* from pIJ12333 (data not shown).

In contrast, different methodology was used to create pIJ12333; instead of amplifying the *tsr* resistance cassette, it was excised as a 1037 bp *NheI* fragment from pIJ6902 and the isolated fragment ligated directly into the single *NheI* site of similarly digested pIJ10706. Primers pSET(Nhe)F and pSET(Nhe)R were designed to amplify across the *NheI* site in pIJ10706. PCR on pIJ10706 revealed the expected 220 bp band, while the successful ligation of *tsr* into the *NheI* site of the vector resulted in the amplification of a 1257 bp fragment (Figure 7.5). Three colonies were checked by PCR across the *NheI* site. In addition, plasmid DNA was purified and high-fidelity PCR amplified a fragment that was sequenced. All three colonies had *tsr* integrated into the *NheI* site. One clone was designated pIJ12333 and introduced into *E. coli* ET12567/pUZ8002 by transformation to be used as the donor for conjugation into *P. alba*.

The four constructs pIJ12330, pIJ12331, pIJ12332 and pIJ12333 were integrated into the *P. alba* genome using hygromycin to select for exconjugants. *P. alba* exconjugants containing the integrated plasmid were subsequently tested for resistance to other antibiotics.

P. alba exconjugants containing pIJ12330 and pIJ12331 were relatively easy to isolate. Four exconjugants from each of the two conjugations were streaked three times on ISP4 agar whilst maintaining selection with hygromycin. Two exconjugants from each conjugation were subject to gDNA extraction. PCR with primers pSET152F and pSET152R confirmed the presence of *aac(3)IV* and *aadA* integrated into the *P. alba* genome. Spores from each exconjugant were tested for resistance to hygromycin, apramycin and spec/strep. *P. alba* harbouring pIJ12330 conferred hygromycin and apramycin resistance, while *P. alba* harbouring pIJ12331 conferred resistance to hygromycin and spec/strep (Figure 7.6). Although direct selection for apramycin or spec/strep resistance in conjugations with pSET152 and pSET Ω was unsuccessful, *P. alba* is evidently able to express these resistance genes, enabling selection to be maintained through use of these antibiotics. However the spec/strep selection appears unreliable. Colonies of WT *P. alba* and of *P. alba* containing pIJ10706 grew through the spec/strep selection after 22 days incubation whereas no such colonies were observed after further incubation of the same strains grown on ISP4 agar containing apramycin. Furthermore the *aac(3)IV* cassette of pIJ12330 appears to confer some cross-resistance to spec/strep (Figure 7.6). In contrast, there is no reported cross-resistance conferred by *aac(3)IV* on spectinomycin or streptomycin in *S. lividans* or *S. coelicolor* (Kieser *et al.* 2000). The enzyme encoded by *aac(3)IV* catalyses the acetyl CoA-dependent acetylation of an amino group of apramycin, thus inactivating the antibiotic. As spectinomycin and streptomycin are also aminoglycosides, AAC(3)IV may conceivably modify one or both of these antibiotics confer some resistance in *P. alba*. Thus, of these two possible additional

Conjugations to mobilise pIJ2330 and pIJ2331 into *P. alba* were less efficient. Only one exconjugant was obtained with pIJ2332 and only two with pIJ12333. All three exconjugants were subjected to gDNA extraction to provide a template for PCR. *vph* and *tsr* genes integrated into the *P. alba* genome were successfully amplified from the appropriate exconjugants (data not shown). Spores from each exconjugant were tested for resistance to hygromycin, viomycin and thiostrepton. *P. alba* harbouring pIJ12332 conferred hygromycin and viomycin resistance, while *P. alba* harbouring pIJ12333 conferred resistance to hygromycin and thiostrepton (Figure 7.7).

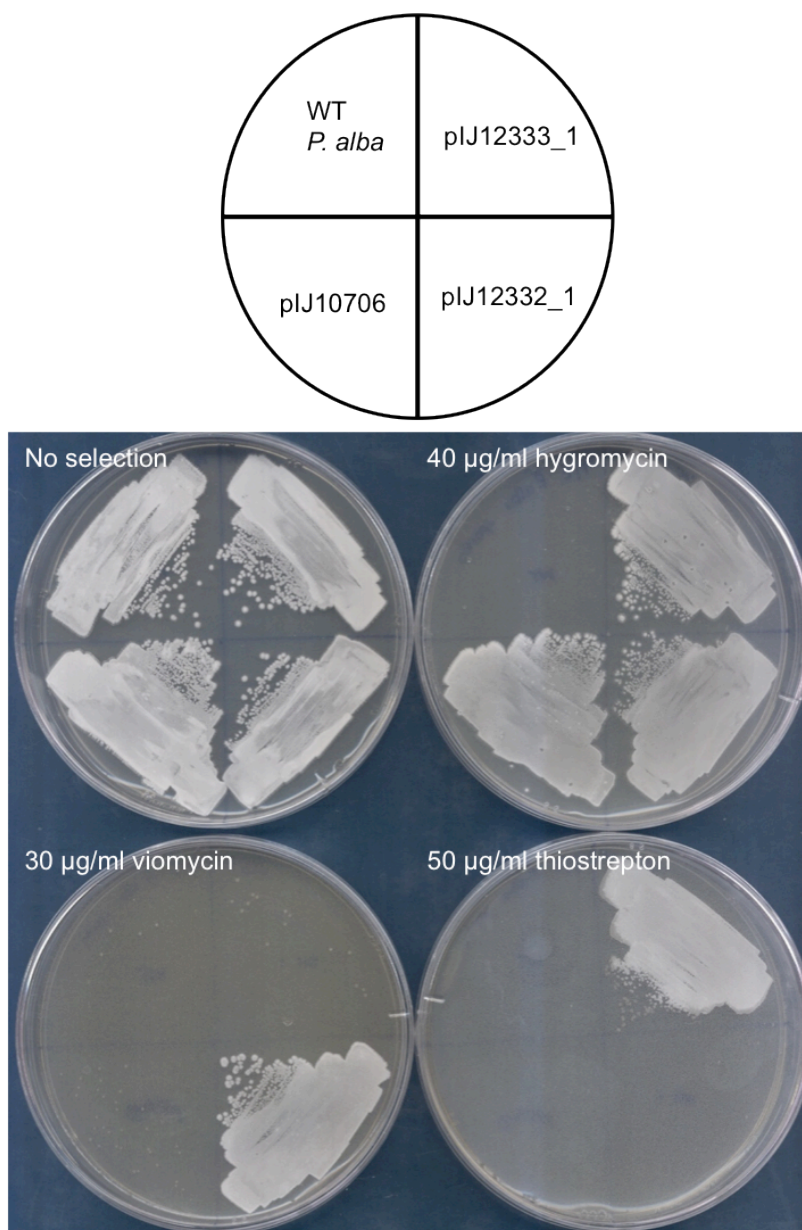


Figure 7.7 : Growth of *P. alba* strains with antibiotic selection.

Spores from exconjugants of *P. alba* with integrated pIJ12332 or pIJ12333 were streaked onto ISP4 agar containing either; no selection, hygromycin, viomycin or thiostrepton. WT *P. alba* and *P. alba* with integrated pIJ10706 were used as controls. Plates were incubated for 7 days at 30 °C then photographed.

Subsequently all four conjugations were repeated but with direct selection for apramycin, spec/strep, viomycin or thioestrepton in the overlay. Control conjugations were also carried out in parallel selecting for hygromycin resistance. Although exconjugants were obtained when selecting for hygromycin, none were obtained when selecting for the alternative resistance marker. The reason why direct selection with these antibiotics is not successful remains unclear. It could be that the expression of the resistance genes was not high enough to enable growth through primary selection, but is sufficient for maintenance.

These experiments indicated that the hygromycin resistance cassette worked most effectively in *P. alba*. This cassette was therefore used to make targeted deletions in the planosporicin biosynthetic gene cluster located in the *P. alba* genome. A second resistance cassette is also needed to select for integration of a plasmid harbouring a WT copy of a *psp* gene in order to complement the mutation *in trans*. Apramycin resistance is widely used in *Streptomyces* genetics, and many tools are available containing this marker, making it the obvious choice. pSET152 and pIJ10706 differ only in the coding sequences of the resistance genes they carry (apramycin and hygromycin, respectively). Both plasmids not only share the same integration site, but also the same promoter region for the selective marker (it had been shown previously that the apramycin gene promoter gave greater resistance to hygromycin than the native promoter when used to transcribe *hyg*, which is why it was used in the construction of pIJ10706 (S. O'Rourke, personal communication)). On the basis that *aac(3)IV* is indeed expressed in *P. alba* but not at a high enough level to enable primary selection, attempts were made to select with lower concentrations of apramycin. The minimum inhibitory concentration (MIC) of apramycin in ISP4 agar was first investigated. *P. alba* was sensitive to an overlay containing 0.25 mg apramycin in 1 ml water (giving a final concentration of 10 µg/ml in a 25 ml ISP4 plate), which is five times lower than the recommended amount of apramycin in *Streptomyces* conjugations. While no exconjugants were obtained when conjugations of pSET152 into *P. alba* were overlaid with 1 ml of water containing 1.25 mg of apramycin (50 µg/ml final concentration), several sporulating *P. alba* colonies grew through an overlay with reduced amounts of apramycin. Ten colonies were obtained from the 1 ml overlay with 0.3 mg apramycin (12 µg/ml final concentration), four colonies from 0.4 mg apramycin (16 µg/ml final concentration), and three colonies from 0.5 mg apramycin (20 µg/ml final concentration). These putative exconjugants were streaked out three times on ISP4 agar containing 25 µg/ml nalidixic acid and 50 µg/ml apramycin. gDNA was isolated from four of these exconjugants (representing each reduced level of selection) and used in the PCR with primers pSET152F and pSET152R; in each case a 323 bp fragment of pSET152 was amplified. Thus reducing the quantity of apramycin used in the overlay to a level only marginally above the MIC allowed the growth of apramycin resistant exconjugants. Repeated conjugations with pSET152 revealed that occasionally patches of the

conjugation mixture would grow through 1 ml overlays containing 0.3 or 0.4 mg apramycin. Thus the integration of complementation constructs based on pSET152 was subsequently selected with 1 ml overlays containing 0.5 mg apramycin.

7.3 Targeted deletions within the planosporicin gene cluster of *P. alba*

Recombination-based genetic engineering (recombineering) enables the replacement of a section of a bacterial chromosome with a selectable marker. The canonical procedure developed for *Streptomyces* at the John Innes Centre by Bertold Gust enabled the creation of precise gene replacements *in vivo* (Gust *et al.* 2003). This efficient method circumvents time consuming digestion and ligation-mediated techniques *in vitro*. The procedure can be split into three steps. Firstly, primers are designed which have 19 or 20 nt 3' ends that anneal to a disruption cassette and 39 nt 5' extensions corresponding to the sequence flanking the gene to be replaced. These primers amplify the disruption cassette containing an *oriT* and a resistance gene which is selectable both in *E. coli* and the target organism, in this case *Planomonospora*. Secondly, the amplified fragment is used to transform *E. coli* BW/pIJ790 harboring a cosmid carrying the gene to be disrupted. As described in Chapter 6, pIJ790 expresses λ RED genes which prevent the degradation of small, linear DNA fragments. As a result, the target gene is replaced with the PCR-amplified DNA. Thirdly, the cosmid containing the disrupted gene is conjugated into the target strain (here *Planomonospora*) via the methylation-deficient *E. coli* donor strain ET12567/pUZ8002. The non-transmissible pUZ8002 acts *in trans* in *E. coli* to mobilise the *oriT* present in the resistance cassette (described in Chapter 6). A double crossover event replaces the target gene with the resistance cassette, creating a targeted deletion mutant. However this strategy depends on a reasonably high frequency of recombination in the targeted organism. Prior to this work, this had not been demonstrated for *P. alba*.

Chapter 5 gave a detailed bioinformatic analysis of the 15 genes constituting the planosporicin biosynthetic gene cluster. These genes were shown to be sufficient for heterologous production of planosporicin in *Nonomuraea* ATCC 39727 in Chapter 6. The demonstration that genes within the cluster are necessary for planosporicin production in the native producer would provide further proof that this gene cluster does indeed encode the biosynthetic machinery for planosporicin production, and would more definitively define the minimal gene set. For this purpose, the gene encoding the planosporicin precursor peptide, *pspA*, was replaced with a resistance cassette using the recombineering method described above. In contrast, *pspB* and *pspC* were not targeted for replacement. There is much experimental evidence showing that LanB is required for dehydration of serine and threonine and LanC for correct (Me)Lan bridge formation. For

example, the deletion of *pepB* and *pepC* from *Staphylococcus epidermidis* 5 led to accumulation of Pep5 prepeptide in the cells without excretion of processed peptide. A *pepC*-deletion clone did not excrete correctly matured Pep5 but it did produce fragments from which serine and threonine were absent but unmodified cysteine residues were present, implying Dha and Dhb had been formed (Meyer *et al.* 1995). Likewise the expression of his-tagged nisin precursors in *nisB* and *nisC* mutant *L. lactis* strains revealed that nisin precursors from the strain lacking NisB activity were totally unmodified, whereas nisin precursors from the strain lacking NisC activity, but having NisB activity, contained dehydrated residues but were devoid of (Me)Lan bridges (Koponen *et al.* 2002). Further evidence for these functions comes from the heterologous expression of combinations of modification enzymes and prepropeptide chimeras in non-producing strains as well as *in vitro* enzymology discussed in Chapter 1. Thus, due to time limitations, this Chapter will focus on targeting only *psp* genes whose biosynthetic function was less certain. These include the three ABC transporters; *pspTU*, *pspYZ* and *pspEF*, and one gene with unknown function, *pspV*.

7.3.1 *pspA*

7.3.1.1 PCR analysis

To confirm the function of the proposed biosynthetic gene cluster in the native host, the proposed *pspA* structural gene was targeted using recombineering as described in Section 7.3. In this first attempt to make a targeted deletion in the *P. alba* genome, three different resistance cassettes were used. All three resistance cassettes are based on the structure: P1-FRT-*oriT*-resistance gene-FRT-P2, in which P1 and P2 are universal priming sites (identical in all disruption cassettes) and FRT is the FLP-recombinase recognition target. The cassette from pIJ773 contains the resistance gene *aac(3)IV*, pIJ778 contains *aadA* and pIJ10700 contains *hyg*. Disruption primers PspAdisruptF and PspAdisruptR were designed containing sequences which matched all three disruption cassettes. PspAdisruptF consists of 39 bp from the sense strand of *pspA* ending in the ATG start codon followed by 20 bp matching the left end of the disruption cassette. PspAdisruptR consists of 39 bp from the anti-sense strand ending in the TGA stop codon followed by 19 bp matching the right end of the disruption cassette. The pIJ773 cassette was used to replace *pspA* in pIJ12321 and pIJ12322 to create pIJ12334 and pIJ12335, respectively. The pIJ778 and pIJ10700 cassettes were used to replace *pspA* in pIJ12321, creating pIJ12336 and pIJ12337, respectively. These constructs were conjugated into *P. alba* and exconjugants were selected with the appropriate antibiotic.

Despite many conjugation attempts with pIJ12334 and pIJ12335, no apramycin-resistant *P. alba* colonies emerged through the overlay on conjugation plates. Somewhat surprisingly, several spectinomycin and streptomycin resistant *P. alba* colonies did emerge after conjugation with pIJ12336 and grew well when streaked on ISP4 agar containing nalidixic acid and spec/strep. However, PCR on gDNA extracted from two exconjugants revealed the presence of the WT *pspA* gene, with no band corresponding to $\Delta pspA::778$. It is likely these colonies acquired spontaneous mutations conferring resistance to 200 $\mu\text{g/ml}$ spectinomycin and 10 $\mu\text{g/ml}$ streptomycin.

Likewise, no hygromycin-resistant *P. alba* exconjugants were obtained after conjugation with pIJ12337. Yet Section 7.2.3 had indicated that the hygromycin resistance cassette would be the most effective for making *P. alba* deletion mutants. Hygromycin resistant exconjugants containing the hygromycin resistance gene integrated into the chromosome were readily isolated from conjugation plates, demonstrating that small, integrative vectors can be conjugated into *P. alba*. When making deletion mutants using recombineering, constructs containing large regions of homologous DNA (e.g. cosmid clones) are often used. While enhancing the probability of recombination, such constructs are likely to conjugate less efficiently than smaller plasmids, potentially precluding the isolation of the desired exconjugants. To address this concern and to potentially improve the efficiency of conjugation, the size of the homologous recombination vector was reduced. pIJ12337 contains 22299 bp of sequence homologous to the region upstream of *pspA* and 15295 bp of sequence homologous to the region downstream of *pspA*. In many actinomycetes, just 2 kb of homologous sequence is sufficient for detectable homologous recombination. Therefore primers were designed to amplify $\Delta pspA::hyg$ with either ~2 kb homologous sequence either side (creating a ~5.5 kb fragment) or ~4 kb homologous sequence (creating a ~9.5 kb fragment). LongAmp polymerase was used to amplify the fragments which were cloned into the commercial vector pGEM-T, creating pIJ12338 and pIJ12339, respectively (Figure 7.8). In parallel, *pspA* was also targeted in pIJ12328, leaving ~8 kb homologous sequence upstream and downstream of *pspA*, creating pIJ12340. All three of these constructs were successfully conjugated into *P. alba*. Four exconjugants from each conjugation were streaked three times on ISP4 agar containing nalidixic acid and hygromycin. They were then grown in liquid culture and gDNA prepared from the resulting mycelium. These gDNA preparations were assessed by PCR using the primers 1289FlanA, which anneals to the left of *pspA*, and 1289RlanA, which anneals within *pspA*. Thus if the WT gene had been successfully deleted, no band would be observed. In parallel, another PCR was carried out using primers *pspA*confirmF and *pspA*confirmR that flank *pspA* in the WT *P. alba* genome. If the WT gene had been replaced by the hygromycin cassette, the amplified fragment would shift from 848 bp to 2059 bp. This shift in band size was observed in nine out of twelve clones, with the three remaining clones

showing both amplified bands, suggesting that only a single-crossover recombination event had occurred (Figure 7.9). Of the four exconjugants from the conjugation with pIJ12338, two had a double-crossover recombination event, entirely removing the WT *pspA* gene from the genome. Of the four exconjugants from the conjugation with pIJ12339, three were double-crossovers. Of the four exconjugants from the conjugation with pIJ12340, all were confirmed as true $\Delta pspA::hyg$ mutants. This perhaps reflects the greater likelihood of homologous recombination occurring when a larger section of homologous DNA is present.

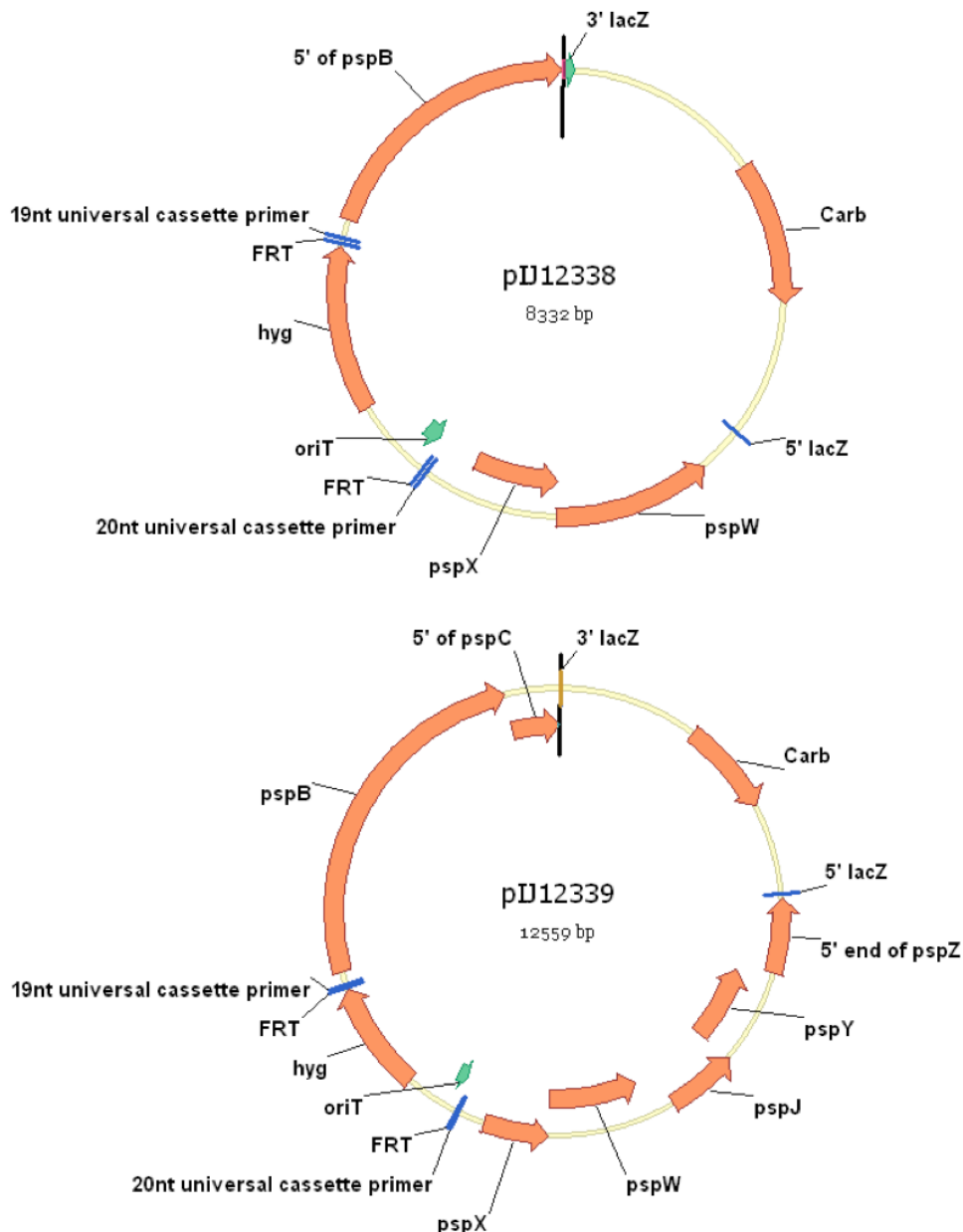


Figure 7.8 : Schematic of reduced constructs used to make a targeted deletion of *pspA*. pIJ12338 and pIJ12339 consist of the pGEM-T vector backbone ligated to the pIJ10700 cassette flanked by regions homologous to sequences either side of *pspA*. pIJ12338 has 2 kb of flanking homology, while pIJ12339 has 4 kb.

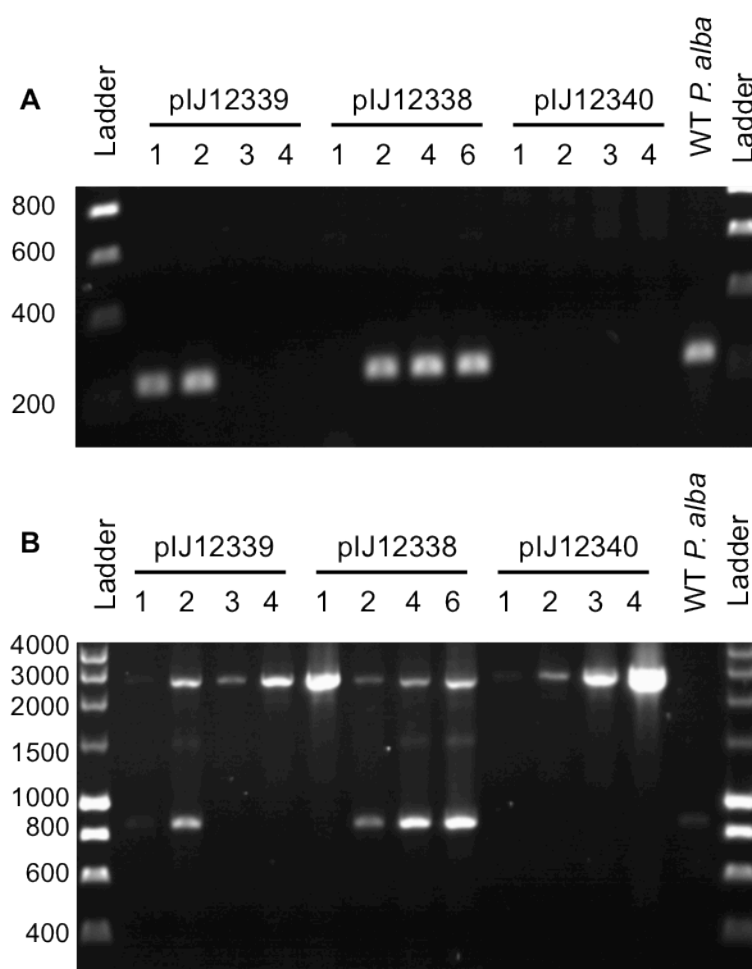


Figure 7.9 : PCR confirmation of the deletion of *pspA* from the *P. alba* chromosome.

This targeting was achieved through conjugation with pJ12338, pJ12339 or pJ12340. Genomic DNAs extracted from four *P. alba* exconjugants from each conjugation were used as templates to amplify *pspA* using two sets of primers. **A**; 1289FlanA and 1289RlanA anneal within *pspA* to amplify 230 bp if the WT gene is present. **B**; PspAconfirmF and PspAconfirmR anneal either side of the disrupted region to amplify 2720 bp $\Delta pspA::hyg$ or 848 bp WT *pspA*. Genomic DNA from WT *P. alba* was used as a control. PCR products were run on a 1 % agarose gel by electrophoresis. The ladder is Hyperladder I (Bioline) with band sizes annotated in bp.

7.3.1.2 Bioassay and MALDI-ToF analysis

One exconjugant corresponding to a double crossover event with each construct was tested for planosporicin production by bioassay and MALDI-ToF. All three exconjugants produced no antibiotic compounds when assayed against *M. luteus*. Likewise, no peptides with a mass corresponding to planosporicin were detected in MALDI-ToF (Figure 7.10). Thus the replacement of *pspA* with P1-FRT-*oriT*-*hyg*-FRT-P2 abolished production of planosporicin.

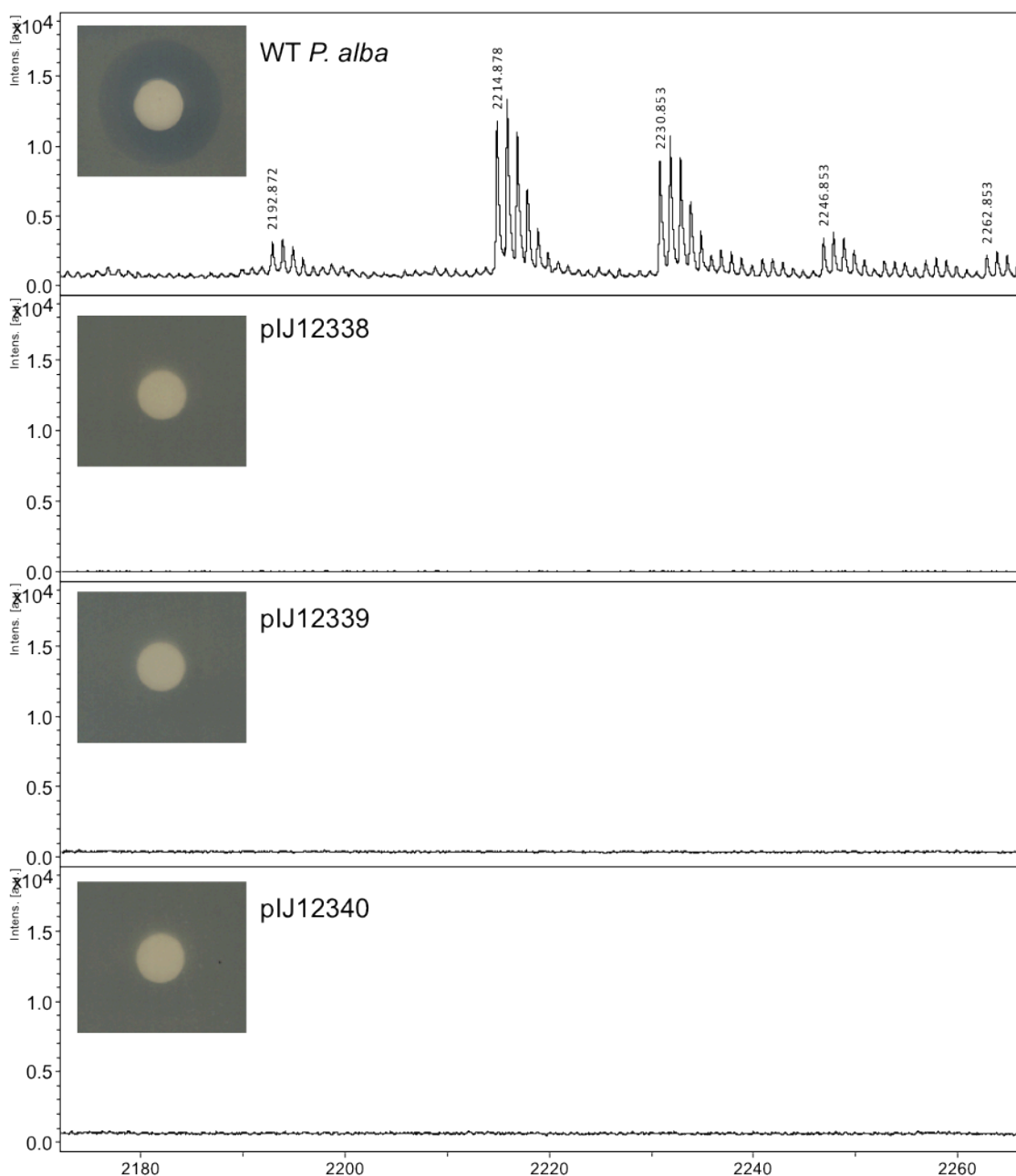


Figure 7.10 : Abolition of planosporicin production due to the targeted deletion of *pspA* in *P. alba*.

The deletion was made through conjugation with three different constructs; pIJ12338, pIJ12339 and pIJ12340. The three *P. alba* $\Delta pspA::hyg$ mutants and a WT *P. alba* control were cultured in AF/MS medium. Supernatant samples were taken after 5-8 days of growth and tested by bioassay and MALDI-ToF. 40 μ l of supernatant was applied to antibiotic assay discs which were laid onto a lawn of *M. luteus*. The plate was incubated for 36 hours at 30 °C before zones of inhibition were photographed (inset). The same supernatants were analysed by MALDI-ToF mass spectrometry. Intensity is given on the y-axis in arbitrary units (au) and the mass/charge ratio (*m/z*) on the x-axis. The monoisotopic mass of each *m/z* peak is labelled.

The strain constructed using pIJ12340, M1302, was used in all subsequent studies of the $\Delta pspA$ mutant. M1302 had a WT morphology when grown on agar and in liquid media. *P. alba* M1302 was grown for a two week period in AF/MS liquid medium and samples of culture supernatant taken and analysed for activity against *M. luteus*. No antibiotic activity was detected. Thus under these conditions any antibiotic compound produced by WT *P. alba* can be assumed to be planosporicin.

7.3.2 *pspYZ*

7.3.2.1 PCR analysis

pIJ12328 was PCR-targeted using the P1-FRT-*oriT*-*hyg*-FRT-P2 cassette from pIJ1700 to replace *pspYZ* using the method described in Section 7.3. This resulted in pIJ12328 $\Delta pspYZ::hyg$, subsequently named pIJ12523, which was then conjugated into heat shocked *P. alba* spores. Five of the putative Hyg^R exconjugants were streaked three times on agar plates containing 40 μ g/ml hygromycin. Since the backbone of pIJ12523 carries *neo* gene conferring kanamycin resistance, if pIJ12523 had been incorporated by single-crossover recombination, then the putative exconjugants would also be Kan^R; however if a double-crossover event had occurred, then *neo* as well as the WT *pspYZ* would have been removed, resulting in a Kan^S strain. Consequently, the Hyg^R exconjugants were tested for growth in the presence of 50 μ g/ml kanamycin. All five exconjugants had a Hyg^R Kan^R phenotype. The strains were then grown in liquid medium containing 40 μ g/ml hygromycin but lacking kanamycin to allow for a second recombination event that would remove *pspYZ* and the remainder of the pIJ12523 plasmid, leaving just the pIJ10700 hygromycin cassette. ISP4 agar containing hygromycin was spread with a sample of liquid culture. The resulting plates of single *P. alba* Hyg^R colonies were replica-printed consecutively onto ISP4 agar plates containing kanamycin and then hygromycin. Figure 7.11 demonstrates that after growth in liquid medium, 415/444 colonies (>93 %) were now the required Hyg^R Kan^S double crossovers. This method was used to isolate four mutants *P. alba* $\Delta pspYZ::hyg_7$, 11, 10 and 24. Using gDNA from each exconjugant, a fragment of expected size was amplified by PCR using one primer annealing within the pIJ10700 cassette and one annealing to the region flanking *pspYZ* (data not shown). The Hyg^R Kan^S *P. alba* $\Delta pspYZ::hyg_7$ mutant was assigned the strain number M1305, gDNA from which failed to amplify a product when both primers were within *pspYZ* (Figure 7.12).

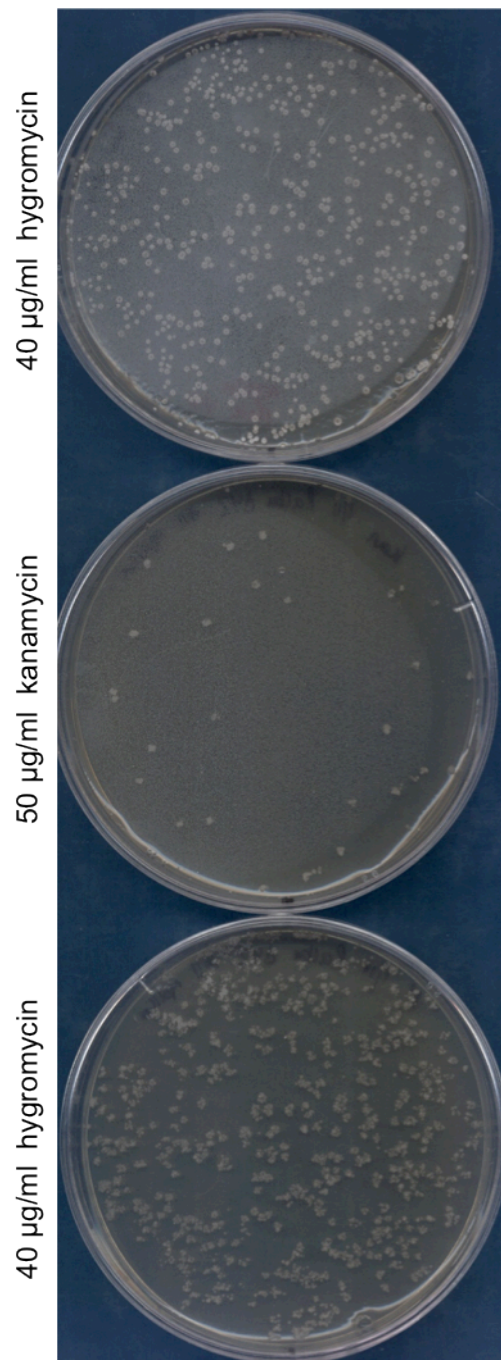


Figure 7.11 : Double crossover mutants isolated after growth in liquid culture.

A Hyg^R Kan^R *P. alba* Δ *pspYZ::hyg* exconjugant was grown for 7 days in liquid ISP4 with hygromycin selection. 50 µl of the culture was spread onto ISP4 containing 40 µg/ml hygromycin and incubated for 7 days at 30 °C (top). Individual sporulating colonies were observed and the whole plate was replica-printed onto ISP4 agar containing 50 µg/ml kanamycin (middle) then 40 µg/ml hygromycin (bottom) and incubated for a further 5 days at 30 °C before being photographed.

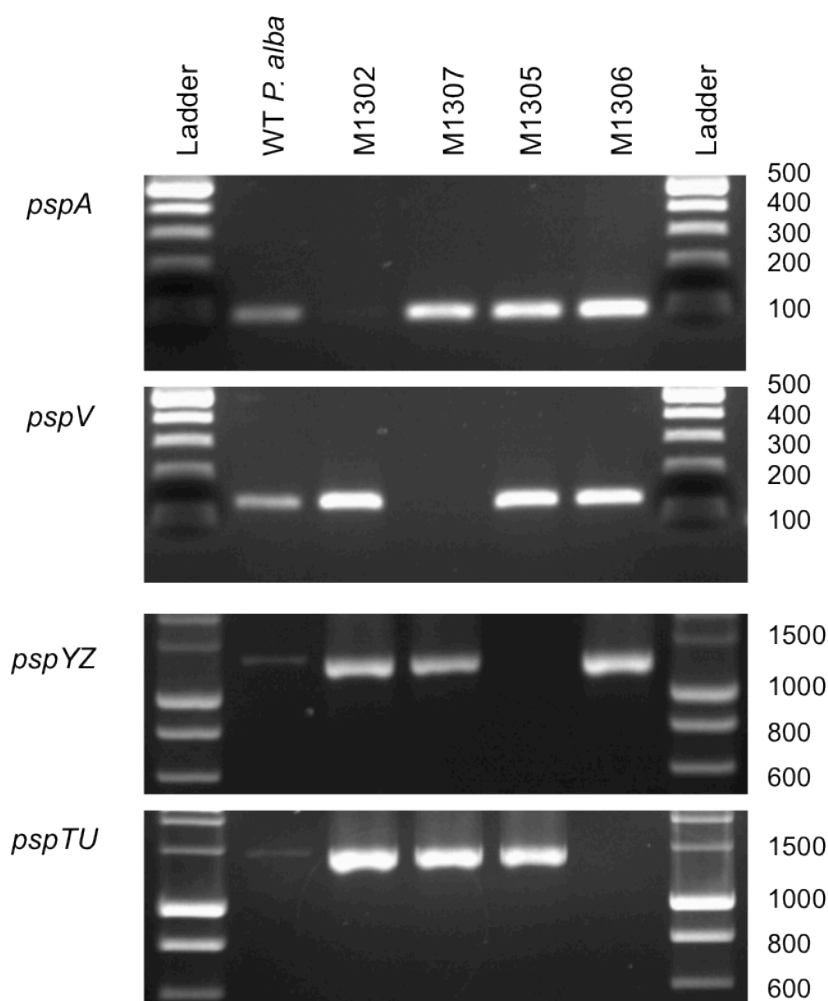


Figure 7.12 : PCR confirmation of four *P. alba* deletion mutants.

Namely M1302 ($\Delta pspA::hyg$), M1307 ($\Delta pspV::hyg$), M1305 ($\Delta pspYZ::hyg$) and M1306 ($\Delta pspTU::hyg$). Genomic DNAs extracted from each strain were used as a template to amplify *pspA*, (*pspAF2* and *pspAR2*; 101 bp), *pspV* (*pspVF* and *pspVR*; 137 bp), *pspYZ* (*pspYF* and *pspZR*; 1218 bp) and *pspTU* (*pspTF* and *pspUR*; 1431 bp) from the *psp* gene cluster. Genomic DNA from WT *P. alba* was used as a control. PCR products were run on a 2 % agarose gel by electrophoresis. The ladder is 100 bp DNA ladder (NEB) or Hyperladder I (Bioline) with band sizes annotated in bp.

7.3.2.2 Bioassay and MALDI-ToF analysis

All four *P. alba* $\Delta pspYZ::hyg$ exconjugants were tested for their ability to produce planosporicin by growing in liquid AF/MS medium. After 4 and 6 days growth, supernatant was sampled and used in a bioassay with *M. luteus* and sent for MALDI-ToF analysis. The WT control strain produced a bioactive compound that created a large zone of clearing due to the inhibition of the growth of *M. luteus*. *P. alba* $\Delta pspYZ::hyg$ clones 7, 10, 11 and 24 induced no such zone of clearing and no masses corresponding to

planosporicin were detected in MALDI-ToF. Figures 7.13 and 7.14 depict the bioassay and MALDI-ToF results for *P. alba* Δ *pspYZ::hyg_7* (assigned strain M1305) after six days growth in AF/MS. In subsequent repetitions of this experiment, a very small halo of *M. luteus* growth inhibition was observed in bioassays with supernatants from *P. alba* M1305. When these samples were sent for MALDI, a very small amount of the 2192 Da planosporicin peptide was detected (samples were spotted twice onto the MALDI plate to provide enough material for detection; data not shown). This observed low level of planosporicin production was more pronounced when using agar bioassay plates. M1305 grown for four days on AF/MS agar medium produced a small zone of inhibition, while the WT strain did not (Figure 7.15). However after longer periods of growth (11 days), the halo from M1305 had not markedly increased, while the WT strain had produced enough planosporicin to create a halo that partially overlapped that of the neighbouring M1305 patch (Figure 7.15). This phenotype was consistent for all four *P. alba* Δ *pspYZ::hyg* exconjugants, indicating that biosynthesis is not abolished in this mutant, and indeed occurs precociously when compared to the WT strain. Methanol extracts of M1305 mycelium failed to reveal an accumulation of active planosporicin in the mycelium through bioassay (Figure 7.16).

7.3.3 *pspTU*

7.3.3.1 PCR analysis

pIJ12321 was PCR-targeted using the P1-FRT-*oriT*-*hyg*-FRT-P2 cassette from pIJ1700 to replace *pspTU* using the method described in Section 7.3. This created pIJ12321 Δ *pspTU::hyg*, designated pIJ12524, which was then conjugated into heat shocked *P. alba* spores. Several hygromycin-resistant *P. alba* colonies were isolated, two of which were streaked out on agar medium containing 40 μ g/ml hygromycin. One of these exconjugants was found to be both Hyg^R and Kan^S. PCR on gDNA extracted from this exconjugant using one primer annealing within the pIJ10700 cassette (*pspTU*10700R) and one annealing in the region flanking *pspTU* (*pspTU*confirmF) amplified a fragment of the expected size. No fragment was amplified with primers annealing within *pspTU* (Figure 7.12). This strain was designated M1306.

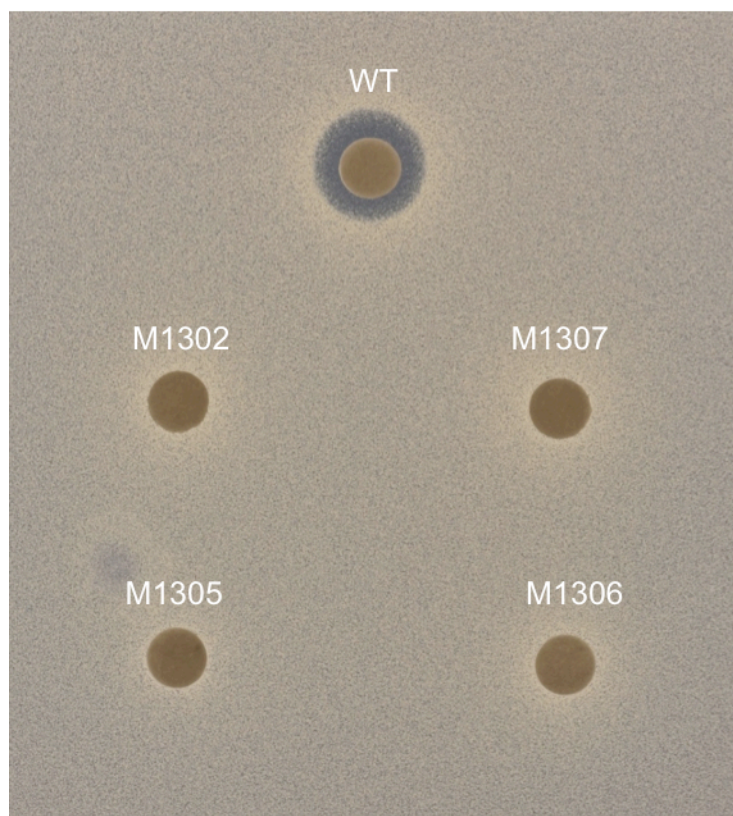


Figure 7.13 : Lack of production of antibiotic compounds from *P. alba* deletion mutants.

The four mutants assayed were M1302 ($\Delta pspA::hyg$), M1305 ($\Delta pspYZ::hyg$), M1306 ($\Delta pspTU::hyg$) and M1307 ($\Delta pspV::hyg$). The above *P. alba* strains were cultured in AF/MS medium for 6 days. A positive control of WT *P. alba* was cultured in parallel. Samples of culture supernatants were tested for antibiotic activity. 40 μ l of supernatant were applied to antibiotic assay discs which were laid onto a lawn of *M. luteus*. The plate was incubated for 36 hours at 30 °C before zones of inhibition were photographed.

7.3.3.2 Bioassay and MALDI-ToF analysis

The two *P. alba* M1306 ($\Delta pspTU::hyg$) clones were tested for their ability to produce planosporicin by bioassay against *M. luteus* and by MALDI-ToF analysis. The clones were grown alongside a WT control in liquid AF/MS medium and sampled after 4 and 6 days growth. Unlike the WT control, no antibiotic compound was detected from the supernatants of *P. alba* M1306 and no masses corresponding to planosporicin were detected in MALDI-ToF (Figure 7.13 and 7.14). This non-production phenotype was consistent when grown on AF/MS agar for 10 days before being overlaid with *M. luteus*,

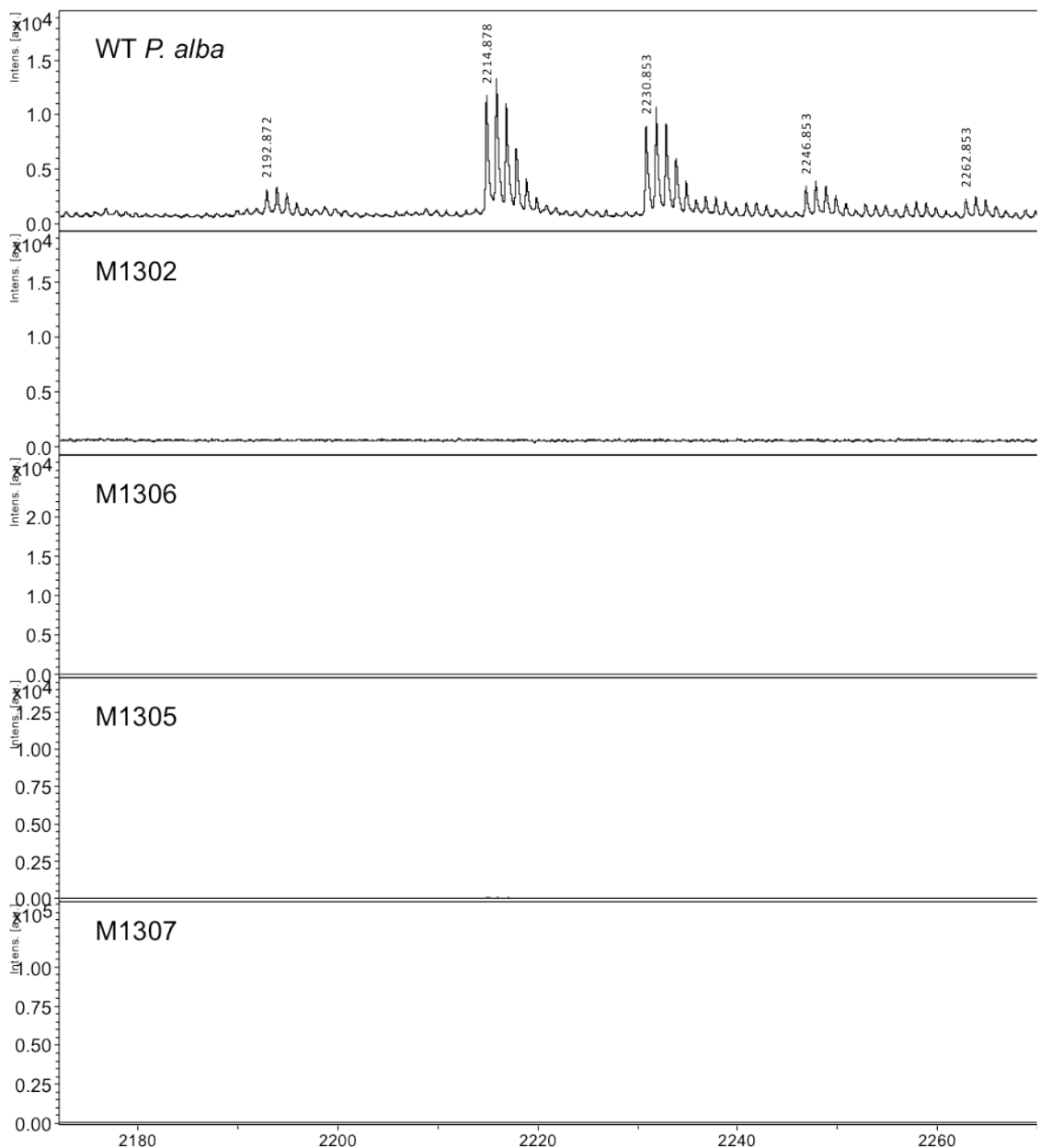


Figure 7.14 : Lack of production of peptides by four *P. alba* deletion mutants.

M1302 ($\Delta pspA::hyg$), M1305 ($\Delta pspYZ::hyg$), M1306 ($\Delta pspTU::hyg$) and M1307 ($\Delta pspV::hyg$) were cultured in AF/MS medium for 6 days. WT *P. alba* was grown in parallel as a positive control. Supernatant samples were taken after 6 days of growth and analysed by MALDI-ToF mass spectrometry. Intensity is given on the y-axis in arbitrary units (au) and the mass/charge ratio (m/z) on the x-axis. The monoisotopic mass of each m/z peak is labelled.

In other lantibiotic clusters, deletion of the ABC transporter responsible for lantibiotic export did not prevent biosynthesis, but led to accumulation of the lantibiotic in the cytoplasm. For example, deletion of *nisT* from the nisin gene cluster led to a build-up of processed nisin in the cytoplasm of *L. lactis* N8 (Qiao and Saris 1996). To determine if

deletion mutations which prevent planosporicin secretion into the supernatant are simply preventing export, mycelial extracts were tested for antimicrobial activity by bioassay. However, methanol extracts of *P. alba* M1306 mycelium did not reveal any antibiotic activity (Figure 7.16). MALDI-ToF analysis of these methanol extracts did not reveal the presence of inactive planosporicin with the leader peptide attached.

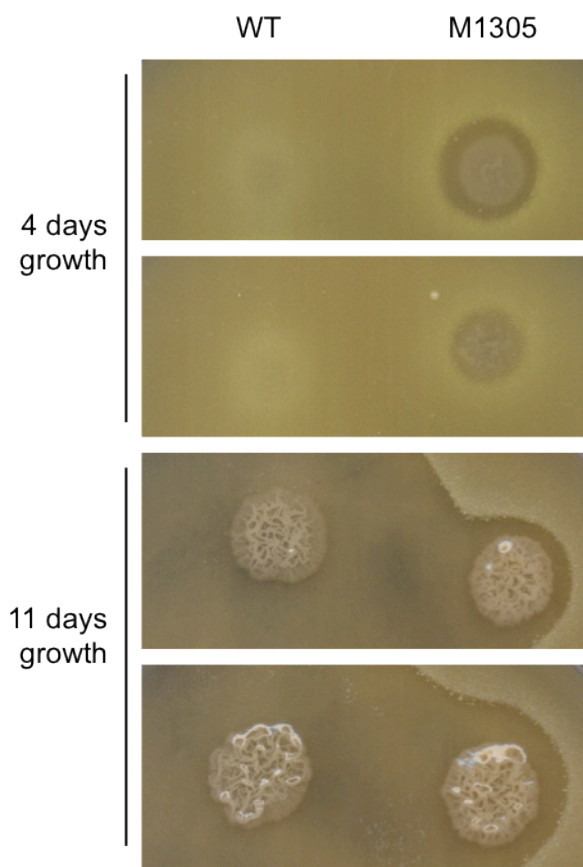


Figure 7.15 : Production of antibiotic compounds by *P. alba* M1305 in comparison to WT *P. alba* when grown on agar medium.

10 ul of the relevant seed culture was patched onto solid AF/MS plates in duplicate and incubated at 30 °C. After 4 days or 11 days, plates were overlaid with SNA containing *M. luteus*. The plates were incubated for 72 hours at room temperature before zones of inhibition were photographed.

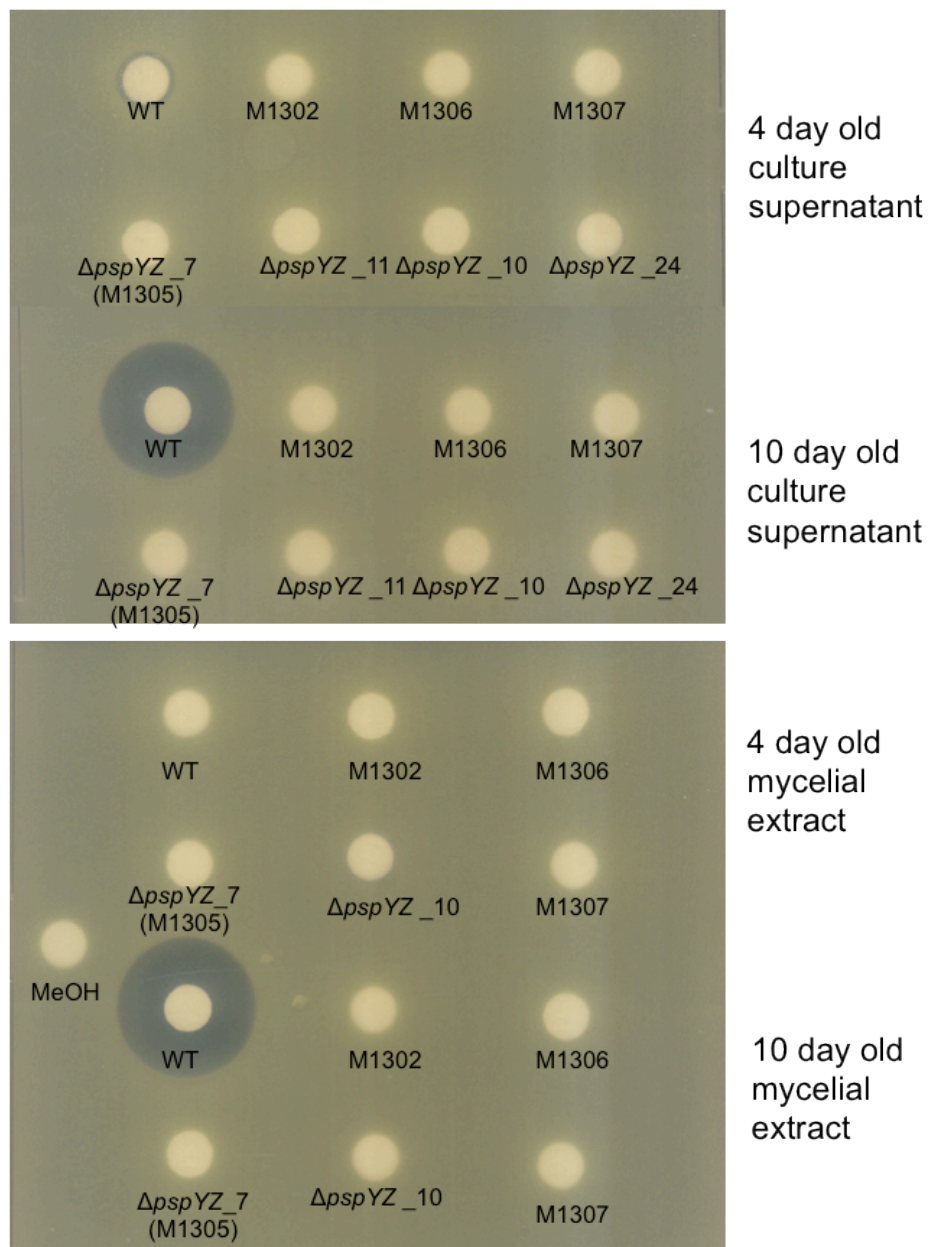


Figure 7.16 : Attempted detection of antibiotic compounds either secreted into the medium or accumulated in the cytoplasm.

P. alba deletion mutants M1306 ($\Delta pspTU::hyg$), M1307 ($\Delta pspV::hyg$) and M1305 ($\Delta pspYZ::hyg$) (as well as three other *P. alba* $\Delta pspYZ::hyg$ exconjugants) were cultured in AF/MS medium alongside M1302 ($\Delta pspA::hyg$) as a negative control and WT *P. alba* as a positive control. Each strain was harvested after 4 or 10 days growth at 30 °C, 250 rpm. Samples of culture supernatants (top panel) and methanol extracts from the mycelium (bottom panel) were tested for antibiotic activity. 40 μ l of each sample was applied to antibiotic assay discs and allowed to dry. A methanol only control was included to check for any inhibition caused by the solvent. Each disc was applied to a lawn of *M. luteus* which was incubated for 36 hours at 30 °C before zones of inhibition were photographed.

7.3.4 *pspV*

7.3.4.1 PCR analysis

In Chapter 4, pIJ12322 (cosmid F13-1) was sequenced and found to lack the 3' end of *pspV*. In Chapter 6, a conjugative derivative of this construct, pIJ12325, was integrated into the *Nonomuraea* chromosome but no planosporicin was detected. Thus in a heterologous host *pspV* appears to be necessary for planosporicin production. Therefore it was deemed interesting to determine if *pspV* was also necessary for planosporicin production in the natural producer. To this end, a number of constructs were made. As *pspV* is located at the right hand border of the *psp* gene cluster, the reduced versions of pIJ12321, i.e. pIJ12327, pIJ12328 and pIJ12329, do not contain any sequence downstream of *pspV* thus leaving no flanking sequence for homologous recombination. It was anticipated that smaller constructs would be more easily mobilised into *P. alba*. Consequently, pIJ12321, which contains 7.5 kb downstream of *pspV*, was targeted to remove all genes upstream of ORF -11 (Figure 7.17). The resistance cassette was subsequently removed by digestion with *Xba*I and re-ligation of the cosmid. This reduced construct was further PCR-targeted using the P1-FRT-*oriT*-*hyg*-FRT-P2 cassette from pIJ1700 to replace *pspV* using the method described in Section 7.3. This created pIJ12321 Δ orf-19 to orf-12, Δ *pspV*::*hyg*, designated pIJ12527, which was then conjugated into heat shocked *P. alba* spores. Although several *P. alba* colonies grew on the conjugation plates, none grew well when restreaked on agar medium containing 40 μ g/ml hygromycin. Consequently, an additional construct was made. Section 7.3.1.1 showed that recombination is possible in *P. alba* with just 2 kb of flanking homologous sequence. LongAmp polymerase was used to amplify Δ *pspV*::*hyg* from pIJ12527 along with >2 kb of adjacent sequence either side. The resulting 6.2 kb Δ *pspV*::*hyg* fragment, with *hyg* flanked by sequences homologous to the *psp* cluster, was cloned into pGEM-T, creating pIJ12530. Eventually, after several attempted conjugations with this construct, one putative exconjugant was isolated that grew well when streaked on agar medium containing 40 μ g/ml hygromycin and which was sensitive to kanamycin. gDNA isolated from a liquid culture of this clone was assessed by PCR using primers flanking *pspV*. This revealed a bandshift from the WT 1858 bp *pspV* to the 2125 bp Δ *pspV*::*hyg*. No amplification product was obtained with primers annealing within *pspV* (Figure 7.12). *P. alba* Δ *pspV*::*hyg* was designated M1307.

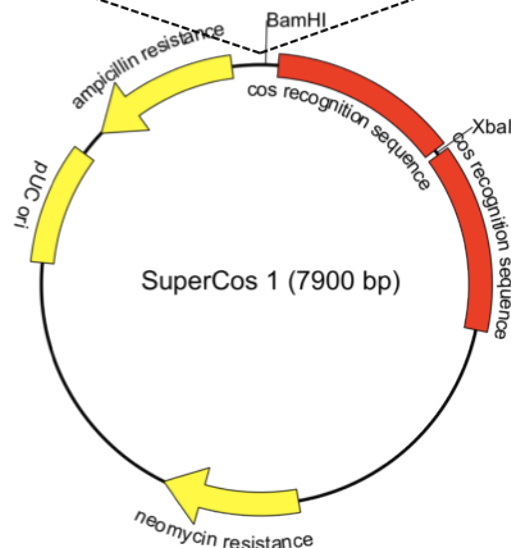
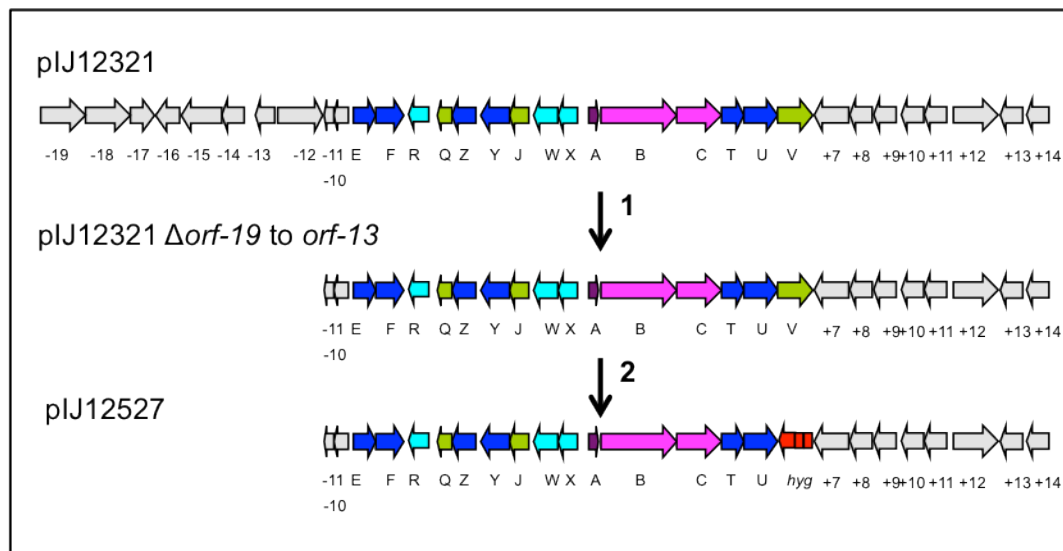


Figure 7.17 : Schematic illustrating the construction of a reduced construct to use to generate a *pspV* deletion mutant in *P. alba*.

In step 1, the insert DNA of cosmid pIJ12321 was PCR-targeted to remove *orf-19* to *orf-12*. In step 2, PCR targeting replaced *pspV* with P1-FRT-*oriT*-*hyg*-FRT-P2 cassette (in red) from pIJ1700.

7.3.4.2 Bioassay and MALDI-ToF analysis

The one *P. alba* Δ *pspV*::*hyg* exconjugant isolated was assayed for the ability to inhibit *M. luteus* and for the presence of peptides in the planosporicin mass range by MALDI-ToF analysis. *P. alba* M1307 grown on AF/MS agar for 10 days did not produce any compounds which inhibit *M. luteus* growth, unlike the large halo of inhibition around the

WT strain (data not shown). Two clones of M1307 were grown alongside a WT control in liquid AF/MS medium and sampled after 4, 5, 6, 7, 8 and 11 days of growth. Unlike, the WT control, no antibiotic activity was detected in the supernatants of *P. alba* M1307 and no masses corresponding to planosporicin were detected upon MALDI-ToF analysis (Figure 7.13 and 7.14). No antimicrobial activity was observed in methanol extracts of M1307 mycelium either (Figure 7.16).

If PspV functions as a scaffold constraining the prepeptide into a conformation for the binding of modification enzymes, these modifications may not occur in *P. alba* M1307. For every dehydration reaction, the mass of the lantibiotic precursor decreases by 18 Da (whereas subsequent cyclisation causes no mass change). Therefore, mass spectrometry can be used to distinguish between a precursor that is not dehydrated and one that is. For example mass spectroscopy on tryptic digests (as well as bioassay, SDS-PAGE and N-terminal sequencing) showed that nisin precursors from a strain lacking NisC activity, but having NisB activity, contained dehydrated residues but no (Me)Lan bridges (Koponen *et al.* 2002). Leader peptide removal occurs after completion of the post-translational modifications, so may not occur if (Me)Lan formation is prohibited. The theoretical mass of the planosporicin prepeptide is 5610.69 Da and the theoretical mass of the post-translationally modified prepeptide (i.e. with seven dehydrations) is 5484.616 Da. Thus MALDI-ToF was also performed in a high mass range on both supernatants and methanol extracts. However no peptides corresponding to the planosporicin prepeptide were detected. These results correspond to those obtained with *Nonomuraea* containing pIJ12325, i.e. containing the *psp* gene cluster but with a truncated *pspV* (Chapter 6).

7.3.5 *pspEF*

7.3.5.1 PCR analysis

pIJ12328 has only ~750 bp of sequence upstream of *pspE* which is unlikely to be sufficient for efficient recombination. In contrast, pIJ12327, lacking the genes downstream of *pspV*, has ~15 kb of DNA upstream of *pspE*. Consequently, pIJ12327 was PCR-targeted using the P1-FRT-oriT-hyg-FRT-P2 cassette from pIJ1700 to replace *pspEF* using the method described in Section 7.3. This created pIJ12321 Δ *pspEF::hyg*, designated pIJ12527, which was then conjugated into heat shocked *P. alba* spores. Although several *P. alba* colonies grew on the conjugation plates under hygromycin selection, none grew well when restreaked onto selective media. An additional construct was therefore made. Primers *pspEF_2kbF* and R2 were used to amplify a single 6.5 kb fragment consisting of Δ *pspEF::hyg* with >2 kb of flanking sequence from pIJ12527. This PCR product was cloned into pGEM-T, creating pIJ12529, which was subsequently

conjugated into *P. alba*. One exconjugant was isolated which grew well when streaked on hygromycin selective medium. gDNA was isolated from a liquid culture of the strain and PCR analysis revealed the presence of both *pspEF* and $\Delta pspEF::hyg$, indicating the occurrence of a single-crossover event. However, when patched first onto kanamycin selective plates then onto hygromycin selective plates it was found to be Hyg^R Kan^S. Thus it appears a deletion event had occurred that had removed the vector backbone but left both the WT *pspEF* and the disrupted $\Delta pspEF::hyg$ alleles. Despite further attempted conjugations, no exconjugants were obtained.

7.4 Generation of ‘scar’ mutants through FLP-mediated excision of the disruption cassette

The FRT sites flanking the ends of the disruptions cassettes are recognition targets for the FLP-recombinase. FLP-mediated recombination removes the central part of the disruption cassette, leaving an “in frame” (multiple of three) 81 bp ‘scar’ sequence whose reading frame lacks stop codons. FLP-mediated recombination enables the repeated use of the same resistance marker to make multiple gene replacements in the same cosmid or strain. In this work, FLP-mediated excision of the disruption cassette was desirable in order to generate non-polar, unmarked in-frame deletions in which expression of downstream genes was unaffected.

The method involves the use of *E. coli* DH5 α containing BT340, the temperature-sensitive FLP recombination plasmid. BT340 encodes the FLP recombinase and is temperature sensitive for replication (replicates at 30 °C), so FLP synthesis and loss of the plasmid are induced at 42 °C (Cherepanov and Wackernagel 1995). Expression of the FLP-recombinase in *E. coli* removes the central part of the disruption cassette in the cosmid, leaving behind the ‘scar’ sequence (as described in Chapter 2) (Datsenko and Wanner 2000). The scarred cosmid can then be used to replace resistance cassette inserts in the *P. alba* chromosome with the unmarked ‘scar’ sequence through homologous recombination.

7.4.1 *pspA*

As described in Section 7.3, *P. alba* M1302 was constructed by replacing *pspA* with the P1-FRT-oriT-hyg-FRT-P2 resistance cassette. This generated a replacement mutant marked with the hygromycin resistance cassette. It seemed likely that this mutation would have polar effects on the expression of *pspB*, *pspC*, *pspT*, *pspU* and *pspV* located

downstream of and co-transcribed with *pspA*. The original replacement primers were carefully designed to end in the ATG start and TGA stop codons of *pspA*, thus retaining the same reading frame after gene replacement. Thus the hygromycin cassette could be removed by FLP-recombination to allow in frame transcriptional read-through into downstream genes. It was also envisaged that this scar mutant could be complemented not only with *pspA* but also with other *lanA* genes and perhaps also fusion constructs using the leader peptide of *pspA* fused to the propeptide gene of other lantibiotics.

There were two possible methods to generate the required *P. alba* Δ *pspA*::scar mutant. Both used the canonical method described in Chapter 2 to excise the hygromycin resistance cassette and *oriT* from between the FRT sites in pIJ12340 (pIJ12328 Δ *pspA*::(*oriT*-*hyg*)). For both approaches, pIJ12340 was first introduced into *E. coli* DH5 α /BT340 by transformation and FLP-mediated recombination induced as described in Chapter 2 yielding Hyg^S Kan^R clones. Removal of the hygromycin resistance cassette was confirmed by high-fidelity PCR using flanking primers (PspAconfirmF and PspAconfirmR) to generate PCR products that were then confirmed by Sanger sequencing. The resulting plasmid, pIJ12328 Δ *pspA*::(scar), was subjected to restriction analysis to ensure that no other gross deletions or rearrangements had occurred in the cosmid during FLP-mediated recombination. The methods differ in the manner in which the *oriT* is re-introduced to the pIJ12328 Δ *pspA*::(scar) plasmid. Usually, the construct would be re-introduced into *E. coli* BW25113/pIJ790 allowing the *bla* gene to be replaced with *oriT* and a selectable resistance marker not already present in the plasmid. However, as the *hyg* marker is already present in pIJ12340 (pIJ12328 Δ *pspA*::(*oriT*-*hyg*)) and selection for apramycin in conjugations is unreliable, an alternative method was used. This method was recently used to scar the *garA* gene of the actagardine cluster by introducing the scar cosmid into WT *A. garbadiensis* (Boakes *et al.* 2009). This work described how single-crossover exconjugants were obtained and subcultured through six successive rounds of growth in liquid media without selection to allow a second recombination event to occur. Using this method, it would be possible to introduce the scar cosmid into WT *P. alba* using hygromycin to select for single crossover events and then grow non-selectively to allow double crossovers. To this end, the *bla-oriT-hyg-bla* cassette from pIJ10701 was amplified using blaF and R primers and the 1535 bp fragment was isolated from an agarose gel. This fragment was used to introduce *oriT* and the hygromycin resistance marker into the backbone of pIJ12328 Δ *pspA*::(scar) in place of *bla* to create pIJ12531. This cosmid was checked by PCR and restriction analysis before transformation into the donor strain *E. coli* ET12567/pUZ8002 for mobilisation into WT *P. alba* via conjugal transfer.

Two hygromycin-resistant *P. alba* colonies were isolated from the conjugation plates and streaked three times on ISP4 containing 25 μ g/ml nalidixic acid and 40 μ g/ml hygromycin.

Two clones from each exconjugant were inoculated into liquid ISP4 media without antibiotic selection and grown for 7 days before the culture was inoculated 1 in 10 into a second liquid culture. This method was repeated between six and eight times for each of the four clones before a sample of culture was plated onto solid ISP4. After 5 days growth, colonies from this plate were patched onto ISP4 containing 40 µg/ml hygromycin and ISP4 without selection. Cells sensitive to hygromycin were patched onto plates of D/seed media to generate mycelia for colony PCR. Out of 96 colonies patched, 30 were sensitive to hygromycin, of which 21 were checked by colony PCR using primers *pspAconfirmF* and *pspAconfirmR*; unfortunately in all 21 cases the primers amplified the 848 bp WT *pspA* and not the 719 bp scar sequence. PCR on gDNA prepared from four colonies confirmed this result. As an alternative assay method, nine of the hygromycin sensitive clones were grown in liquid AF/MS medium to screen for lack of antibiotic production. However all clones produced planosporicin. Therefore, although this method was reported to be successful in generating scar mutants in *A. garbadiensis*, it was unsuccessful in *P. alba*. All hygromycin sensitive colonies had reverted to WT after integration of the scarred cosmid.

On reflection, these results are not particularly surprising. Growth experiments show that young *P. alba* mycelium is susceptible to planosporicin. Immunity occurs either through the expression of immunity genes or through age, as mature colonies no longer elongate through growing a new cell wall, so planosporicin is no longer effective. The double crossover event which leaves a scar in place of *pspA* likely occurs at a similar frequency to the excision of the entire construct. The removal of *pspA* would abolish production of planosporicin and thus prevent the self-induction of planosporicin immunity genes. Any mycelial compartments that instead reverted to WT would produce planosporicin which could kill any actively growing *pspA* scar mutants. Although planosporicin production in ISP4 is not routinely observed, a halo of inhibition is sometimes seen when *P. alba* is grown in this medium, so production is possible. In an attempt to avoid this, hygromycin resistant single crossovers that had been streaked three times on ISP4 containing 25 µg/ml nalidixic acid and 40 µg/ml hygromycin were then streaked a further three times on ISP4 lacking any selection. Each time, well separated single colonies were selected so that any colonies which had reverted to WT would not be close enough to any scarred mutants to kill them. Forty colonies were patched onto ISP4 containing 40 µg/ml hygromycin then ISP4 lacking selection. Fifteen were hygromycin sensitive and patched onto D/seed agar for growth without selection. Once again, colony PCR revealed that all of the hygromycin sensitive colonies had reverted to WT instead of generating a scar mutant. Time limitations prevented alternative methods being pursued.

7.5 Complementations

7.5.1 Complementing biosynthetic mutants

All *psp* mutants were made through the targeted replacement of the gene of interest with the hygromycin resistance cassette. Attempts to make scar mutants failed as described in Section 7.4, so instead, the marked mutants were used. Attempts at complementation were performed *in trans* with the WT gene. This required the use of a second resistance cassette. Section 7.2.3 described how apramycin-resistant *P. alba* exconjugants containing pSET152 integrated into the chromosome were obtained if the conjugation was overlaid with 20 µg/ml nalidixic acid and just 20 µg/ml apramycin. Thus all complementation constructs made during this work were based on the vector pSET152.

7.5.1.1 *pspA*

To complement deletion of the gene encoding the planosporicin precursor peptide, *pspA* was introduced in *trans* with expression from the native promoter P_{pspA} (defined as the region between *pspX* and *pspA*) through integration of pIJ12532. Chapter 2 describes the creation of pIJ12532 through the cloning of P_{pspA} together with *pspA* into pSET152. pIJ12532 was conjugated into *P. alba* M1302 and exconjugants were selected with 20 µg/ml apramycin.

Three exconjugants were streaked out three times on ISP4 media containing 50 µg/ml apramycin and were subsequently inoculated into liquid culture from which gDNA was prepared. The gDNA preparations were assessed by PCR using primers which anneal within pIJ12532 (pSET152F2 and pSET152R2) to amplify the 727 bp band corresponding to *pspA* integrated into the *P. alba* chromosome. Two representative clones of each exconjugant were tested for planosporicin production along with WT *P. alba* as a positive control. Supernatants collected from 5, 6 and 7 day old AF/MS liquid cultures were tested for the ability to inhibit growth of *M. luteus*. While supernatants from 5 day cultures of WT *P. alba* conferred visible halos, those from *P. alba* M1302 carrying the complementation construct pIJ12532 did not (Figure 7.18). The absence of planosporicin-related peaks was confirmed by MALDI-ToF analysis (Figure 7.19). Furthermore, agar-grown patches of *P. alba* M1302 carrying pIJ12532 also failed to create a halo of *M. luteus* inhibition. The genes *pspABCTUV* are likely to constitute one operon, with expression driven from the *pspA* promoter. A polar effect of the *pspA* deletion marked with the hygromycin cassette on the expression of downstream genes is the most likely explanation for the lack of complementation. The proteins encoded by *pspBC* are likely to be involved in (Me)Lan bridge formation so are very likely to be essential. Furthermore, Section 7.3.3 and Section

7.3.4 revealed that *pspTU* and *pspV* are both essential for planosporicin biosynthesis. One solution would be to complement the *pspA* mutation with *pspABCTUV*. However this fragment is approximately 8 kb, making the incorporation of mutations during PCR amplification highly likely. The absence of conveniently positioned restriction sites prevented the excision of the *pspABCTUV* fragment. However, even this approach may not have been successful. Recent work demonstrated that the deletion of *mibA* from the *mib* gene cluster in *M. corallina* could not be complemented. In the *mib* cluster, *mibABCDTUV* is likely to be transcribed as one operon. Deletion analysis revealed that *mibTU* and *mibV* were not essential for microbisporicin production. Thus attempts were made to complement the targeted *M. corallina* $\Delta mibA::aac(3)IV$ mutant with *mibA* and *mibABCD* under control of the native P_{mibA} promoter. However neither construct restored microbisporicin production (Foulston 2010).

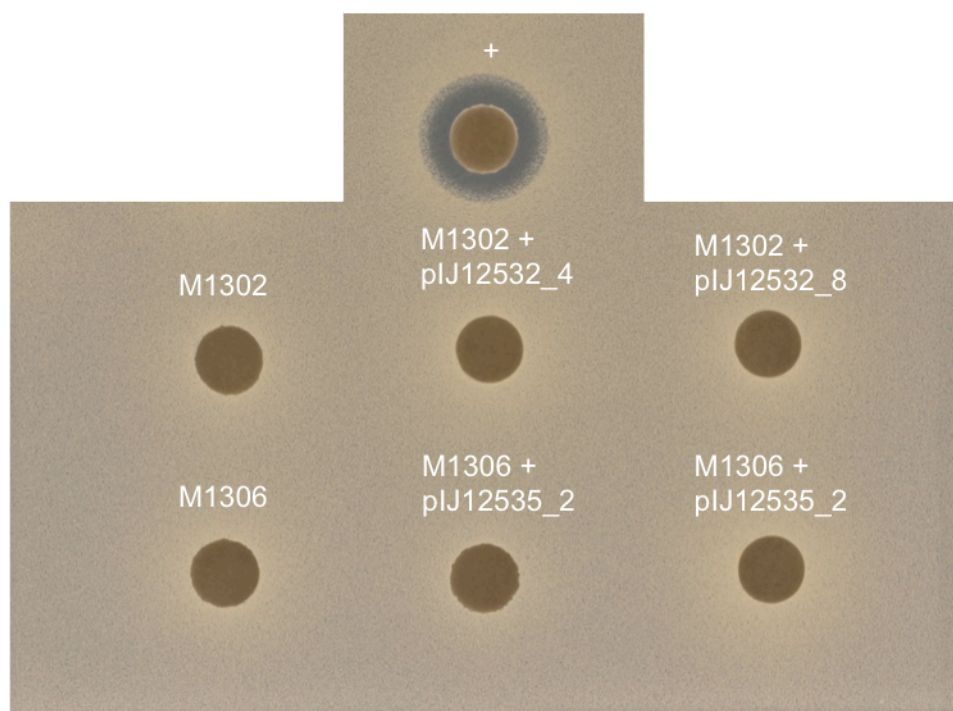


Figure 7.18 : Lack of production of antibiotic compounds from *P. alba* deletion mutants complemented in *trans* with the WT gene.

P. alba M1302 ($\Delta pspA::hyg$) with integrated pIJ12532 (P_{pspA} -*pspA*) and *P. alba* M1306 ($\Delta pspTU::hyg$) with integrated pIJ12535 (P_{pspA} -*pspTUV*) were cultured in AF/MS medium for 5 days. Samples of culture supernatants were tested for antibiotic activity. 40 μ l of supernatant were applied to antibiotic assay discs which were laid onto a lawn of *M. luteus*. A sample containing planosporicin (+) was used as a positive control in the assay. The plate was incubated for 36 hours at 30 °C before zones of inhibition were photographed.

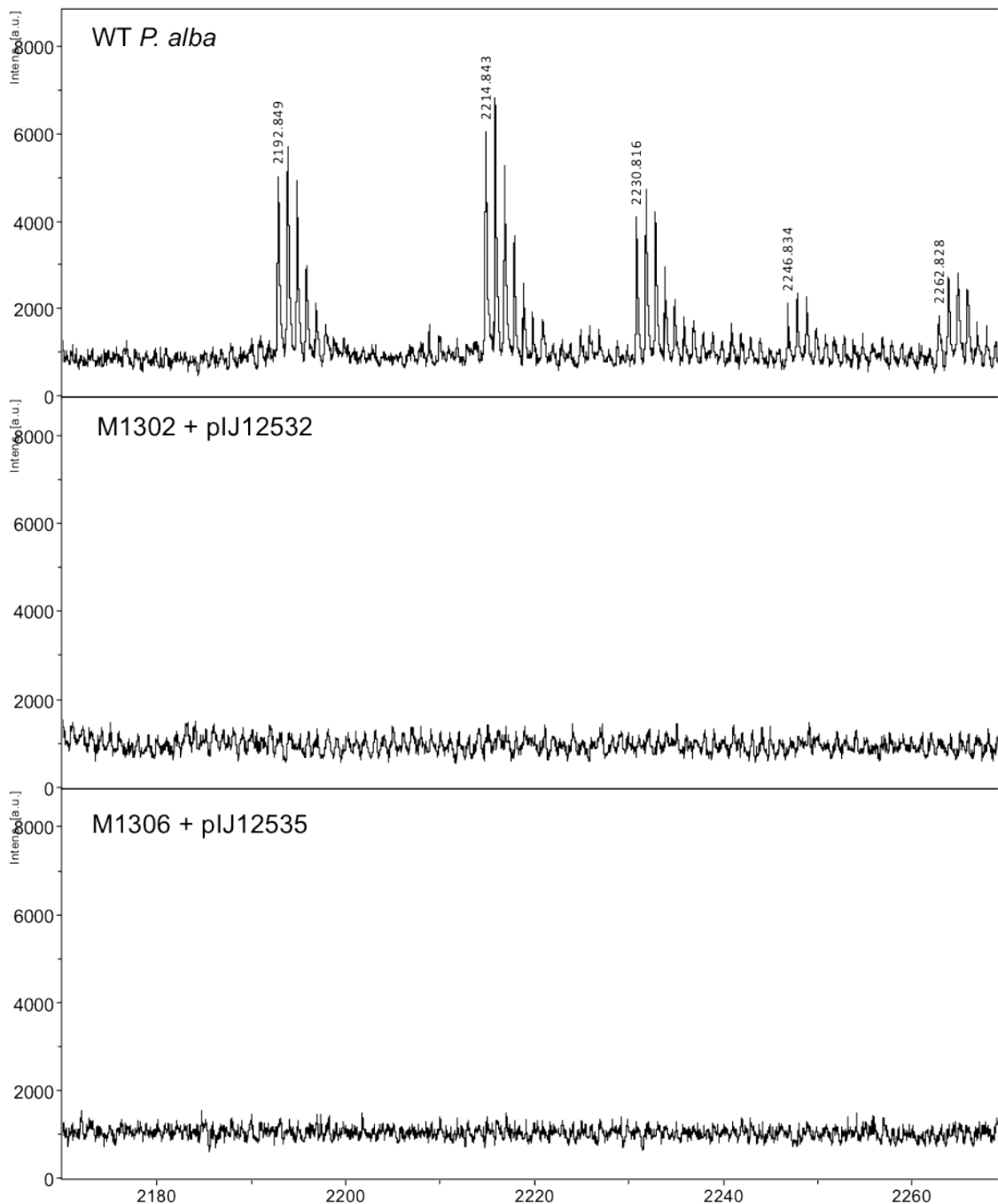


Figure 7.19 : Lack of production of peptides from *P. alba* deletion mutants complemented in *trans* with the WT gene.

P. alba M1302 ($\Delta pspA::hyg$) with integrated pIJ12532 ($P_{pspA}-pspA$) and *P. alba* M1306 ($\Delta pspTUV::hyg$) with integrated pIJ12535 ($P_{pspA}-pspTUV$) were both cultured in AF/MS medium. WT *P. alba* was grown in parallel as a positive control. Supernatant samples were taken after 5 days of growth and analysed by MALDI-ToF mass spectrometry. Intensity is given on the y-axis in arbitrary units (au) and the mass/charge ratio (m/z) on the x-axis. The monoisotopic mass of each m/z peak is labelled.

7.5.1.2 *pspTU*

To complement deletion of the ABC transporter genes *pspTU*, several constructs were introduced *in trans* with expression from the native promoter. *pspABCTUV* are likely to constitute one operon, with expression driven from the *pspA* promoter, P_{*pspA*}. Chapter 2 details the creation of pIJ12533 through the cloning of P_{*pspA*} into pSET152. *pspTU* was amplified using the primers pspTUF_SpeI and pspTUR_XbaI, while *pspTUV* was amplified using the primers pspTUF_SpeI and pspTUVR_XbaI, in both cases using a high-fidelity polymerase. The amplified fragments were cloned into pGEM-T. Four colonies from each ligation were subjected to high fidelity PCR with the primers pSET152F and pSET152R. The amplified fragments were subjected to Sanger sequencing to identify clones in which no mutations had occurred. *pspTU* and *pspTUV* were liberated from pGEM-T using a double digest with XbaI and SpeI. The 2066 bp *pspTU* and 3345 bp *pspTUV* fragments were isolated from an agarose gel and ligated between the SpeI and XbaI sites of pIJ12533 to create pIJ12534 and pIJ12535, respectively. All three complementation constructs were sequenced before transformation into *E. coli* ET12567/pUZ8002. Conjugation of these donor strains with *P. alba* M1306 was attempted several times, during which the conjugation plates were overlaid with 20 µg/ml apramycin.

Although several *P. alba* colonies grew through the overlay, very few sporulated, many grew poorly and did not grow when streaked onto ISP4 containing a higher concentration of apramycin. These colonies were likely to be the recipient *P. alba* strain growing through the low level of apramycin selection. However, one of the three conjugations was successful. One exconjugant was obtained through conjugation of *P. alba* M1306 with *E. coli* ET12567/pUZ8002 harboring pIJ12535 containing the cloned *pspTUV* genes. This exconjugant was streaked three times on ISP4 media containing 50 µg/ml apramycin and 25 µg/ml nalidixic acid before inoculation into liquid culture to grow mycelium from which gDNA was prepared. This gDNA was assessed by PCR using primers within pIJ12535 to demonstrate that the construct had integrated successfully into the *P. alba* M1306 chromosome. Two clones from this exconjugant were patched onto plates of AF/MS agar but no halo of inhibition was observed when overlaid with *M. luteus* after 5, 6 or 7 days of growth. Likewise these clones were cultured in AF/MS liquid medium, but no antibiotic production was detected when 5, 6, 8 or 13 day old supernatant samples were assayed against *M. luteus* (Figure 7.18). These samples were subsequently sent for MALDI-ToF analysis and the absence of planosporicin-associated peaks confirmed (Figure 7.19).

7.5.2 Introduction of a different *lanA* into WT *P. alba*

Section 7.4 described how attempts to make a scarred version of *P. alba* Δ *pspA*::*hyg* was not successful. It was also not possible to complement the marked mutant *P. alba* M1302 in *trans* with the WT *pspA* under control of the P_{pspA} . If either of these methods were successful, the logical extension would be to complement the relevant Δ *pspA* mutant with a structural peptide from a different lantibiotic cluster. The close similarity between the planosporicin and microbisporicin gene clusters (described in Chapter 5) made the structural gene from the *mib* cluster, *mibA*, the obvious choice. Microbisporicin is produced from a single gene cluster, containing one prepropeptide gene, *mibA* (Foulston and Bibb 2010). In addition to the introduction of (Me)Lan bridges this prepropeptide is subject to a number of unusual post-translational modifications. Microbisporicin refers to four lantibiotic-like compounds which are 107891 A1 (2246.71±0.06 Da), 107891 A2 (2230.71±0.06 Da) (Lazzarini *et al.* 2005; Castiglione *et al.* 2008), MF-BA-1768 α 1 (2214.880 Da) and MF-BA-1768 β 1 (2180.805 Da) (Lee 2003). The mass difference between these compounds corresponds to the level of modification of the prepropeptide. 107891 A1 is the fully modified peptide containing a chlorinated tryptophan and dihydroxyproline. 107891 A2 is 16 Da smaller than 107891 A1, due to the loss of one oxygen atom from Pro₁₄. The loss of the second oxygen atom creates the non-hydroxylated form of microbisporicin, MF-BA-1768 α 1. While MF-BA-1768 β 1 is 34 Da smaller than MF-BA-1768 α 1 due to the loss of the chlorine atom on Trp₄, creating the non-hydroxylated and non-chlorinated form of microbisporicin.

Constructs pIJ12138 and pIJ12362 were kindly provided by L. Foulston. These plasmids are based on pIJ10706 in which *mibA* or *mibABCD* is under control of the native P_{mibA} promoter. These constructs were mobilised by conjugal transfer into WT *P. alba* where integration into the chromosome was selected with hygromycin. Six exconjugants from each conjugation grew well when streaked three times on ISP4 containing 25 µg/ml nalidixic acid and 40 µg/ml hygromycin. Four exconjugants were moved into liquid culture and the resulting mycelium used to prepare gDNA for PCR analysis. Amplification of the 115 bp *mibA* (primers LF025F and LF025R) from *P. alba* containing pIJ12138 confirmed integration of the construct into the chromosome. Likewise, amplification of both the 115 bp *mibA* and 159 bp *mibD* (primers LF026F and LF026R) from *P. alba* containing pIJ12362 confirmed the integration of this construct. Four exconjugants from each conjugation were grown in liquid AF/MS and MV media alongside a *P. alba* WT control. In MV medium, no antibiotic compounds were produced (as assessed by failure of culture supernatants to inhibit a lawn of *M. luteus*), even by the WT control. In AF/MS medium, the WT control and two of the four *P. alba* pIJ12138 exconjugants produced an antibiotic compound. The other two exconjugants also failed to produce a bioactive compound in a

repeat assay. Of the four *P. alba* pIJ12362 exconjugants, only one produced a bioactive compound in AF/MS medium.

In the absence of the *mib* genes responsible for the hydroxylation and chlorination of microbisporicin, it was expected that only the MF-BA-1768 β_1 form of microbisporicin would be visible in the spectrum. This has an expected m/z peak for [M+H⁺] of 2181.805 Da, a peak for [M+Na⁺] of 2203.805 Da and a peak for [M+K⁺] of 2219.805 Da. These masses are similar to the mass of planosporicin. In culture supernatants from the native producer *M. corallina*, the non-hydroxylated and non-chlorinated form of microbisporicin is not usually detected in WT supernatant (Foulston 2010). The presence of chlorine in the other three forms of microbisporicin alters the isotope signature such that peak abundance is 3rd>2nd>1st. However the anticipated absence of chlorination in microbisporicin produced by *P. alba* means that peak abundance would be 2nd>1st>3rd, the same as for planosporicin associated peaks. In all MALDI-ToF spectra, the 2192 Da, 2214 Da and 2230 Da peaks of planosporicin were present while the 2181 Da, 2203 Da and 2219 Da peaks of non-hydroxylated, non-chlorinated microbisporicin were absent. As anticipated, peaks corresponding to the hydroxylated and chlorinated forms of microbisporicin were also absent. Thus it was not possible to produce microbisporicin in *P. alba* through the heterologous expression of *mibA* or *mibABCD*. There are many possible explanations. Perhaps the promoter P_{mibA} cannot be recognised in *P. alba* so is not expressed. Even if the construct is expressed, perhaps the modification enzymes encoded by the *psp* gene cluster do not share sufficient sequence similarity with the Mib enzymes to act on both prepropeptides.

7.6 Discussion

The successful targeted deletion of planosporicin biosynthesis genes demonstrates that the intrinsic homologous recombination efficiency of *P. alba* is sufficient to use PCR targeting, which was developed based on the recombination efficiency of *Streptomyces*. Furthermore the direct isolation of double-crossover mutants demonstrated that *P. alba* is naturally quite recombinogenic.

The deletion of *pspA* proved its essentiality in planosporicin production. Likewise the deletion of the *pspTU* ABC transporter genes also abolished planosporicin production. This was unexpected as the deletion of the equivalent *mibTU* genes from *M. corallina* did not stop microbisporicin production (Foulston and Bibb 2010). However, as *P. alba* M1306 ($\Delta pspsTU::hyg$) could not be complemented with *pspTUV*, there may be a second site mutation which prevents production in this particular mutant. Deletion of the *pspYZ* ABC transporter genes severely reduced the amount of planosporicin secreted out of

mycelium. In liquid culture, often no production is observed, whereas a small halo is consistently observed in solid bioassays. There is no accumulation of planosporicin in the mycelium, suggesting this transporter is not responsible for export of the mature lantibiotic. The deletion of *pspV* indicates that it too is essential for planosporicin production. It is possible that PspV may act as a scaffold or chaperone for the prepropeptide to protect it from degradation and present the substrate to modification enzymes. Yet there was no indication of the planosporicin prepropeptide or smaller degradation products in wide-range MALDI-ToF of *P. alba* M1307 supernatants. In *M. corallina*, MibV has been implicated in the chlorination of microbisporicin. However planosporicin has a valine at this position and no chlorinated residues elsewhere on the peptide. These four deletion mutants of *psp* genes proposed to encode proteins with a biosynthetic function all gave a phenotype of greatly reduced or totally abolished planosporicin secretion. Extraction of mycelium with methanol did not reveal a significant accumulation of planosporicin in the mycelium of any of these mutants.

The inability to delete *pspEF* could indicate that they are essential for *P. alba* growth. Bioinformatic analysis from Chapter 5 indicates that *pspEF* may provide an immunity mechanism. However the *mibEF* genes of the microbisporicin cluster were successfully deleted from *M. corallina* and this Δ *mibEF* mutant was complemented by the *in trans* expression of *mibEF* (Foulston and Bibb 2010). The Δ *mibEF* strain showed severely reduced and delayed microbisporicin biosynthesis. Furthermore, there was also severe growth inhibition when exposed to WT supernatant containing microbisporicin. Thus it would appear that *mibEF* encode an immunity mechanism and that when disrupted, a negative feedback loop acts to reduce production (Foulston and Bibb 2010).

Disappointingly, a number of *P. alba* deletion mutants could not be complemented through *in trans* expression of the WT gene. Although PCR analysis indicated that the rest of the gene cluster was intact, Southern blots would be needed to be certain that no rearrangements had occurred in the *P. alba* chromosome during the construction of the deletion mutants. However, even a Southern blot cannot rule out the presence of a second site mutation introduced when the gene was deleted. This may be more likely in *P. alba* mutants generated through the introduction of ~2 kb PCR-generated fragments on both sides of a resistance cassette with subsequent screening for double crossover recombination events. As the PCR generated fragments were not sequenced, the 2 kb flanking fragments might have contained several mutations, which, depending on the sites of the crossovers, could have introduced additional mutations upstream (and/or downstream) of the targeted knockouts. This method was used to generate the first *P. alba* Δ *pspA::hyg* knockout mutant and the only *P. alba* Δ *pspV::hyg* mutant, M1307. However another *P. alba* Δ *pspA::hyg* mutant was constructed using a cosmid-based

construct and this strain (M1302) also failed to demonstrate complementation. Time restrictions prevented attempts to complement the *P. alba* Δ *pspV::hyg* mutant (M1307) in this work.

The lack of complementation is unlikely to be due to a polar effect on downstream genes, as the complementation constructs were carefully designed to include the deleted gene and all genes downstream in the same operon and sequenced to ensure no mutations were introduced during construction. However there may be other problems with the complementation constructs. There may be a lack of activity in the *pspA* and *pspJ* promoter regions which prevents transcription. One solution would be to use a constitutive promoter such as the promoter region of the erythromycin resistance gene (*ermE*) from *S. erythraea*, (Bibb *et al.* 1985). The upregulated variant of this promoter, *ermE**p, has been used frequently as a strong constitutive promoter for native and heterologous genes in several actinomycetes. For example, the triketide lactone synthase, DEBS1- TE, was cloned with the strong heterologous *ermE** promoter in place of the promoter for the activator gene. This directed highly efficient production of the PKS protein in *Streptomyces cinnamonensis* (Wilkinson *et al.* 2002). A recent example of a lantibiotic gene expressed from *ermE** comes from the complementation of *S. coelicolor* M1146 derivatives containing a version of *cyp* gene cluster in which one biosynthetic gene was deleted. The WT copy of each gene was cloned downstream of the constitutive *ermE** promoter and EF-Tu RBS of pIJ10257, which integrates into the Φ BT1 phage attachment site. Complementation with the WT copy of the relevant gene restored heterologous expression (Claesen and Bibb 2010). However there is no certainty that the *ermE** promoter would be active in *P. alba*. Indeed a study in which a number of promoters which showed high level expression in *Streptomyces*, such as *ermE**p, were used in *Actinoplanes friuliensis* indicated that these promoters had poor expression levels in *A. friuliensis* (Wagner *et al.* 2009). It may be more useful to use the 454 genome scanning data to identify promoter sequences for *P. alba* genes likely to show high or constitutive activity, such as those for ribosomal proteins or elongation factors. It may be that the change in genomic context has a detrimental effect on expression levels. It is likely that expression of *psp* genes in the WT strain is coordinated in a distinct ratio which may be altered when expression of the complementation construct occurs at a different chromosomal location. This could be tested by restoring the WT gene at the original locus by homologous recombination. However this method could mask the involvement of a second site mutation which, if present near to the deleted gene, would also be replaced by the reparative recombination event.

In addition, it may be that there are additional regulatory features within the cluster which are interrupted during *in trans* expression. For example in the Pep5 gene cluster of *S.*

epidermidis, there is a RNA secondary structure between *pepI* and *pepA* which stabilises the *pepI* transcript for Pepl expression (Pag *et al.* 1999). Although a large section of the up and downstream regions were cloned along with the WT gene in the *psp* complementation constructs, other factors may affect mRNA stability of the *in trans* transcripts compared to the WT transcripts. Difficulties in fully complementing mutant phenotypes *in trans* may be a common problem with lantibiotic gene clusters with complex regulatory mechanisms, where the stoichiometry of individual gene products may be essential for effective regulation. For example, the $\Delta mibA$ and $\Delta mibV$ deletions in the *mib* cluster could not be complemented with the WT gene *in trans* (Foulston and Bibb 2010). Problems were also encountered when complementing deletion mutations in the actagardine gene cluster (Bell 2010).

This Chapter describes the creation of four *P. alba* mutants through the deletion of likely biosynthetic genes. Each deletion completely abolished planosporicin production, demonstrating that the corresponding genes are essential for biosynthesis but giving little insight into their precise function. Chapter 8 uses the approach described in this Chapter to make three further *P. alba* deletion mutants all affected in genes likely to play a role in regulation, the consequence of which is then investigated further.

7.7 Summary

P. alba was found to be tractable for genetic studies, and constructs were introduced by both (presumed) site-specific integration and homologous recombination.

The efficiency of conjugation was negatively correlated with the size of the construct (however some relatively small constructs were never successfully mobilised into *P. alba*).

Constructs carrying the hygromycin resistance cassette were readily mobilised into *P. alba* through conjugal transfer and exconjugants were selected with hygromycin at four times the MIC. Apramycin-resistant *P. alba* exconjugants could only be selected with apramycin at two times the MIC.

The targeted deletion of *pspA* marked with the hygromycin resistance cassette abolished production of planosporicin as determined by bioassay and MALDI-ToF analysis. This mutant could not be complemented in *trans* with WT *pspA*.

The generation of a 'scar' version of the *pspA* deletion mutant was not successful.

The deletion of *pspYZ* abolished production of planosporicin as determined by bioassay and MALDI-ToF analysis.

The deletion of *pspTU* abolished production of planosporicin as determined by bioassay and MALDI-ToF analysis. This mutant could not be complemented in *trans* with WT *pspTUV*.

The deletion of *pspV* abolished production of planosporicin as determined by bioassay and MALDI-ToF analysis.

Introduction of *mibA* and *mibABCD* into WT *P. alba* did not result in the detection of microbisporicin-related peaks when assessed by MALDI-ToF analysis.

Chapter 8 : Mutational analysis of planosporicin regulation

8.1 Introduction

The regulation of secondary metabolic pathways occurs at two levels; it can be pathway-specific (e.g. SARPs and LALs), or be through pleiotropic regulatory genes (e.g. BldD regulates spore formation and antibiotic production) (Bibb 2005). In the latter case, multiple signals are integrated through different regulators to ensure that a pathway is expressed only under the right conditions. The regulated expression of antibiotic biosynthetic genes clusters may also be essential to ensure the mechanism of producer-immunity is in place before high levels of antibiotic production commence.

The regulation of secondary metabolite production has been characterised in a range of actinomycetes. Several pleiotropic regulatory factors have been identified, such as the stringent factor ppGpp, as well as several pathway-specific regulatory protein families, such as the SARPs and LALs (as discussed in Chapter 1). The use of response regulators to coordinate regulation is common in many lantibiotic gene clusters from low-GC Gram-positive organisms. In *L. lactis*, regulation of nisin biosynthesis is mediated through the NisR response regulator in combination with the sensor histidine kinase NisK as a two-component regulatory system (Kuipers *et al.* 1995). In contrast, the proposed use of an ECF sigma factor, cognate anti-sigma factor and a transcriptional regulator in the *psp* cluster represents a less common mechanism of secondary metabolite regulation, and form the focus of this Chapter.

Chapter 6 demonstrated the successful heterologous production of planosporicin in *Nonomuraea*. This implied that not only were all of the essential biosynthetic enzymes encoded by the cloned gene cluster, but also that all of the necessary regulatory elements needed to synthesise planosporicin were also present. As stated in Chapter 7, conjugating into mycelia often results in exconjugants growing that contain a mixture of WT and engineered genomes. Consequently, genetic analysis of individual biosynthetic genes was carried out in the native producer, for which spores are available for conjugations. This Chapter continues the analysis of *psp* gene function through the specific deletion of putative regulatory genes to investigate the mechanism of regulation of the *psp* gene cluster.

8.2 Targeted deletions within the planosporicin gene cluster of *P. alba*

8.2.1 *pspX*

8.2.1.1 PCR analysis

pIJ12328 was PCR-targeted using the P1-FRT-*oriT*-*hyg*-FRT-P2 cassette from pIJ1700 to replace *pspX* using the method described in Chapter 7. This created pIJ12328 Δ *pspX*::*hyg*, named pIJ12351, which was conjugated into heat shocked *P. alba* spores. 48 sporulating *P. alba* colonies were visible just one week after the overlay with antibiotic selection. Four of these putative exconjugants were streaked three times on ISP4 agar containing 25 μ g/ml nalidixic acid and 40 μ g/ml hygromycin. One colony from each of the four exconjugants was transferred to liquid ISP4 medium with hygromycin selection to grow mycelium from which gDNA was isolated. PCR with the primers Δ ECFconfirmF and R, which anneal either side of the targeted region, amplified the expected 1971 bp Δ *pspX* fragment in all four exconjugants (data not shown). However in exconjugant 3, the 981 bp WT *pspX* fragment was also amplified, indicating that this exconjugant had resulted from a single crossover event. This was confirmed by the observed resistance to 50 μ g/ml kanamycin of the clone. Indeed when the entire plate of 48 exconjugants was replica plated onto ISP4 agar containing 50 μ g/ml kanamycin and then ISP4 agar with 40 μ g/ml hygromycin, 19 colonies were resistant to kanamycin and 29 colonies were sensitive to kanamycin (Figure 8.1). Kanamycin sensitivity indicates that the *neo* gene on the cosmid backbone is no longer present, most likely as the result of a double crossover event, and that the WT *pspX* gene has been replaced with the pIJ10700 cassette containing *hyg* and *oriT*. This was the largest number of exconjugants observed for any *P. alba* conjugation. On this basis it appears that 40 % of exconjugants were single crossovers and 60 % were double crossovers. Thus *P. alba* appears to be naturally fairly recombinogenic.

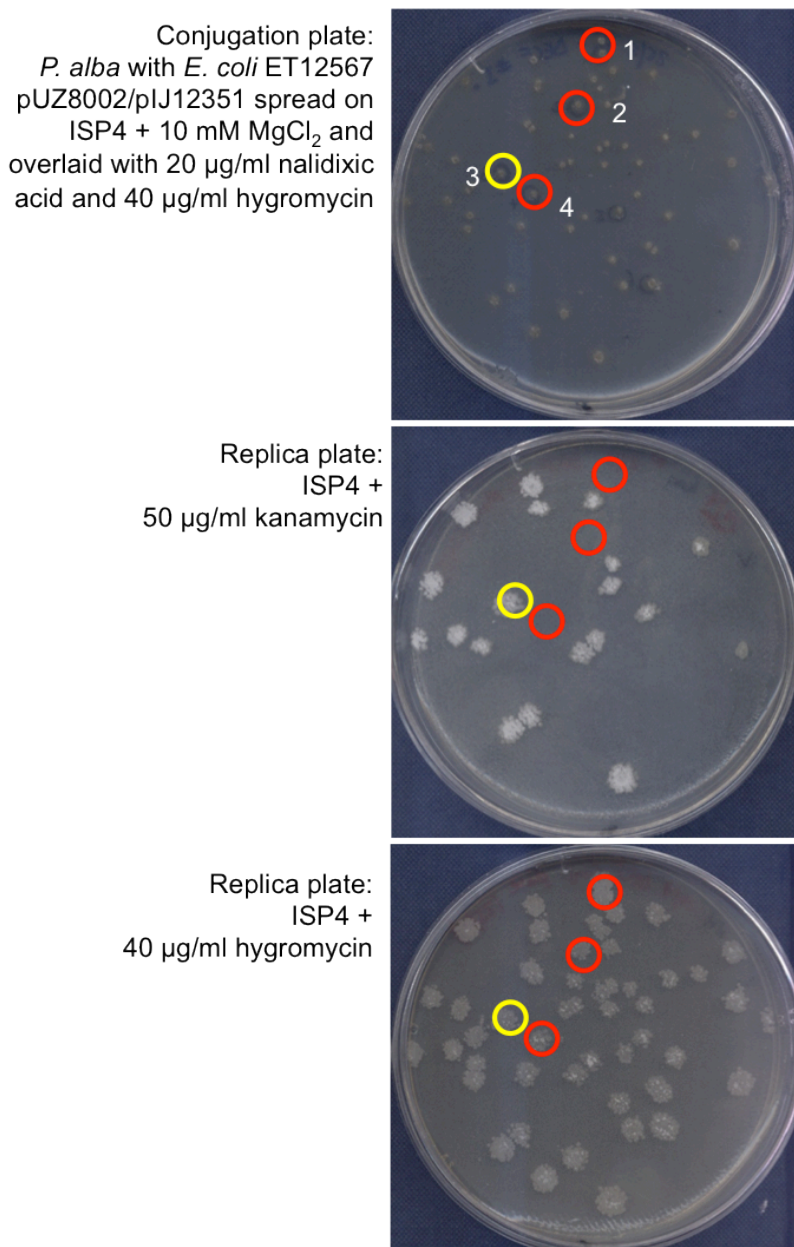


Figure 8.1 : Generation of *P. alba* Δ *pspX*::*hyg* exconjugants.

P. alba spores were subjected to heat shock at 50 °C for 10 minutes, cooled then mixed with *E. coli* ET pUZ8002 harboring pIJ12351 and plated on ISP4 agar containing 10 mM MgCl₂. After 20 hours at 30 °C, plates were overlaid with 1 ml H₂O containing antibiotics to give a final concentration of 25 µg/ml nalidixic acid and 40 µg/ml hygromycin (top). After a further 7 days incubation at 30 °C, this conjugation plate was replica-printed using a velvet onto ISP4 agar containing 50 µg/ml kanamycin (middle) then onto ISP4 agar containing 40 µg/ml hygromycin (bottom). All plates were incubated at 30 °C for a further 5 days before being photographed. 48 sporulating Hyg^R *P. alba* colonies are visible on the conjugation plate (top). Replica plating onto kanamycin and then hygromycin revealed 19 are also Kan^R, while 29 are Kan^S. The four colonies originally picked for analysis are numbered 1-4. Numbers 1, 2 and 4 are Hyg^R Kan^S while number 3 is Hyg^R Kan^R.

8.2.1.2 Bioassay and MALDI-ToF analysis

P. alba Δ *pspX*::*hyg*_1, 2 and 4, kanamycin-sensitive double crossovers with the pIJ10700 cassette in place of *pspX*, were tested for planosporicin production after growth in AF/MS liquid medium for seven days, both by bioassay and by MALDI-ToF analysis. All three exconjugants had a WT morphology when grown on agar or in liquid media, but produced no antibiotic compounds when culture supernatants were assayed against *M. luteus* (Figure 8.2). Likewise no peptides with a mass corresponding to planosporicin were detected by MALDI-ToF analysis (data not shown). Thus the replacement of *pspX* with P1-FRT-*oriT*-*hyg*-FRT-P2 abolished production of planosporicin. The *P. alba* Δ *pspX*::*hyg*_1 exconjugant, named M1303, was used in all subsequent analyses of the Δ *pspX* mutant. Figure 8.3 confirms that *P. alba* M1303 does not produce a compound with antibiotic activity, while Figure 8.4 confirms the lack of detection of planosporicin-associated peaks upon MALDI-ToF analysis of a culture supernatant.

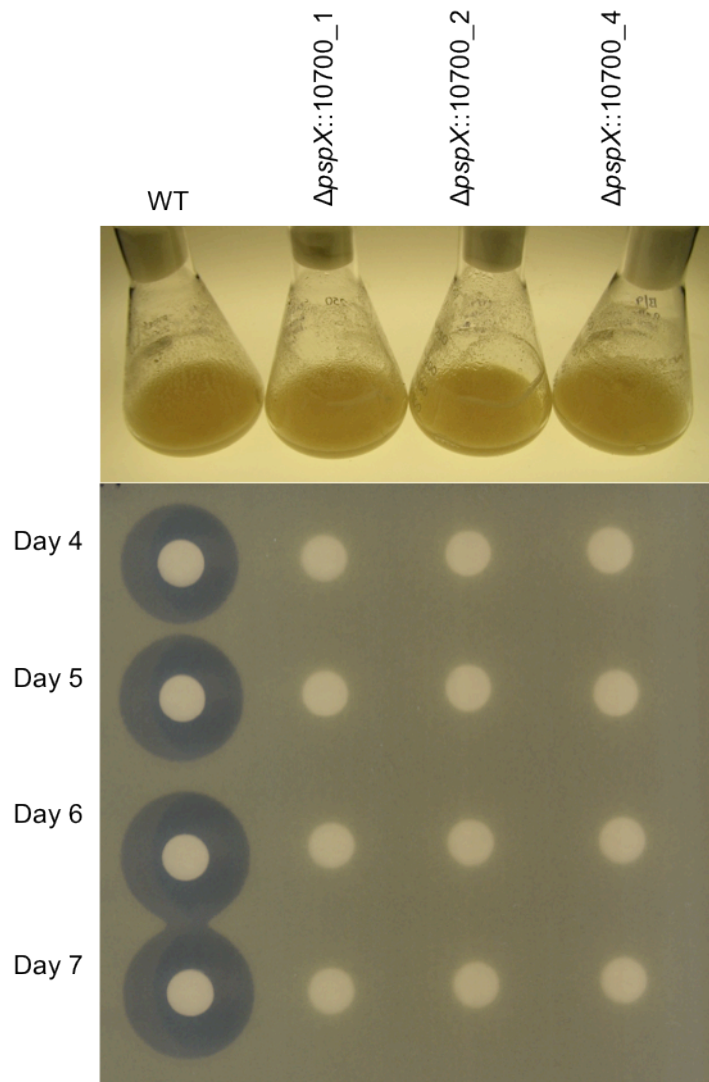


Figure 8.2 : Lack of production of antibiotic compounds from three *P. alba* $\Delta pspX::hyg$ exconjugants.

The three *P. alba* exconjugants were cultured in AF/MS medium for 7 days. A positive control of WT *P. alba* was cultured in parallel. Samples of culture supernatants taken after 4, 5, 6 and 7 days were tested for antibiotic activity. 40 μ l of supernatant was applied to each antibiotic assay disc which were allowed to dry before being applied to a lawn of *M. luteus*. The plate was incubated for 36 hours at 30 °C before zones of inhibition were photographed.

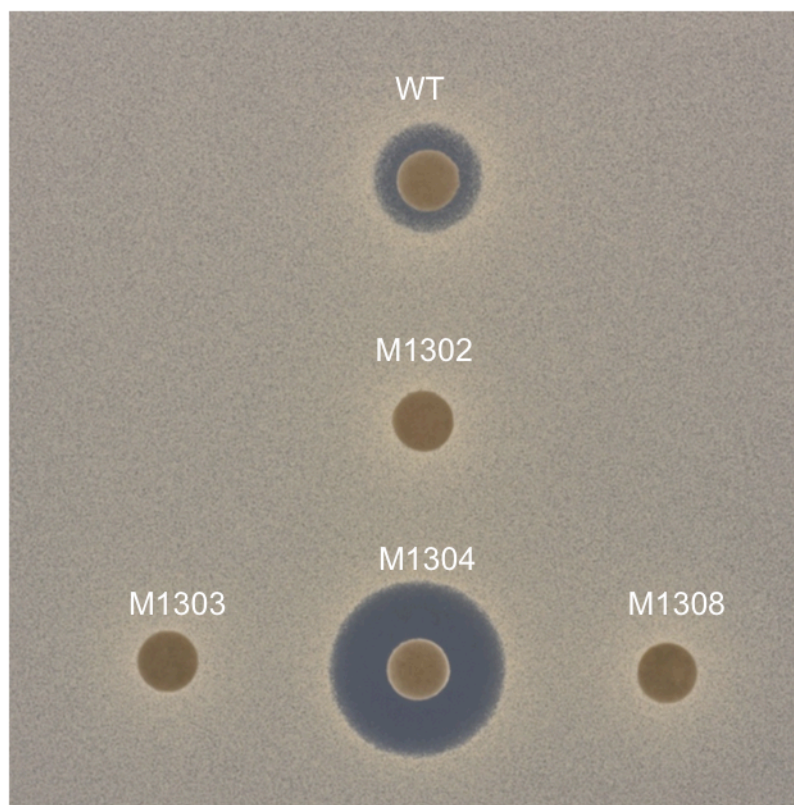


Figure 8.3 : Production of antibiotic compounds from four *P. alba* deletion mutants. Namely M1302 ($\Delta pspA::hyg$), M1303 ($\Delta pspX::hyg$), M1304 ($\Delta pspW::hyg$), and M1308 ($\Delta pspR::hyg$). The above *P. alba* strains were cultured in AF/MS medium for 6 days. A positive control of WT *P. alba* was cultured in parallel. Samples of culture supernatants were tested for antibiotic activity. 40 μ l of supernatant was applied to antibiotic assay discs which were laid onto a lawn of *M. luteus*. The plate was incubated for 36 hours at 30 °C before zones of inhibition were photographed.

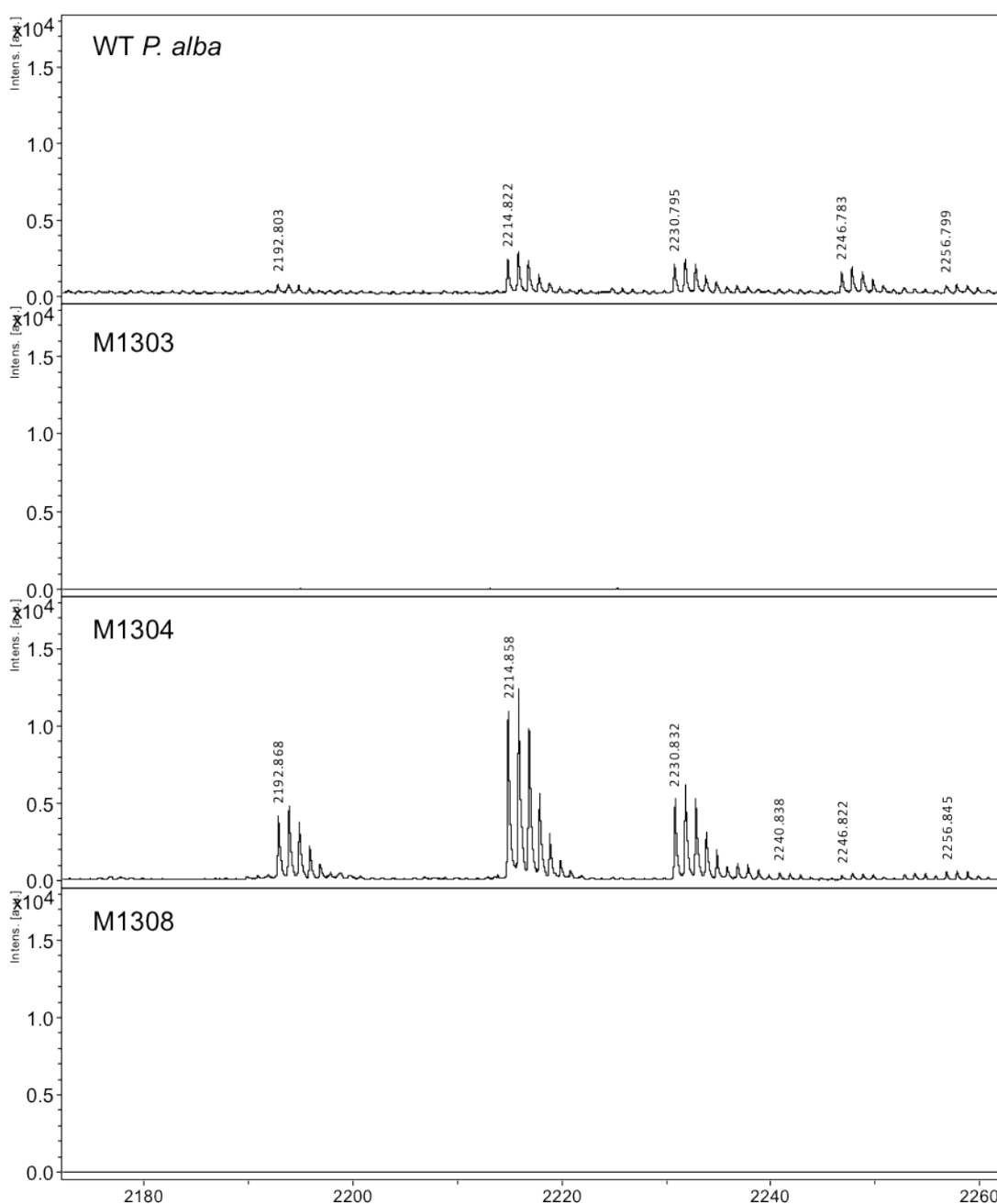


Figure 8.4 : Production of peptides from three *P. alba* deletion mutants.

Namely M1303 ($\Delta pspX::hyg$), M1304 ($\Delta pspW::hyg$), and M1308 ($\Delta pspR::hyg$). These four *P. alba* strains were cultured in AF/MS medium for 6 days. A positive control of WT *P. alba* was cultured in parallel. Supernatant samples were taken after 6 days of growth and analysed by MALDI-ToF mass spectrometry. Intensity is given on the y-axis in arbitrary units (au) and the mass/charge ratio (m/z) on the x-axis. The monoisotopic mass of each m/z peak is labelled.

8.2.2 *pspW*

8.2.2.1 PCR analysis

pIJ12328 was PCR-targeted using the P1-FRT-*oriT*-*hyg*-FRT-P2 cassette from pIJ1700 to replace *pspW* using the method described in Chapter 7. This created pIJ12328 Δ *pspW*::*hyg*, designated pIJ12522, which was conjugated into heat shocked *P. alba* spores. Five exconjugants were streaked three times for growth on hygromycin selective plates. These Hyg^R exconjugants were tested for growth in the presence of 50 μ g/ml kanamycin. Three out of the five exconjugants were Kan^S; *P. alba* Δ *pspW*::*hyg*_3, *P. alba* Δ *pspW*::*hyg*_5 and *P. alba* Δ *pspW*::*hyg*_3(2). Sensitivity to kanamycin suggested that they had resulted from a double-crossover recombination event that had replaced *pspW* in the *P. alba* chromosome with the pIJ10700 cassette. The two Kan^R exconjugants were likely to reflect single-crossover insertion of pIJ12523, containing the *neo* gene, into the *P. alba* chromosome. After growth in liquid medium containing hygromycin but lacking kanamycin, single colonies were assayed for kanamycin sensitivity. One of the Kan^R exconjugants yielded a Kan^S segregant, creating *P. alba* Δ *pspW*::*hyg*_17. PCR analysis of gDNA isolated from these four strains demonstrated a shift in band size from 1318 bp, corresponding to WT *pspW*, to 1917 bp, corresponding to Δ *pspW*::*hyg*, when using primers annealing to the regions flanking *pspW* (data not shown), whereas no amplification product was observed when both primers annealed within *pspW* (data not shown; see Figure 8.5 for analysis of one of the clones). *P. alba* Δ *pspW*::*hyg*_3 was assigned the strain number M1304.

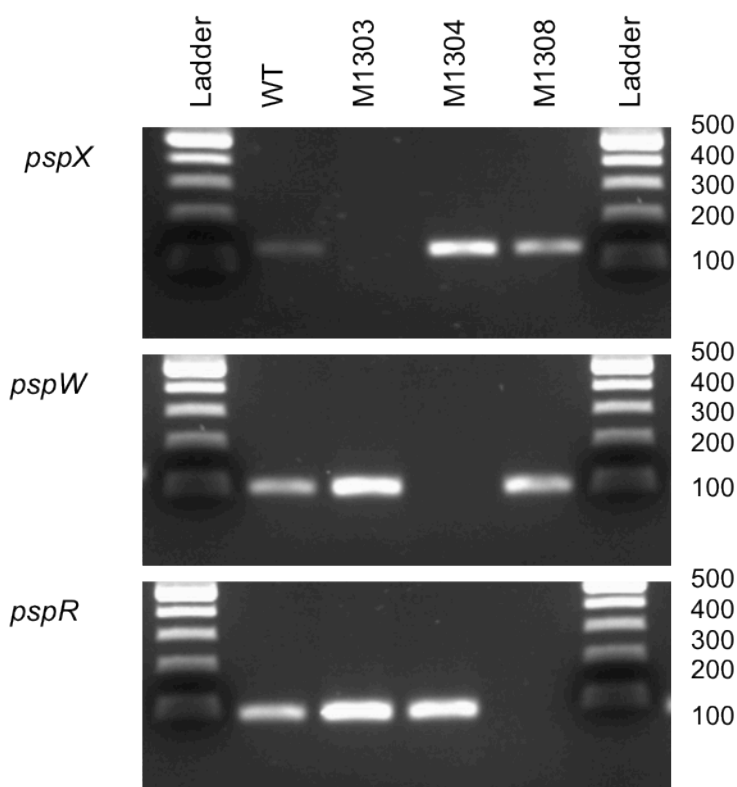


Figure 8.5 : PCR confirmation of three *P. alba* deletion mutants.

Namely M1303 ($\Delta pspX::hyg$), M1304 ($\Delta pspW::hyg$), and M1308 ($\Delta pspR::hyg$). Genomic DNAs extracted from each strain were used as templates to amplify *pspX*, (*pspXF* and *pspXR*; 133 bp), *pspW* (*pspWF* and *pspWR*; 102 bp) and *pspR* (*pspRF* and *pspRR*; 93 bp) from the *psp* gene cluster. Genomic DNA from WT *P. alba* was used as a positive control. PCR products were run on a 2 % agarose gel by electrophoresis. The ladder is 100 bp DNA ladder (NEB) with band sizes annotated in bp.

8.2.2.2 Bioassay and MALDI-ToF analysis

The four *P. alba* $\Delta pspW::hyg$ clones were tested for their ability to produce planosporicin. Supernatant samples from all four AF/MS liquid cultures inhibited the growth of *M. luteus* and MALDI-ToF analysis confirmed the presence of planosporicin. Figure 8.3 depicts the bioassay result for supernatant from *P. alba* $\Delta pspW::hyg_3$ (strain M1304) compared to WT supernatant. In agar diffusion assays, antibiotic activity is correlated with the square of the inhibition zone radius around the antibiotic assay disc (Drugeon *et al.* 1987). The difference in zone size between the WT and M1304 sample implies that deletion of the anti-sigma factor gene *pspW* markedly enhanced planosporicin production. Figure 8.4 shows the peaks obtained from MALDI-ToF analysis in which the y-axis scale (peak intensity in arbitrary units) is the same for all samples. Although not strictly quantitative, peak intensity is clearly much higher with the supernatant from *P. alba* M1304 compared

to WT *P. alba*. To confirm this result, approximately equivalent numbers of *P. alba* WT and M1304 spores were patched onto AF/MS agar plates. WT *P. alba* usually only starts to produce planosporicin after 5 or 6 days incubation, whereas after just 2 days, *P. alba* M1304 produced an antibiotic compound which inhibited the growth of *M. luteus* (Figure 8.6). This early production of planosporicin is unlikely to be due to an enhanced rate of growth or development, as *P. alba* M1304 grew at a similar rate to the WT strain on agar and in and liquid medium. The early production of planosporicin in the M1304 mutant is investigated further in Section 8.5.

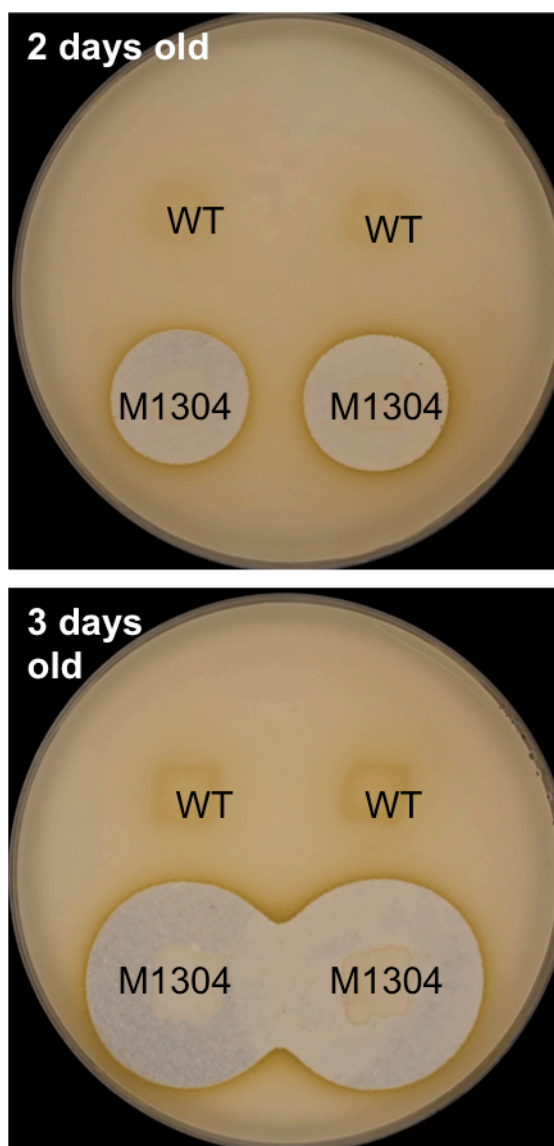


Figure 8.6 : Production of antibiotic compounds by *P. alba* M1304 and the WT strain. Approximately 10^6 *P. alba* WT and M1304 spores were patched on AF/MS agar plates and incubated at 30 °C. After 2 or 3 days, plates were overlaid with SNA containing *M. luteus*. The plates were incubated for a further 36 hours at 30 °C before zones of inhibition were photographed.

8.2.3 *pspR*

8.2.3.1 PCR analysis

pspR is located towards the left hand border of the *psp* gene cluster. In the reduced cosmid pIJ12328 there is just 2.5 kb of sequence upstream of *pspR*. In contrast, in cosmid pIJ12357, there is 16 kb of sequence upstream of *pspR* and 13 kb downstream (the 7 kb downstream of *pspV* were removed to create pIJ12327 in Chapter 6). Consequently, to provide larger flanking homologous regions for recombination, pIJ12357 was chosen for PCR targeting of *pspR*, which was replaced with the P1-FRT-*oriT*-*hyg*-FRT-P2 cassette from pIJ10700. The resulting deleted construct, pIJ12525, was mobilised into *P. alba* via conjugation with *E. coli* ET12567/pUZ8002/pIJ12525. After several conjugation attempts, one exconjugant grew well when streaked onto ISP4 media containing hygromycin. This exconjugant was inoculated into ISP4 liquid medium containing hygromycin and gDNA was prepared. PCR amplification using primers flanking *pspR* amplified a 2394 bp fragment corresponding to $\Delta pspR::hyg$ instead of the 1569 bp WT *pspR* fragment (data not shown). Likewise, primers annealing within *pspR* amplified a fragment from all other *psp* mutants but not from *P. alba* $\Delta pspR::hyg$, subsequently named *P. alba* M1308 (Figure 8.5). In an attempt to create a second *P. alba* $\Delta pspR::hyg$ exconjugant, primers *pspR_2kbF* and *pspR_2kbR* were used to amplify ~2 kb either side of $\Delta pspR::hyg$ using pIJ12525 as template. The resulting 6.3 kb fragment consisted of a region of the *psp* cluster with the resistance cassette from pIJ10700 in place of *pspR* which was cloned into pGEM-T to create pIJ12528. It was thought the relatively small size of this vector compared to pIJ12525 would increase conjugation efficiency. Despite many attempts to conjugate this construct into *P. alba*, no further *P. alba* $\Delta pspR::hyg$ exconjugants were identified. Sequencing of the *oriT* cassette present in pIJ12528 confirmed the absence of mutations that may have prevented conjugal transfer.

8.2.3.2 Bioassay and MALDI-ToF analysis

P. alba M1308 ($\Delta pspR::hyg$) was inoculated from a seed culture into AF/MS liquid medium alongside a WT control, and supernatants were sampled after 5 and 7 days. The M1308 supernatant did not inhibit growth of *M. luteus* (Figure 8.3). The absence of planosporicin-associated peaks was confirmed by MALDI-ToF analysis (Figure 8.4). Thus it would appear that *pspR* is a key regulator of planosporicin biosynthesis. This idea is discussed further in Section 8.5.

8.3 Complementations

8.3.1 Complementing regulatory mutants

All mutants were made through the targeted replacement of a *psp* gene with the hygromycin resistance cassette. Two deletion mutants in proposed regulatory genes abolished planosporicin production. Attempts were made to complement these mutations *in trans* through the integration of a construct containing the WT gene. As in Chapter 7, all complementation constructs were based on the vector pSET152, which integrates into the chromosome presumably at the Φ C31 site and carries *aac(3)IV* conferring apramycin resistance. Thus the Hyg^R mutants were conjugated with complementation constructs and exconjugants were selected with apramycin.

8.3.1.1 *pspX*

The P_{pspX} promoter (defined as the region between *pspA* and *pspX*) and *pspX* were amplified with primers to introduce a *Bam*HI site upstream of P_{pspX} and a *Xba*I site downstream of *pspX*. This fragment was cloned into pGEM-T and sequenced to check that no mutations had been introduced. P_{pspX} -*pspX* was subsequently ligated into *Bam*HI plus *Xba*I cleaved pSET152 to create pIJ12539. This plasmid was mobilised into *P. alba* M1303 through conjugal transfer and exconjugants selected with 20 μ g/ml apramycin.

Over 20 *P. alba* colonies growing through the overlay were streaked out on ISP4 agar medium containing 50 μ g/ml apramycin and 20 μ g/ml nalidixic acid. However, only one grew well on the higher level of apramycin and was subsequently inoculated into liquid medium and gDNA prepared. This gDNA was assessed by PCR using primers annealing within pIJ12539 (pSET152F and pSET152R2) and the expected 1122 bp band corresponding to P_{pspX} -*pspX* was observed (data not shown). Two clones from this exconjugant were grown in AF/MS liquid medium and the resulting supernatants tested for the production of an antibiotic compound capable of inhibiting the growth of *M. luteus*. The presence of the complementation construct pIJ12539 in *P. alba* M1303 restored planosporicin production, as demonstrated in the bioassay (Figure 8.7) and by MALDI-ToF analysis (Figure 8.8).

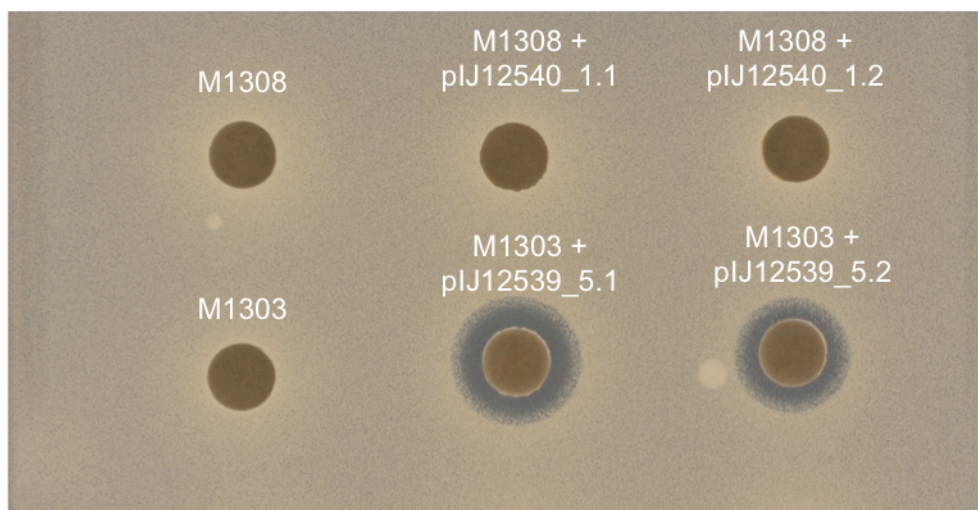


Figure 8.7 : Production of antibiotic compounds from *P. alba* deletion mutants complemented in *trans* with the WT gene.

Two clones of *P. alba* M1308 ($\Delta pspR::hyg$) with integrated pIJ12540 ($P_{pspR}-pspR$) and *P. alba* M1303 ($\Delta pspX::hyg$) with integrated pIJ12539 ($P_{pspX}-pspX$) were cultured in AF/MS medium for 5 days. Samples of culture supernatants were tested for antibiotic activity. 40 μ l of supernatant was applied to antibiotic assay discs which were laid onto a lawn of *M. luteus*. The plate was incubated for 36 hours at 30 °C before zones of inhibition were photographed.

8.3.1.2 *pspR*

The region between *pspQ* and *pspR* was defined as the native promoter P_{pspR} and was amplified along with *pspR* using primers that introduced *Bam*HI and *Xba*I sites at the 5' and 3' ends, respectively, of *pspR*. This fragment was cloned into pGEM-T, confirmed by sequencing, the *Bam*HI/*Xba*I fragment excised and ligated into the *Bam*HI/*Xba*I site of pSET152 to create pIJ12540. This plasmid was mobilised into *P. alba* M1308 by conjugal transfer, using 20 μ g/ml apramycin to select exconjugants.

One putative *P. alba* M1308 pIJ12540 exconjugant was observed to grow well when streaked on 50 μ g/ml apramycin. Genomic DNA prepared from this exconjugant was assessed by PCR using primers annealing within the integrated pIJ12540 construct. Amplification of the 1474 bp band corresponding to *pspR* confirmed the presence of pIJ12540 in this *P. alba* strain. Two clones from this exconjugant were grown in AF/MS liquid medium alongside a positive control of WT *P. alba*. After 5, 6, 8, 13 and 19 days of incubation, samples of supernatant were tested but no antibiotic compounds were detected in bioassay with *M. luteus*. Subsequent repetitions of this assay gave the same result. Five day old supernatants consistently failed to yield halos of inhibition (Figure 8.7) and MALDI-ToF analysis failed to identify planosporicin-associated peaks (Figure 8.8).

Likewise 5, 6 and 7 day old patches of mycelium grown on AF/MS agar also failed to reveal a halo of inhibition when overlaid with *M. luteus* (data not shown).

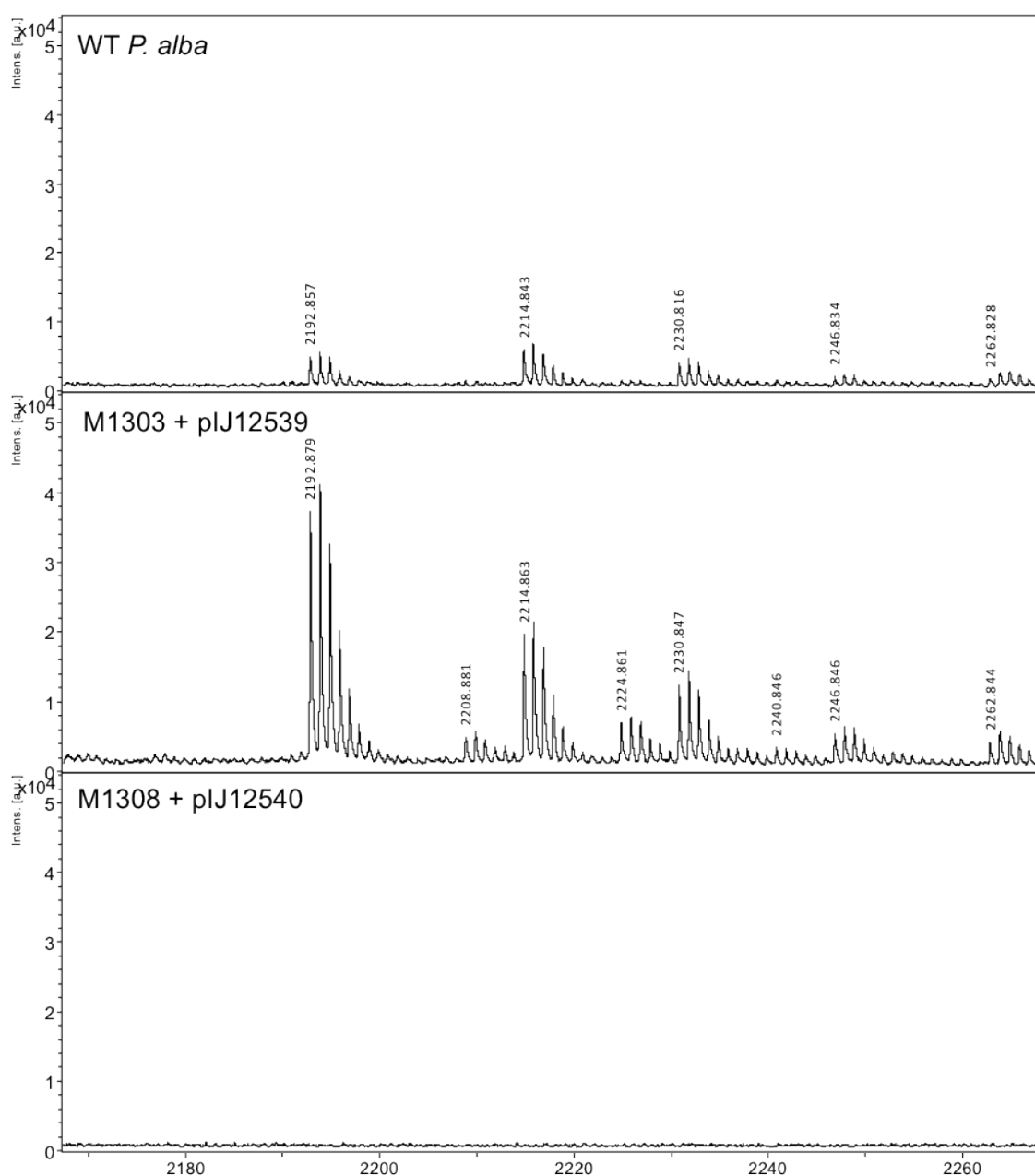


Figure 8.8 : Production of peptides from *P. alba* deletion mutants complemented in *trans* with the WT gene.

P. alba M1308 ($\Delta pspR::hyg$) with integrated pIJ12540 (P_{pspR} -*pspR*) and *P. alba* M1303 ($\Delta pspX::hyg$) with integrated pIJ12539 (P_{pspX} -*pspX*) were cultured in AF/MS medium. Supernatant samples were taken after 5 days of growth and analysed by MALDI-ToF mass spectrometry. Intensity is given on the y-axis in arbitrary units (au) and the mass/charge ratio (m/z) on the x-axis. The monoisotopic mass of each m/z peak is labelled.

8.4 Computational analysis

In addition to the bioinformatic analysis of the coding regions of the *psp* gene cluster in Chapter 5, additional programs were used to identify regions that might be associated with biologically active sites such as promoter motifs, activator/repressor binding sites, attenuators or RNA processing sites, all of which may play a role in the regulation of planosporicin biosynthesis.

8.4.1 Dyad symmetries

Clone Manager (Sci-Ed) was used to identify two areas of DNA whose base pair sequences are inverted repeats of each other. These dyad symmetries could form stable RNA stem-loop or hairpin structures, or act as a binding site for transcriptional regulators. The program first searches for sequences that are repeated in the opposite direction on the same strand and then uses RNA energy parameters to determine which loops would be most stable.

8.4.1.1 Attenuators

Clone Manager reported that the largest and most stable RNA hairpin loop structure which could be formed from the *psp* gene cluster occurred between *pspA* and *pspB*. This loop has the stem sequence: GCCCGUGGGGCCGGGC and has a ΔG value of -35.4 (Figure 8.9). This putative structure is located 14 bp downstream of the *pspA* stop codon and is likely to represent a transcription terminator. The hairpin loop would form in the mRNA strand during transcription, causing RNA polymerase to dissociate from the DNA template, resulting in some level of transcription termination.

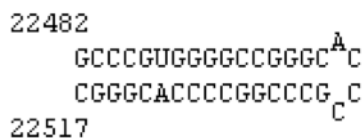


Figure 8.9 : The 36 bp hairpin loop located between *pspA* and *pspB*.

Numbers correspond to the position in the pIJ12321 insert DNA sequence. *pspA* is located from 22297-22467 and *pspB* is 22561-25764.

8.4.1.2 Binding sites

Chapter 5 identified PspR as a transcriptional regulator of the LuxR family. PspR may exist in the cell as a homodimer, in which case an inverted repeat may be evident of a cognate binding site(s). For example the LuxR dimer (together with RNA polymerase) binds the *lux* box, a 20 bp inverted repeat (Stevens *et al.* 1994). It was proposed that the PspR binding site could be identified as a palindromic region in the promoter regions of the relevant *psp* operons. Dyad symmetries were only observed in the promoter regions of *pspEF* and *pspJYZQ*. The putative P_{*pspEF*} site is a 10 bp inverted repeat with the sequence CGCCGGGC^{CG}, separated by 17 nucleotides with one mismatch (the underlined G at position 7 is a T in the repeat sequence). The putative P_{*pspJYZQ*} site is an 11 bp inverted repeat with the sequence CGGCCGGCGGC, separated by 44 nucleotides. Thus neither appear likely to represent genuine binding sites for PspR.

8.4.2 Promoter motifs

As described in Chapter 5, *pspX* is likely to encode an ECF sigma factor. The consensus motif found at many ECF-dependent promoters differs from those of other σ^{70} proteins. This motif is characterised by a highly conserved 'AA' motif in the -35 region that induces a DNA geometry appropriate for sigma factor binding (Lane and Darst 2006). ECF-dependent promoters are more generally characterised by an 'AAC' motif in the -35 region and a 'CGT' motif in the -10 region (Staron *et al.* 2009).

The annotation of the *psp* gene cluster in Chapter 5 was based on a distinctive pattern of GC content across codons that allows the identification of protein coding regions; translational start sites were identified by defining an appropriate RBS upstream of a start codon (Bibb *et al.* 1984; Kieser *et al.* 2000). On the basis of this annotation, the *psp* genes were grouped into five operons; *pspABCTUV*, *pspXW*, *pspJYZQ*, *pspR* and *pspEF*. The five promoter regions were analysed using MEME (Multiple EM for Motif Elicitation) (Bailey and Elkan 1994). Figure 8.10 depicts the alignment of the consensus sequences identified upstream of *pspE*, *pspJ*, *pspX* and *pspR*. This consensus motif was derived by MEME analysing four intergenic regions from the *psp* gene cluster (*ORF-10-pspE*, *pspR-pspQ*, *pspJ-pspW* and *pspX-pspA*) together with five from the *mib* gene cluster (*orf1-mibJ*, *mibO-mibQ*, *mibQ-mibR*, *mibX-mibA* and *mibV-mibE*). Other intergenic regions of the *psp* cluster were also analysed (e.g. *pspZ-pspQ*) but a consensus motif was not identified. The consensus motif in the *mib* regions had previously been identified as characteristic for ECF sigma factor binding with 'GAACC' at -35 and 'GCTAC' at -10 (Foulston and Bibb 2011). Table 8.1 shows that the 'GAACC' at -35 is also present in four *psp* promoters, while the conserved GXXXC motif at -10 is characterised by a preference for GCT, which is most clearly present in the *pspE* and *pspJ* promoter regions.

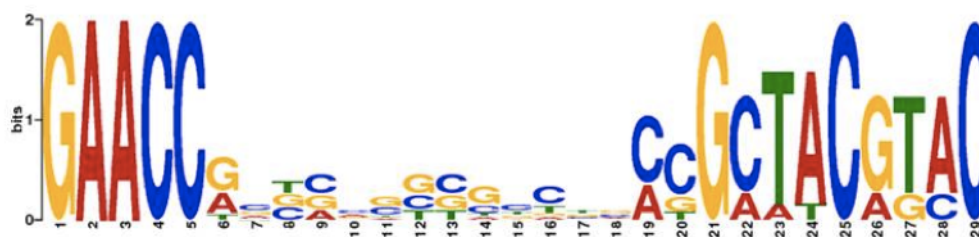


Figure 8.10 : A sequence logo of the consensus sequence compiled by MEME.

The motif represents four promoter regions of the planosporicin gene cluster (upstream of *pspE*, *pspJ*, *pspX* and *pspR*) with five regions of the microbisporicin cluster (upstream of *mibJ*, *mibQ*, *mibR*, *mibX* and *mibE*).

Overall, four *psp* promoter regions contain a consensus motif similar to those commonly recognised by ECF sigma factors (Table 8.1, Figure 8.10). This is in contrast to the classical consensus promoter sequences of primary sigma factors, the -35 'TTGACA' and -10 'TATAAT' regions. The only operon which is entirely missing an ECF sigma factor binding motif is *pspABCTUV*. The only other regulator in the cluster is the LuxR-type transcriptional regulator PspR. Active PspR may be a homodimer, in which case an inverted repeat may be evident as a binding site. Section 8.4.1.2 failed to identify a binding site for the PspR homodimer upstream of *pspA*. In the *mib* cluster, a 'GAACC' motif was identified 89 nucleotides upstream of *mibA*. This is the same as the -35 motif for other *mib* operons. However this conservation was not extended to the 'GCTAC' motif observed at the -10 region of other *mib* operons. On this basis it was proposed that the *mibABCDTUV* operon is regulated by a member of the σ^{70} family, possibly a vegetative sigma factor (Foulston and Bibb 2011). In the *psp* cluster, the 'GAACC' motif was only observed in the promoter regions of *pspE*, *pspR*, *pspJ* and *pspX*. There was no such motif in the *pspA* promoter region. For *pspEF* and *pspJYZQ*, the motif for ECF sigma factor binding is conserved at both the -35 and -10 regions, whereas for *pspX* and *pspR*, only the -35 region is highly conserved (Figure 8.10). The variation in conservation of the ECF sigma factor consensus sequence could imply that operons *pspEF* and *pspJYZQ* are highly expressed by σ^{PspX} , while *pspR* and *pspXW* are less so. This would fit with the observation that many ECF sigma factors auto-regulate their own expression (Staron *et al.* 2009), and that they play a regulatory, not catalytic role, in cell physiology. Alternatively, the regulatory genes *pspXW* and *pspR* may be recognised by a different member of the σ^{70} family.

	bp from GAACC to start	p-value	Promoter sequence from consensus sequence to ATG/GTG start codon
<i>pspE</i>	66	1.81e ⁻¹²	<u>GAACCACGCAGCCGGCGGCCGCTACGTAC</u> AGGGCGTGACCACGATCGGCCGAGCAAAGGAGACACGGTG
<i>pspJ</i>	122	2.64e ⁻¹¹	<u>GAACCACGGCTCTGCGGACC</u> <u>GCTACGTAC</u> CAGGTACCGGAAGGCAGGACGCCGACCGGGCGGAAGCGGGGTTCC CTCCACGAGGGACGTCGCCCGGCCGCGACCGACCAGGGAAGGCC ACGACATG
<i>pspX</i>	34	1.09e ⁻¹⁰	<u>GAACC</u> <u>GATCGCGGTTTTCCGG</u> <u>A</u> <u>TACGTAC</u> ATGGCGTG
<i>pspR</i>	383	5.97e ⁻¹⁰	<u>GAACCACGAACGTGATCAAG</u> <u>G</u> <u>AATCGGCC</u> CGACAATGATGATCACCGAACTCACGGGGGACCGCATTGCGCCGG CACGCTTCGCAGGCTGTCGAACTCAACGTCCTGGCCACTCCGGG CTGAAAGCCCACTTCGACTCACCGAACACCTCCCTCAACACGATCC GCTCCGGCCTGGACCACGTGCACTTCGTGACAGTGAGGACCTCG CCGCGTCCGACGAGTCACTCGACCTCCCGCTGCTGACACCGGTCA CCAGCGAGCTCCAGGCCGAGGGACTGCGCACCTGCGCGTACGG AACCCGCTGGCCCTGCTCATCGCGTCAACACCGACGTCTCGGGA TTCCAGACCTACGACGCGATCCGCTCGGGCGCGAATTTCTGT

Table 8.1 : Promoter motifs identified by MEME.

The five *psp* promoter regions were entered, of which four were used in the alignment; P_{pspE} , P_{pspJ} , P_{pspX} and P_{pspR} . The consensus sequence identified by MEME is underlined, in which the -35 and -10 consensus sequences are in red.

8.5 Transcriptional analysis

The regulation of the *psp* gene cluster was investigated through analysis of the transcription of the cluster in five *P. alba* strains: M1302 ($\Delta pspA::hyg$), M1303 ($\Delta pspX::hyg$), M1304 ($\Delta pspW::hyg$), M1308 ($\Delta pspR::hyg$) and the WT strain. The five strains were inoculated in duplicate to the same OD₄₅₀ and grown in parallel in AF/MS liquid medium. OD₄₅₀ was monitored at regular intervals to create a growth curve (Figure 8.11). M1308 grew at approximately the same rate as the WT strain, whereas M1302, M1303 and M1304 grew at a faster rate. However all strains reached the same final optical density at stationary phase. Consequently four time points of 41, 73, 113 and 162 hours were used in order to cover the different growth phases of all five strains.

Culture supernatants from each time point were used to assess when production of planosporicin began in each strain. As observed previously, M1302, 1303 and M1308 did not produce planosporicin. There was no halo of *M. luteus* inhibition even in bioassays with supernatant collected after more than 11 days growth (Figure 8.11).

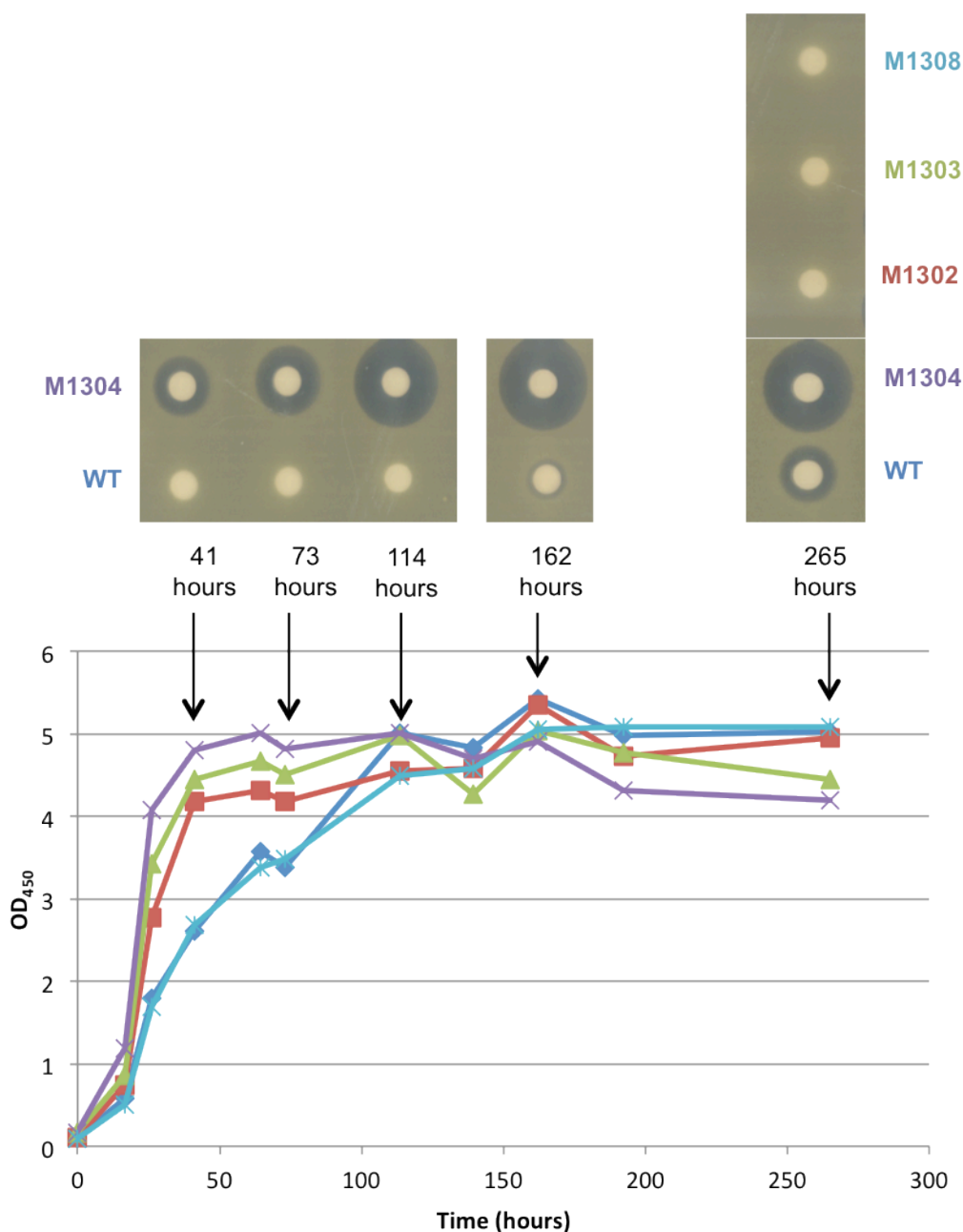


Figure 8.11 : Growth and planosporicin production in five *P. alba* strains.

Growth was measured by determining optical density at 450 nm averaged for two independent cultures of each strain at each time point; *P. alba* WT (dark blue diamonds), M1302 Δ pspA::hyg (red squares), M1303 Δ pspX::hyg (green triangles), M1304 Δ pspW::hyg (purple crosses) and M1308 Δ pspR::hyg (light blue stars). Mycelium was collected after 41, 73, 114 and 162 hours for RT-PCR analysis. Supernatant was collected at 41, 73, 114, 162 and 265 hours to be assayed for antibacterial activity. 40 μ l of supernatant was applied to antibiotic assay discs which were laid onto a lawn of *M. luteus*. The plate was incubated for 36 hours at 30 °C before zones of inhibition were photographed.

Two methods of RNA extraction were trialled on WT *P. alba* mycelium grown in two different media and sampled at two different time points. Good quality RNA was prepared using either the pestle and mortar method or silica beads with FastPrep (methods described in Chapter 2). Gel electrophoresis of approximately 1 µg RNA revealed both the 23S and 16S rRNA bands, and often larger mRNA aggregates (Figure 8.12). RNA was extracted from mycelium sampled from one of the duplicate cultures after 41, 73, 113 and 162 hours using the pestle and mortar method. For all 20 RNA extractions, a control PCR experiment using primers RT_hrdBF and RT_hrdBR confirmed the absence of DNA contamination. Quantification by Nanodrop showed that all RNA samples gave a 260/280 ratio between 1.9 and 2.3, and a 260/230 ratio greater than 1.8, and in most cases also greater than the 260/280 reading, as expected for highly purified RNA. Gel electrophoresis of approximately 1 µg RNA revealed in most cases the 23S rRNA band was larger than the 16S rRNA band, indicating that the samples have not suffered significant degradation. Subsequently, the RNA samples were used as templates for reverse transcription as detailed in Chapter 2. cDNA from each reverse transcriptase reaction was diluted 1 in 10 in H₂O and determined by Nanodrop analysis to contain approximately 80-110 ng cDNA.

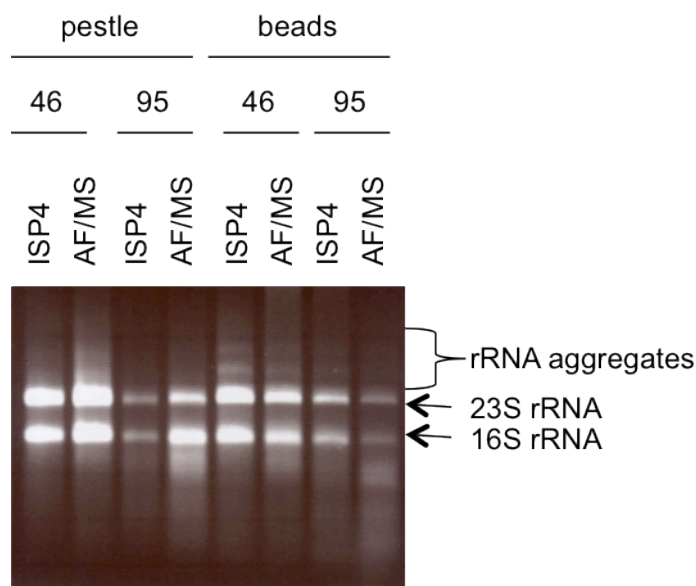


Figure 8.12 : Appearance of RNA extracted from *P. alba* separated by gel electrophoresis. RNA was extracted from WT *P. alba* cultured in two different types of media (ISP4 or AF/MS) at two different time points (46 and 95 hours) through two different methods. One method used a pestle and mortar to break open cells while the other used silica beads in a FastPrep machine (described in Chapter 2). Both methods gave good quality RNA (with a 260/280 ratio of 1.94-2.11 and a 260/230 ratio >1.99). Approximately 1-2 µg of RNA was run on a 1 % agarose gel to reveal bands of 23S and 16S rRNA as well as larger rRNA aggregates (labelled).

8.5.1 Reverse transcriptase (RT) PCR

Analysis of *psp* gene expression at the four time points shown in Figure 8.11 was initially investigated by PCR of cDNA (Reverse Transcriptase (RT) PCR). As a positive control for each cDNA template, a gene with 87 % nucleotide identity across 1036 nucleotides of 3' sequence of *hrdB* from *S. coelicolor* was identified from *P. alba* 454 sequence data. *hrdB* encodes the major vegetative sigma factor σ^{70} and is commonly used as a control in RT-PCR and quantitative (q)RT-PCR experiments in *Streptomyces* sp. (Hesketh *et al.* 2009). Primers annealing to this constitutively expressed gene served as a positive control for cDNA synthesis. In addition, primers annealing to the rRNA gene set *rrnD* were used as an additional positive control. Further primers were designed to amplify short (80-150 bp) fragments from the transcripts of *psp* genes (Chapter 2). RT-PCR was performed on RNA from each of the four time points indicated in Figure 8.11. PCR reactions performed on RNA samples that had not been treated with reverse transcriptase served as controls for DNA contamination. To get an overview of transcription within the cluster, primers were used which anneal within each of the proposed *psp* operons (Figure 8.13).

The expression of *hrdB* occurred at roughly equal levels in all five strains at each time point (Figure 8.13). The lower expression levels in later time points may represent the decrease in transcription that would generally be expected during stationary growth. All of the negative control samples (from cDNA synthesis reactions lacking reverse transcriptase) did not amplify a visible PCR product, demonstrating that the RNA samples were not contaminated with genomic DNA (Figure 8.13).

Overall, the strongest level of expression of *psp* genes occurred in strain M1304 (Figure 8.13). There is also a good level of expression of all *psp* genes in the WT after 162 hours growth, corresponding with the observed production of planosporicin in bioassay analysis (Figure 8.13). At earlier WT samples, expression of *pspE*, *pspR*, *pspQ*, *pspY*, *pspJ*, *pspW*, *pspX* and *pspA* is evident. The hairpin loop between *pspA* and *pspB* observed in Section 8.4.1.1 likely acts as a transcriptional attenuator, reducing transcription of *pspBCDTUV* relative to *pspA*. Expression levels of *pspE*, *pspQ*, *pspY*, *pspJ*, *pspT* and *pspV* were lower in strains M1302, M1303 and M1308 compared to M1304 and the WT strain (Figure 8.13).

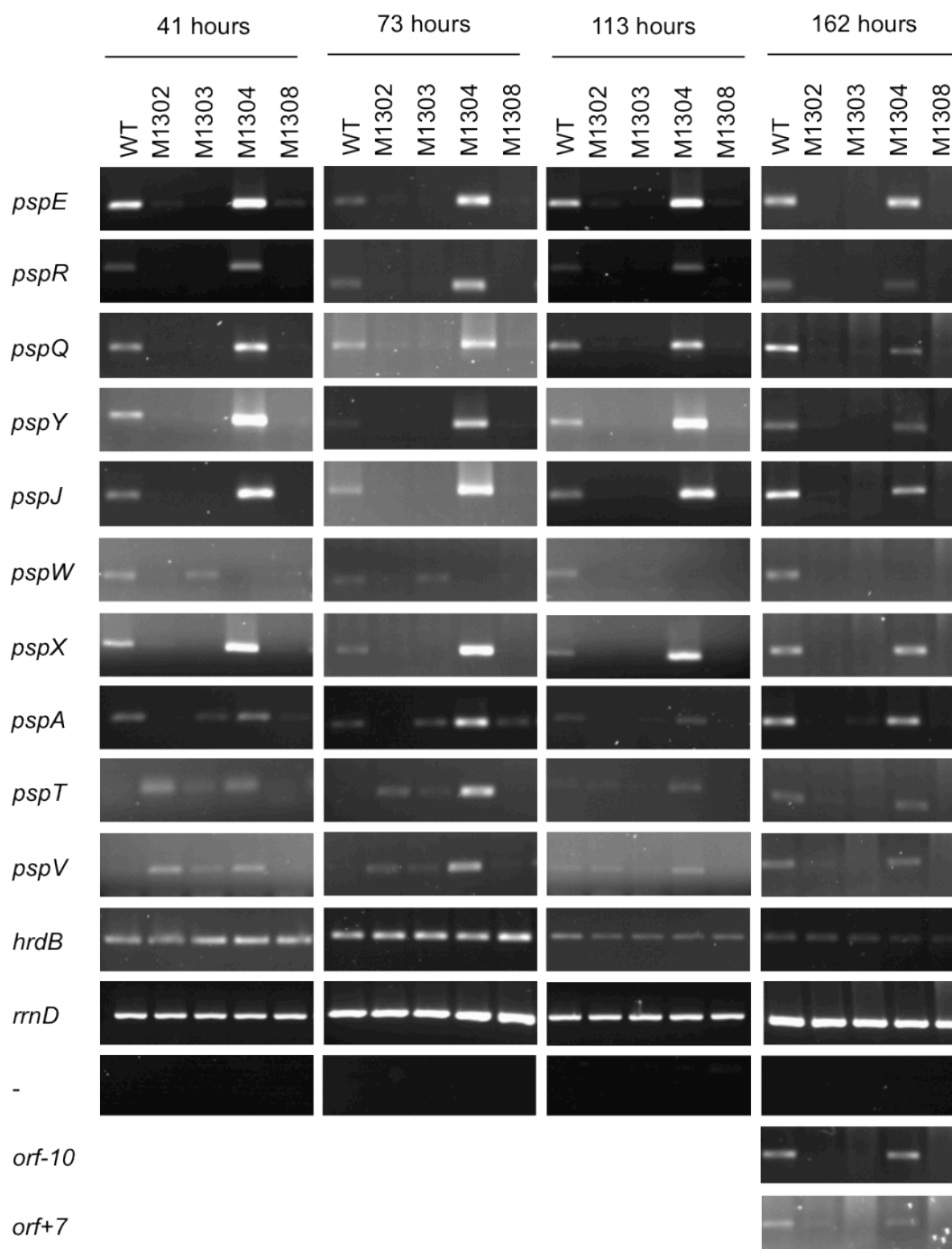


Figure 8.13 : RT-PCR on four *P. alba* strains.

Namely M1302 ($\Delta pspA::hyg$), M1303 ($\Delta pspX::hyg$), M1304 ($\Delta pspW::hyg$), M1308 ($\Delta pspR::hyg$) and WT. RNA extracted from each strain at four different time points was used as a template for cDNA synthesis. Primers annealing within *pspE*, *pspR*, *pspQ*, *pspY*, *pspJ*, *pspW*, *pspX*, *pspA*, *pspT* and *pspV* were used to assess the level of transcription of these genes at each time point. Primers annealing within *hrdB* and *rrnD* were used as positive controls. RNA from cDNA synthesis reactions lacking the RT enzyme was used as a template with primers for *hrdB* in a negative control reaction (marked '-'). Primers annealing within *orf-10* and *orf+7* were used to assess the transcription of genes flanking the *psp* cluster. PCR products were run on a 2 % agarose gel by electrophoresis.

As anticipated, the *pspA* transcript was completely absent from strain M1302, and appeared to decrease from a level similar to the WT at 41 and 73 hours, to a level lower than the WT by 113 and 162 hours in both strain M1303 and M1308. Interestingly, M1302 also expresses *pspT* and *pspV*. It would appear that the introduction of the pIJ10700 cassette in place of *pspA* does not have a polar effect on the transcription of the rest of the *pspABCTUV* operon. On the contrary, the promoter in front of the hygromycin resistance gene in the pIJ10700 cassette may be driving higher levels of expression than in the WT strain. In addition, *psp* transcript levels were markedly reduced in strains M1303 and M1308. Reassuringly, the *pspX* transcript was not detected in strain M1303, and likewise M1308 did not contain a detectable level of *pspR* transcription, while transcription of *pspW* was detected in M1303, indicating that the introduction of the pIJ10700 cassette in place of *pspX* did not completely abolish *pspW* expression. Overall, it is apparent that *pspA*, *pspX* and *pspR* are all required for efficient transcription of most of the *psp* gene cluster.

It is interesting to note that expression of *orf-10* and *orf+7* also appear to be coordinately regulated with the rest of the cluster. After 165 hours incubation expression is detectable in WT and M1308 but not in M1302, M1303 and M1308 (Figure 8.13). However an essential role in planosporicin production seems unlikely; Chapter 6 demonstrated that the cluster of 15 genes extending from *pspE* to *pspV* was sufficient for planosporicin production in *Nonomuraea*.

8.5.2 Quantitative (q) PCR

RT-PCR, which is at best semi-quantitative, had indicated that levels of *psp* transcription were similarly and markedly reduced in both M1303 and M1308 (Section 8.5.1), indicating that *pspX* and *pspR*, respectively, were both essential for expression of the *psp* cluster. However, a more quantitative analysis might reveal differences in expression levels between M1303 and M1308 that could indicate specific roles for each regulatory protein. Consequently, quantitation of *psp* gene expression after 73 and 162 hours of growth was determined by Quantitative (q) PCR.

Primers were selected which annealed to representatives from each operon predicted in Section 8.4. Table 2.5 of Chapter 2 lists the primers used in RT-PCR and Section 2.20.6 describes the qPCR protocol. Products were detected using the fluorogenic dye SYBR Green that upon excitation emits a strong fluorescent signal when bound to double-stranded (ds)DNA. However, as SYBR Green will bind to any dsDNA in the reaction, follow up assays were used to ensure only the desired product was present. Melting curve analysis was used to measure the temperature-dependent dissociation between DNA

strands through the resultant large reduction in SYBR green fluorescence. The melting curves from each plate confirmed that just one product was amplified with each primer pair. Likewise, samples checked by gel electrophoresis revealed one band of the expected size for each primer pair. The quantitation graphs revealed that the three independent qPCR analyses of each RNA sample took off after similar numbers of qPCR cycles. As well as cDNA templates, each primer pair was used in reactions with different dilutions of *P. alba* genomic DNA to generate a standard curve onto which a linear regression line was plotted. The R^2 value is a measure of goodness-of-fit of linear regression (where an R^2 of 1.0 means all points lie exactly on a straight line with no scatter). The regression line for all standard curves had an R^2 value >0.98 , indicating the regression equation can be reliably used to convert threshold cycle (Ct) values into transcript copy number. Primer efficiency varied from 85 – 95 % for all primer pairs used. Occasionally some contamination with gDNA was observed in the blank controls (no template), however this was at a very low level, only detectable after >35 cycles.

The raw Ct values were averaged across the triplicates to get a mean value which was converted into the copy number of cDNA in each sample using the appropriate regression equation. These values were normalised to the expression of the endogenous control gene, *hrdB*, in each sample to take account of differing cDNA concentrations in each sample. Initial analysis of this data revealed that the level of expression within an operon varied, whereas it would be expected to be consistent for all constituent genes. The one exception would be if an attenuator is present, for example after *pspA*. This unexpected variation is likely due to variable amplification efficiencies using different primer pairs during qPCR. To counteract this observed variability, the expression level of each gene in the different mutants was expressed as a percentage of the WT expression level.

Figure 8.14 shows the level of transcription of 11 genes from the *psp* gene cluster in four *P. alba* deletion mutants. The increased expression of all *psp* genes in *P. alba* M1304 is immediately apparent. After 73 hours of growth, WT *P. alba* had not yet started to produce planosporicin, while M1304 exhibited 5.1 - 37.5 fold increases in expression of 10 *psp* genes (*pspW* is deleted from the genome so no expression is detected). These included representatives from each of the five proposed operons. By 162 hours of growth, planosporicin production had commenced in the WT strain and so the increase in transcription in M1304 was reduced compared to earlier time points, but was still 1.1 - 4.5 fold higher than WT levels. The increased expression of *pspT* and *pspV* in M1302 compared to M1303 and M1308 may be due to transcription initiation at the promoter of the hygromycin resistance gene used to replace *pspA*. All of the negative control samples gave values comparable to the water only control, demonstrating that the RNA samples were not contaminated with genomic DNA.

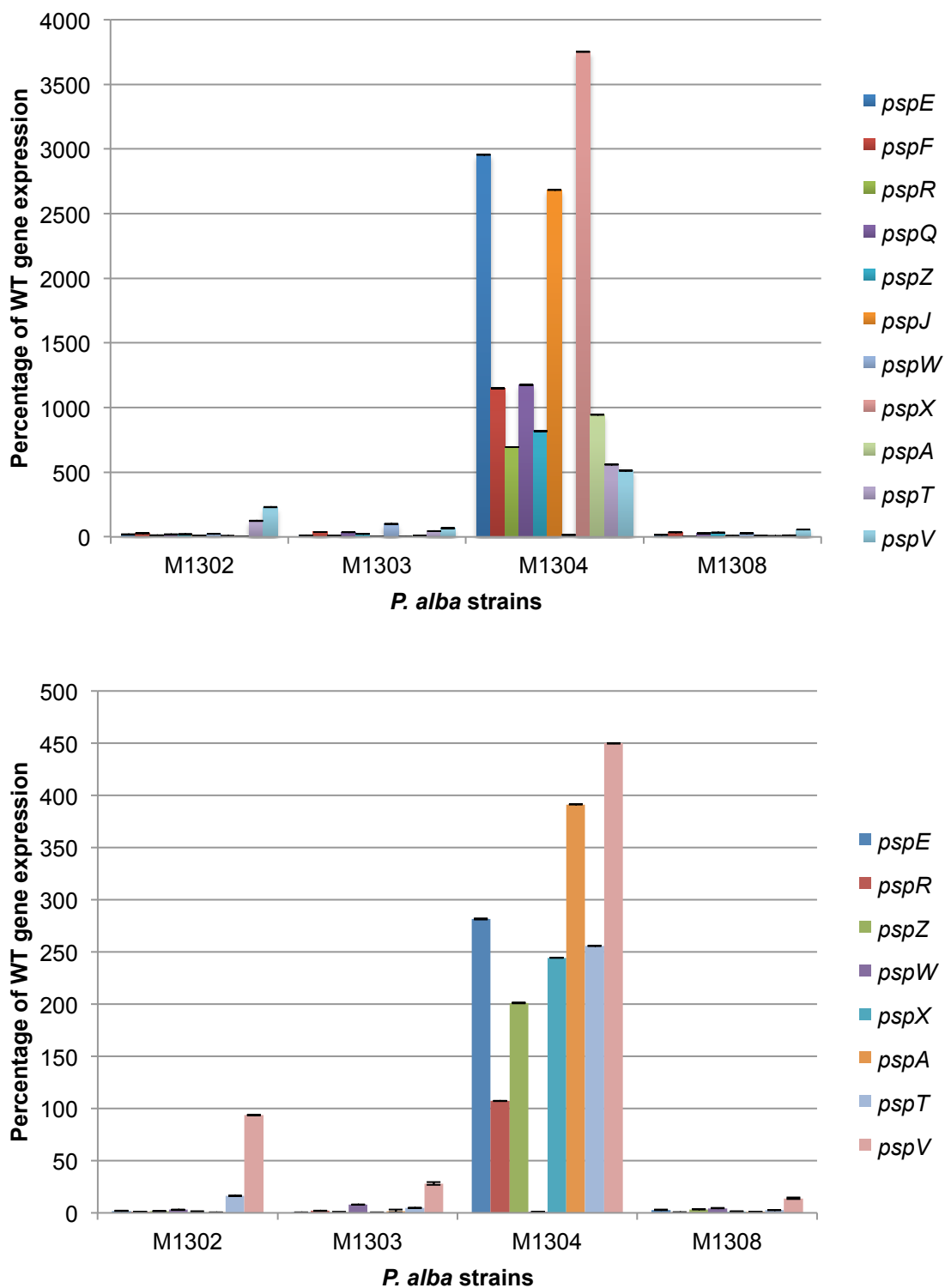


Figure 8.14 : Level of transcription from 11 genes of the *psp* gene cluster in four different *P. alba* strains at two different time points.

cDNA was synthesised from RNA extracted from *P. alba* M1302, M1303, M1304, M1308 and WT after 73 hours (top) and 162 hours (bottom). qPCR used primers which annealed to 11 different genes in the *psp* gene cluster. Primers for *hrdB* were used as a control. The copy number of each *psp* gene was normalised against the *hrdB* control then expressed as a percentage of WT expression. Standard error is displayed in error bars.

In the *psp* cluster there are two positively acting regulatory genes; *pspX* encodes an ECF sigma factor and *pspR* a transcriptional regulator. To analyse their individual roles in regulation, the expression of *psp* genes was compared in *P. alba* strains M1303 ($\Delta pspX$) and M1308 ($\Delta pspR$). Genes markedly down-regulated in one mutant compared to the other might imply direct regulation. Figure 8.15 shows that after 73 hours growth *pspJ*, *pspZ* and *pspE* were down-regulated 3.12-fold, 1.53-fold and 1.36-fold in M1303 compared to M1308, potentially revealing direct regulation by σ^{PspX} . In contrast, *pspA*, *pspT* and *pspV* were down-regulated 2.17-fold, 3.57-fold and 1.18-fold in M1308 compared to M1303, potentially revealing direct regulation by PspR. The same trend is observed after 162 hours growth, with the difference in expression levels even more pronounced with 5.89-fold lower *pspZ* expression and 9.32-fold lower *pspE* expression in M1303 compared to M1308. In addition, *pspW* is also up-regulated in M1303 compared to M1308. However this is likely to be read-through from the promoter of the hygromycin resistance gene in the pIJ10700 cassette used to replace *pspX* in the construction of M1303.

PspX and PspR are likely to directly regulate different *psp* genes independently of each other. If they were absolutely dependent on the presence of the other protein, then expression levels would be the same in both mutants. Therefore, both regulators presumably act independently. This analysis can be summarised to form a model for the regulation of planosporicin biosynthesis (Figure 8.16). An unknown signal activates *pspR* transcription at an unidentified promoter to activate *pspABCTUV* transcription and planosporicin production. Once planosporicin is produced, σ^{PspX} is released from PspW to drive high the level expression of *pspEF*, *pspJYZQ*, *pspXW* and *pspR* (through a second promoter). Thus PspR is the master regulator and σ^{PspX} -PspW form an ECF sigma factor–anti sigma factor complex that responds to sub-inhibitory concentrations of planosporicin. In the absence of PspR (strain M1308), *pspA* expression is decreased, which inhibits *pspX* induction, consequently decreasing expression of σ^{PspX} -regulated genes (*pspJ*, *pspZ*, *pspE*) but not to the same level as when σ^{PspX} is completely absent. This implies there is a basal level of *pspXW* expression even in the absence of planosporicin to trigger the cycle.

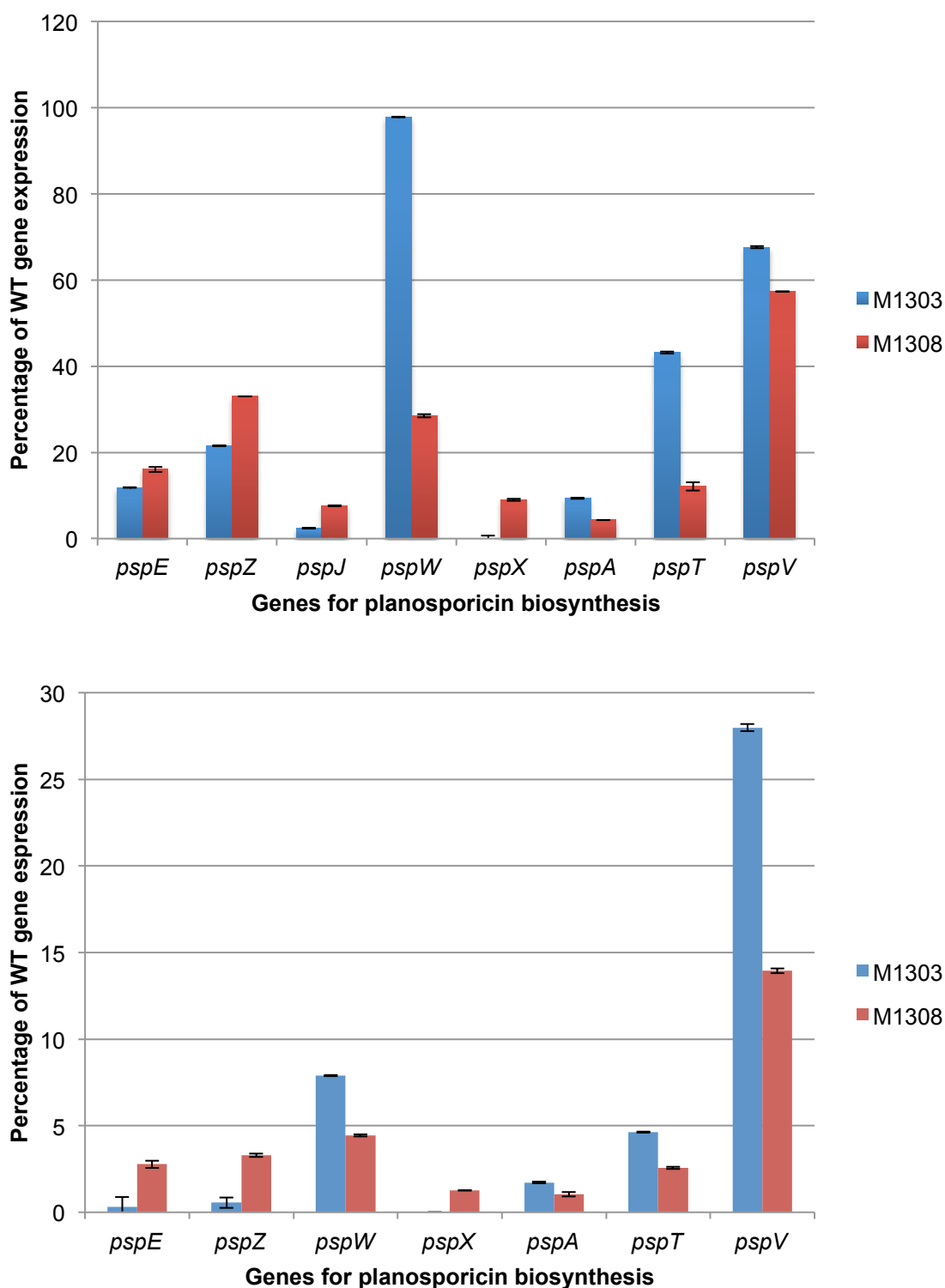


Figure 8.15 : Level of transcription from 7 genes of the *psp* gene cluster in two different *P. alba* strains at two different time points.

cDNA was synthesised from RNA extracted from *P. alba* M1303, M1308 and WT after 73 hours (top) and 162 hours (bottom). qPCR used primers which annealed to 7 different genes in the *psp* gene cluster. Primers for *hrdB* were used as a control. The copy number of each *psp* gene was normalised against the *hrdB* control then expressed as a percentage of WT expression. Standard error is displayed in error bars.

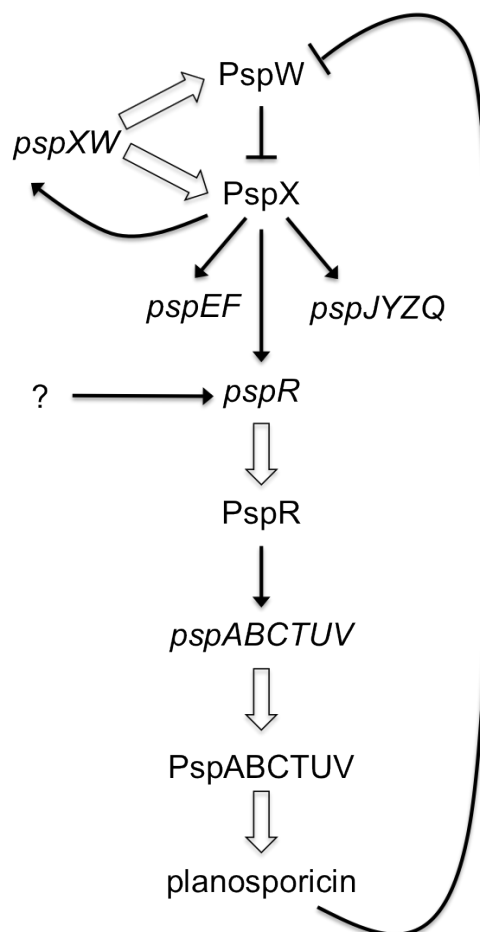


Figure 8.16 : Model for the regulation of the *psp* gene cluster.

Block arrows represent transcription and translation. Black arrows represent activation, whereas inhibition is indicated by a blocked line.

8.6 Discussion

The bioinformatic analysis of the *psp* cluster in Chapter 5 proposed the ECF sigma factor encoded by *pspX* and the transcriptional regulator encoded by *pspR* as a potential regulators of *psp* gene expression. The lack of planosporicin production in deletion mutants of *pspX* and *pspR* in Section 8.2 demonstrates they have an essential role in planosporicin biosynthesis. This is further substantiated in Section 8.5 in which PspX and PspR are found to have distinct roles in the regulation of different *psp* genes.

The results presented in this Chapter suggest a model for the regulation of planosporicin biosynthesis in which PspR functions as a master regulator to promote a sub-inhibitory level of planosporicin production. This induces a feed-forward mechanism mediated through σ^{PspX} that results in high level planosporicin production (Figure 8.16). In the young *P. alba* mycelium, a basal level of expression of *pspXW* results in a small amount of σ^{PspX} sequestered in the membrane by the PspW anti-sigma factor. Transcription of *pspR* mediated through a growth rate-dependent promoter is activated through an as-yet

unknown signal. This results in expression of the *pspABCTUV* operon and low-level planosporicin production which is likely exported through PspTU. Extracellular planosporicin interacts either directly with PspW, or inhibits peptidoglycan biosynthesis at a low level which is perceived by PspW. Inactivation of the anti-sigma factor causes the release of σ^{PspX} which forms a holoenzyme with RNA polymerase to drive high level expression of the entire *psp* gene cluster. This includes *pspEF* which encode an ABC transporter predicted to confer immunity to planosporicin.

A strikingly similar regulatory system is proposed for the microbisporicin gene cluster of *M. corallina* (Foulston and Bibb 2011). As described in Chapter 5, the *psp* gene cluster bears a strong resemblance to the *mib* gene cluster, not only on the basis of percent identity between proteins but also of synteny within the cluster. Foulston and Bibb proposed the ECF sigma factor, MibX, and the transcriptional regulator, MibR, as the key activators of *mib* gene transcription in a feed-forward mechanism. This was based on qPCR data, bacterial two- hybrid experiments (demonstrating that σ^{MibX} and the anti-sigma factor MibW interact strongly in *E. coli*) and the identification of the transcriptional start sites of each *mib* operon. MibR was proposed as the master regulator that induces the low level expression of the major microbisporicin biosynthetic operon through activation of an unidentified promoter triggered by an unknown signal. This would result in the production of microbisporicin without the chlorination and dihydroxylation modifications. Interaction of this less active form of microbisporicin with MibW would result in inactivation of MibW, release of MibX and high level expression of the entire *mib* cluster, resulting in the production of fully modified microbisporicin.

The high level of similarity between the *mib* and *psp* clusters, combined with the qPCR data from M1303 and M1308 suggests that planosporicin production is regulated in a similar manner. To support this model the suggested σ^{PspX} binding sites would need to be corroborated by identifying the transcriptional start site of each *psp* operon by S1 nuclease protection assays. Bacterial two-hybrid experiments could also be used to confirm an interaction between σ^{PspX} and PspW. Gel-shift experiments could be used to indicate the DNA binding site of PspR, which is presumed to lie just upstream of *pspA*. It is interesting to note that the ECF binding motif at -10 of four of the *psp* promoters is conserved as 'GCTAC' in *pspE* and *pspJ* but is only GXXXC in *pspX* and *pspR*, the two regulators. This may reflect a difference in expression level, indicating that expression is focused on biosynthetic enzymes rather than further expression of regulators.

According to this model of *psp* regulation, *pspR* must have two promoters; one recognised by σ^{PspX} and another triggering initial expression of the cluster. It is possible that this second promoter is located a considerable distance upstream, and was not cloned into pSET152 preventing successful complementation of strain M1308.

8.7 Summary

The targeted deletion of *pspX* marked with the pIJ10700 cassette abolished production of planosporicin as determined by bioassay and MALDI-ToF analysis. This mutant was complemented in *trans* with WT *pspX* under control of the native promoter.

The targeted deletion of *pspW* marked with the pIJ10700 cassette markedly increased production of planosporicin as determined by bioassay and MALDI-ToF analysis.

The targeted deletion of *pspR* marked with the pIJ10700 cassette abolished production of planosporicin as determined by bioassay and MALDI-ToF analysis. It was not possible to complement this mutant in *trans* with WT *pspR* using 446 bp upstream of the *pspR* start site as a promoter.

Expression of *mibA* and *mibABCD* in WT *P. alba* did not result in the detection of microbisporicin-related peaks in MALDI-ToF analysis.

A hairpin loop is likely formed in RNA between *pspA* and *pspB*, forming a transcriptional attenuator.

An ECF sigma factor consensus binding motif 'GAACC' was identified in the -35 regions of all *psp* operons except *pspA*. The *pspE* and *pspJ* promoters also contained a conserved 'GCTAC' motif in the -10 region.

RT-PCR demonstrated that expression of the *psp* gene cluster was dependent on *pspA*, *pspX* and *pspR*.

qPCR demonstrated a 5 – 38 fold increase in transcription of the *psp* cluster in *P. alba* M1304 ($\Delta pspW::hyg$) compared to the WT after 3 days growth.

RT-PCR and qPCR demonstrate the differential control of operons within the *psp* cluster by the ECF sigma factor gene *pspX* and the transcriptional regulator gene *pspR*.

A regulatory model is proposed for the pathway-specific regulation of the planosporicin biosynthetic gene cluster by a master regulator (PspR) and an ECF sigma factor–anti sigma factor complex (σ^{PspX} -PspW) that responds to sub-inhibitory concentrations of planosporicin.

Chapter 9 : General discussion

9.1 Planosporicin structure and activity

Lantibiotics form a distinct class of antimicrobial peptides due to their unusual post-translational modifications. These modifications constrain the molecule into a defined structural conformation that provides both stability and biological activity.

Few three-dimensional solution structures have been solved for lantibiotics produced by members of the order Actinomycetales. Two examples are microbisporicin, a type A lantibiotic produced by *Microbispora* sp. (Castiglione *et al.* 2008) and actagardine, a type B lantibiotic produced by *Actinoplanes* sp. (Zimmermann and Jung 1997). The revisions made to the planosporicin structure (detailed in Chapter 3) changed its classification from type B to type A, and revealed significant N-terminal similarity to nisin-like lantibiotics.

A notable feature of lantibiotics is the presence of non-proteogenic amino acids. Against MRSA strain L1400, planosporicin has an MIC of 8 µg/ml, while microbisporicin has an MIC of less than 0.13 µg/ml (Castiglione *et al.* 2008). The lower MIC of microbisporicin may be due to the presence of additional modifications; a chlorinated tryptophan and dihydroxyproline. Yet a *M. corallina* knockout mutant unable to chlorinate tryptophan displayed an increased zone of inhibition in agar diffusion bioassays (Foulston and Bibb 2010). The formation of the inhibition zone is affected by a number of lantibiotic properties, including solubility, stability, and diffusion. Therefore a difference in zone size in an agar diffusion assay is not necessarily indicative of a difference in potency. For example the lantibiotics nisin A and nisin Z differ in just a single amino acid residue and have identical MICs when tested on a number of Gram-positive bacteria, but at concentrations above the MICs, nisin Z displays a larger zone of inhibition compared to nisin A (de Vos *et al.* 1993). These nisin variants have comparable solubility and stability, so it was concluded that nisin Z diffuses more rapidly than nisin A (de Vos *et al.* 1993). Thus it may be that the difference in charge distribution in the non-chlorinated form compared to the chlorinated form of microbisporicin results in an increase in the rate of diffusion in agar, not necessarily an increase in potency.

The rate of diffusion is not the only factor affecting the ease at which the compound reaches the target. Additional difficulties may be encountered when crossing the thick peptidoglycan layer due to interactions with polar, charged and hydrophobic moieties

present on the cell surface. An alternative explanation for the increased potency of microbisporicin compared to planosporicin lies in the structural differences at the C-terminus. Planosporicin is charged at its C-terminus due to the presence of the carboxyl group of Cys₂₄, while in microbisporicin the C-terminal cysteine is modified to form an AviCys modification that is not charged. A recent patent by Sentinella Pharmaceuticals, Inc. demonstrated that carboxamide derivatives of planosporicin have enhanced antibacterial activity (Maffioli *et al.* 2010). This patent describes how the carboxyl groups of Cys₂₄ and Glu₁₄ can be modified to remove the negative charge to create variants with increased potency. However this is not likely to be a strategy that will apply to all lantibiotics. In both the solution and the crystal structure of mersacidin, the N-terminus folds back to interact with the carboxyl group of the glutamic acid residue at position 17 giving a compact structure (Schneider *et al.* 2000). Site-directed mutagenesis exchanged Glu₁₇ for Ala₁₇ creating a derivative that was obtained in good yields but showed markedly reduced activity (Szekat *et al.* 2003). This confirmed that the biological activity of mersacidin relied on the carboxylic acid moiety at position 17, and thus this residue cannot be replaced to reduce overall negative charge with the possibility of increased potency. This may apply to other lantibiotics with a mersacidin-like lipid II binding motif; for example the site-directed mutagenesis of the equivalent Glu residue in haloduracin abolished bioactivity (Cooper *et al.* 2008).

9.2 Planosporicin mode of action

Information from a number of indirect experiments implies that planosporicin inhibits cell wall biosynthesis. Planosporicin severely inhibits peptidoglycan synthesis over that of other macromolecules, and planosporicin treatment of growing bacterial cells result in the accumulation of UDP-linked peptidoglycan precursors (Castiglione *et al.* 2007). Planosporicin also shows activity against VRE so is proposed to bind lipid II through a moiety other than the D-Ala-D-Ala pentapeptide (Castiglione *et al.* 2007).

Thus the classification of planosporicin as a type A, nisin-like molecule (as described in Chapter 3) allows inference into its likely mode of action. The N-terminal 11 residues of planosporicin are identical to those of nisin (except for positions 4 and 6). Homology modelling of epidermin and mutacin on the structure of nisin bound to lipid II revealed that the A and B rings can accommodate a variety of amino acid side-chains apart from those at conserved positions 3 and 7-11 (Hsu *et al.* 2004). Planosporicin is conserved in these regions, indicating that it too is likely to be able to form a cage around the pyrophosphate moiety of lipid II.

However the presence of nisin-like A and B rings is not a pre-requisite for lipid II binding. Mersacidin and actagardine are type B lantibiotics with a globular tertiary structure but they also bind to lipid II. It is thought that the 'CTLTXEC' motif in the C-ring of mersacidin and the B-ring of actagardine confers an ability to bind lipid II (Zimmermann and Jung 1997; Brotz *et al.* 1998). Lacticin 3147 and haloduracin are two-component lantibiotics with no similarity to nisin. The mature A1 component of lacticin 3147 and the mature halo of haloduracin also contain the 'CTLTXEC' motif in the B-ring and have been proposed to bind lipid II (Cooper *et al.* 2008; Zerikly and Challis 2009).

Thus it would appear that lantibiotics from distinct gene clusters have evolved convergently to act upon on the same target. Similarly the glycopeptide vancomycin also acts on lipid II, but through the sequestration of the D-Ala-D-Ala moiety (Barna and Williams 1984). In contrast the precise mechanism of action of other lipid II-targeting antibiotics such as the mannopeptimycins, ramoplanin, plusbacin A₃ and katanosin B, is not yet known (Fischbach 2009). Targeting lipid II is such an effective antibacterial strategy, it is also targeted by antimicrobials produced by other domains of life. Plectasin is a fungal defensin which also binds to the pyrophosphate moiety of lipid II (Schneider *et al.* 2010).

While a key functionality of planosporicin is likely to be the binding of lipid II by its N-terminus, it is likely that most lantibiotics possess multiple activities that combine differently for individual target strains. For example, nisin has a second mode of action; the formation of trans-membrane pores resulting in osmotic shock and cell lysis. This is mediated through its C-terminus. However binding to lipid II by the N-terminus is a key initial step (Hasper *et al.* 2004). Other type A lantibiotics that bind lipid II at the N-terminus vary in their ability to form pores through their C-termini. Epidermin and gallidermin are 20 Å shorter than nisin, limiting pore-formation to only the thinnest of membranes (Bonelli *et al.* 2006).

As yet there is no evidence for the formation of pores in the plasma membrane as a second mode of action for planosporicin. Indeed, the characteristic inhibition of DNA, RNA and proteins observed upon treatment of *S. aureus* with nisin is not observed with planosporicin (Castiglione *et al.* 2007). The mode of action of planosporicin was not investigated during this work due to perceived difficulties in distinguishing specific pore formation with stress-induced lysis (O'Neill *et al.* 2004). To date, there has been little work undertaken to investigate the precise mode of action of planosporicin. The ability to bind lipid II is implied from the strong similarity of the N-terminus to that of nisin.

The specific targeting of a key component of cell wall biosynthesis implies lantibiotics may yet find a niche use in the clinic. So far, only nisin has been used commercially, however

the observed low solubility at physiological pH prevents clinical use (Gowans *et al.* 1952). Instead nisin is widely used as a food preservative with E number E234. The success of other lantibiotics such as microbisporicin and actagardine in pre-clinical trials indicates they may soon enter Phase I clinical trials. Lantibiotics differ completely to the planar heterocycles currently used in the clinic, and may mark a new frontier in antimicrobials. One remaining concern is that treatment with a peptide may result in an immune response. However, there is evidence to suggest that the presence of (Me)Lan bridges prevents the presentation of these molecules to macrophages, preventing specific antibody formation (Hans-Jorg Sahl, personal communication). It is likely that lantibiotics will only ever be successful in treating Gram-positive infections due to the bulky (Me)Lan residues preventing transport across the membrane of Gram-negative bacteria. However there are already exceptions; a lantibiotic produced by *Bifidobacterium longum* has recently been patented due to claims that it successfully inhibits the growth of Gram-negative bacteria under conditions in which the outer membrane is intact (O'Sullivan and Lee 2011).

While many actinomycetes encode lantipeptide gene clusters, very few produce a corresponding compound with antimicrobial activity. Assuming that these gene clusters are expressed at significant levels under laboratory conditions, these lantipeptides must presumably have other biological functions. The vast discrepancy between the concentrations used to demonstrate antimicrobial activity and the concentration present in the soil implies that even lantibiotics that do exhibit antimicrobial activity when tested under laboratory conditions may have an entirely different role in the natural environment, for example as signalling molecules (Goh *et al.* 2002; Shank and Kolter 2009).

9.3 The *psp* gene cluster

9.3.1 Introduction of non-proteogenic amino acids

The proteins PspB and PspC were assigned functions on the basis of their significant similarity to proteins whose function had been determined experimentally. PspB is likely the dehydratase responsible for the dehydration of five serines and two threonines to Dha and Dhb, respectively, while PspC is likely the cyclase responsible for the coupling of a cysteine to five of these dehydrated residues to form four Lans and one MeLan (Figure 9.1). So far, *lanB* and *lanC* genes have only been found in biosynthetic operons encoding lantibiotics. Conversely, all lantibiotics are biosynthesised by one of four characterised pathways; LanB/C, LanM, RamC and LanL (as described in Chapter 1). Until recently, one exception was thought to be the sublancin 168 biosynthetic gene cluster from *B. subtilis* 168. This lantibiotic was originally characterised as containing four bridges; one Lan, one

MeLan and two disulphides (Paik *et al.* 1998). However the sublancin 168 biosynthetic gene cluster encodes no recognisable lanthionine synthases and only BdBB, a thiol-disulfide oxidoreductase, and SunT, an ABC transporter were shown to be essential for production (Dorenbos *et al.* 2002). Recently, the structure of sublancin was revised as tandem mass spectrometry and Q-ToF revealed there were no dehydrated residues or (Me)Lan bridges and therefore sublancin is not a lantibiotic (Oman *et al.* 2011). Instead two disulphide bridges and one additional modification, a glycosylation of Cys₂₂ was observed. An S-glycosyl transferase adds a glucose residue to a cysteine residue, therefore this antibiotic was reclassified from a lantibiotic to a glycopeptide (Oman *et al.* 2011).

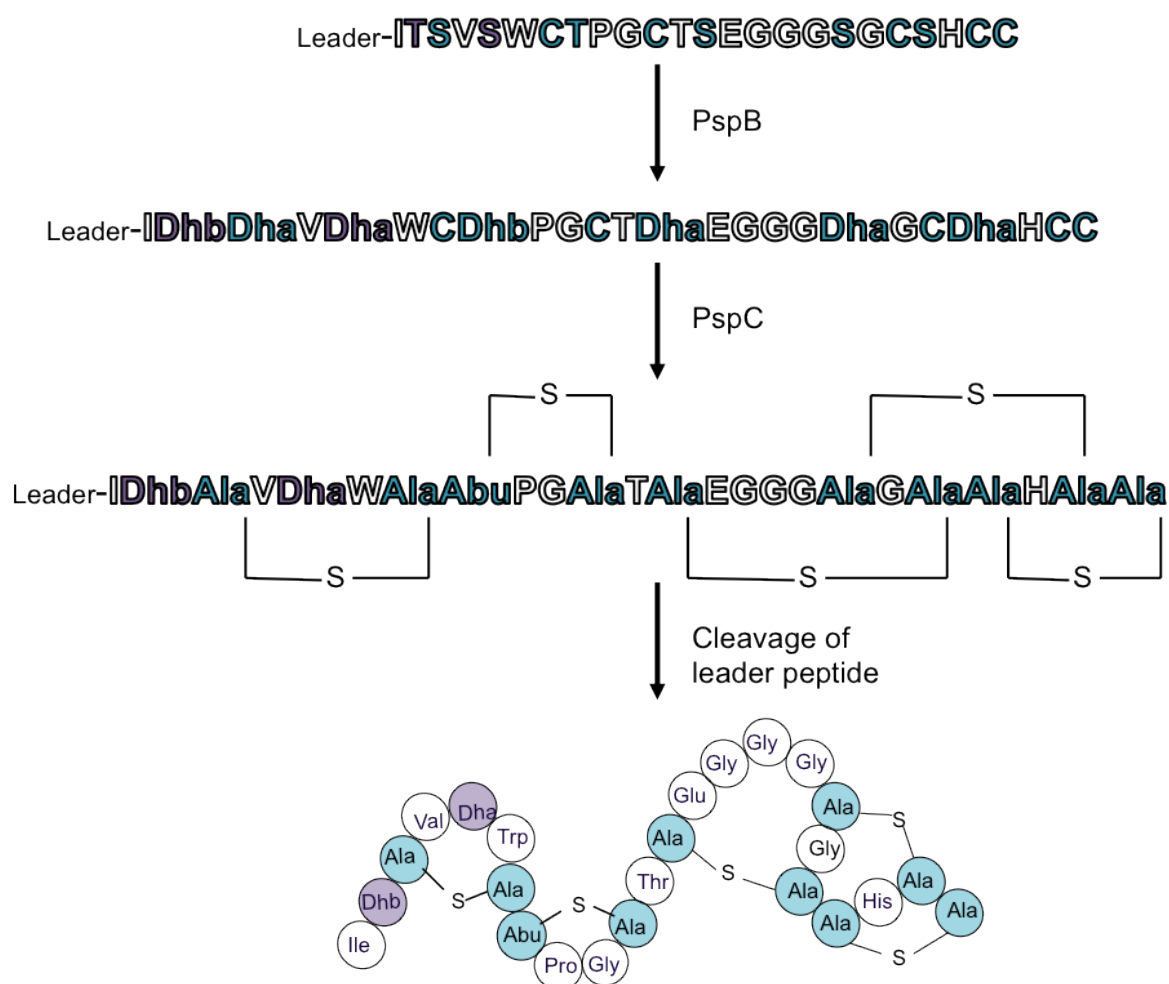


Figure 9.1 : Schematic representation of planosporicin post-translational modifications.

PspB catalyses the dehydration of Ser and Thr into Dha and Dhb, respectively. PspC installs the four Lan bridges and one MeLan bridge by regioselective cyclisation. Removal of the leader sequence (MGISSPALPQNTADLFQLDLEIGVEQSLASPA) by an unidentified protease liberates the active lantibiotic.

Over 50 % identity was observed between PspBC and MibBC. This is much higher than the <30 % identity commonly observed between LanBC enzymes and likely reflects the similar prepropeptide substrates of the *psp* and *mib* gene clusters due to a more recent common ancestral gene cluster. The 24 amino acid propeptides of planosporicin and microbisporicin are dehydrated in identical positions; 2, 3, 5, 8, 13, 18 and 21, and four out of the five (Me)Lan bridges in planosporicin are also present in microbisporicin. An even higher percent identity would likely be observed if the PspA and MibA leader peptides were identical (they show 43 % end-to-end identity), as is the case for subtilin and ericin S prepropeptides from *B. subtilis*. In these biosynthetic clusters EriB and SpaB share 83 % identity while EriC and SpaC share 80 % identity (Stein *et al.* 2002a). Consequently the EriC enzyme is able to complement a SpaC deletion mutant, restoring production of mature subtilin (Helfrich *et al.* 2007).

9.3.2 ATP-binding cassette transporters

All lantibiotic gene clusters encode at least one ABC transporter system proven or presumed to be responsible for transport of the lantibiotic out of the cell, and many also encode a second as an immunity mechanism (Chatterjee *et al.* 2005). Interestingly, the architecture of the transporter responsible for lantibiotic export differs with the producing organism. The NisT of low-GC *Lactococcus* contains fused domains while the MibTU of high-GC *Microbispora* encodes the nucleotide-binding domain (NBD) and transmembrane domain (TMD) domains on separate proteins. However, the dependence on these transporters for export of the lantibiotic differs on a case-by-case basis.

There are several examples in which deletion of the transporter genes had no effect on lantibiotic production, presumably due to complementation by other transporters located in the same gene cluster or elsewhere on the chromosome. The single-component transporters SpaT and PepT, encoded in the subtilin and Pep5 gene clusters, respectively, were each found to be non-essential for production of the respective lantibiotic (Klein *et al.* 1992; Meyer *et al.* 1995). However there was some indication that *spaT* mutants accumulated subtilin intracellularly, implying that it does play a role in transport (Klein *et al.* 1992). Likewise, the deletion of PepT abolished 90 % of Pep5 production, indicating that it can only be partially replaced by other host-encoded translocators (Meyer *et al.* 1995). Epicidin 280 produced by *Staphylococcus epidermidis* BN 280 exhibits 75 % identity to Pep5 produced by *S. epidermidis* 5. Interestingly, the epicidin 280 gene cluster encodes proteins with significant sequence similarity to those of the Pep5 cluster, but lacks a homologue of the PepT ABC transporter (Heidrich *et al.*

1998). The absence of an ABC transporter in the gene cluster implies export occurs through a transporter encoded elsewhere in genome.

A similar phenotype has been observed in high-GC bacteria in which the two-component transporter responsible for lantibiotic export has been deleted. MibTU is predicted to be involved in the export of microbisporicin from *M. corallina*. Deletion of *mibTU* did not abolish microbisporicin production and so it was proposed that other transporters encoded by the *mib* gene cluster (MibYZ, MibEF and/or MibN) could compensate for this deletion (Foulston and Bibb 2010). Likewise, the actagardine gene cluster encodes GarTH, an ABC transporter thought to be involved in lantibiotic export (Boakes *et al.* 2009). Heterologous production of actagardine in *S. lividans* was not affected by the deletion of these genes (Bell 2010).

Equally, there is precedence for the deletion of ABC transporter genes preventing lantibiotic production. NisT is absolutely required for export of nisin; after deletion of *nisT* from *L. lactis*, nisin activity could only be detected if cells were lysed (Qiao and Saris 1996). Likewise *S. cinnamoneus* encodes the two-component ABC transporter CinTH. Deletion of the corresponding genes created a strain unable to produce cinnamycin (S. O'Rourke, personal communication). The ABC transporter PspTU of *P. alba* exhibits a similar phenotype; the production of planosporicin is completely abolished in a *pspTU* deletion mutant. However as only one mutant was isolated and could not be complemented, the existence of a second-site mutation cannot be ruled out.

The ABC transporter PspEF was proposed to function as an immunity mechanism based on its significant alignment to other lantibiotic immunity transporters. In this work, time restrictions prevented deletion analysis of this two-component ABC transporter. Work on immunity mechanisms for other lantibiotics has confirmed that LanFEG transporters confer resistance. For example, deletion of *spaEFG* from *B. subtilis* and of *nisFEG* from *L. lactis* 6F3 increased susceptibility to nisin and subtilin, respectively (Klein and Entian 1994; Siegers and Entian 1995). However viability was maintained through the continued expression of a second immunity mechanism, mediated through a lipoprotein. There does not appear to be a link between this reduction in resistance and compound production, as lantibiotic biosynthesis continued at near-WT levels in both deletion mutants. However, in some cases there appears to be a link with the regulation of lantibiotic production; in-frame deletion of the lipoprotein gene *nisI* resulted in a marked reduction not only in immunity levels but also in the production of nisin (Ra *et al.* 1999).

Many lantibiotic gene clusters rely solely on a LanFEG transporter for immunity, such as the epidermin biosynthetic gene cluster of *S. epidermidis* Tu3298. The *psp* gene cluster does not encode an obvious secondary mechanism of immunity, although PspQ may be a

lipoprotein and PspYZ is an ABC transporter of unknown function. Likewise the microbisporicin gene cluster of *M. corallina* encodes the immunity transporter MibFEG, but the lipoprotein MibQ and unknown ABC transporter MibYZ may also play a role in immunity. Deletion of *mibEF* gave a phenotype in which microbisporicin was only detected at very low levels after prolonged growth, possibly through cell lysis (Foulston and Bibb 2010).

It could be anticipated that deletion of the immunity mechanism would leave the organism vulnerable to its own antibiotic. However there is some evidence that immunity genes may regulate export from the cell or biosynthesis in some way. Alternatively the cluster may encode a mechanism for sensing resistance levels which, when insufficient, prevents antibiotic export or production. This mechanism is likely to be located within the cluster, as heterologous expression of *epiFEG* in addition to the epidermin biosynthetic genes in *Staphylococcus carnosus* resulted in five-fold higher levels of epidermin production compared to *S. carnosus* bearing only *epiABCDQP* (Peschel and Gotz 1996).

The recently characterised microbisporicin gene cluster from *M. corallina* encodes a third ABC transporter, MibYZ, whose function is as yet unknown. Similarly, the planosporicin gene cluster also encodes three ABC transporters, all of which are encoded as separate TMD and NBD proteins. Homologues of *pspYZ* have been observed in other gene clusters as part of a cluster encoding homologs of *pspXWJ*. However *pspYZ* is the first to undergo deletion analysis, revealing severe reduction in the amount of planosporicin production. Interestingly, on solid media, production begins before the WT but never accumulates to a very high level, implying this ABC transporter may function together with PspXW in a mechanism of regulation as discussed in the next Section.

9.3.3 Regulatory genes

The regulatory mechanism coordinating the expression of the planosporicin gene cluster appears dissimilar to most other antibiotic gene clusters. As described in Chapter 1, most low-GC type AI antibiotic gene clusters encode a sensor histidine kinase and a response regulator, forming a two-component system. One well characterised example is the NisRK system, which is induced through a basal level of nisin expression detected at the cell membrane by NisK, which phosphorylates NisR, which in turn activates high level expression of the nisin gene cluster to generate high level production (Kuipers *et al.* 1995; de Ruyter *et al.* 1996). This low level production of the antibiotic may be a common mechanism through which gene cluster expression is induced, as many other clusters also encode a two-component system (Stein *et al.* 2002b; Schmitz *et al.* 2006).

The planosporicin and microbisporicin high-GC type AI lantibiotic gene clusters do not encode a two-component system. Instead there is pair of genes encoding an ECF sigma factor and an anti-sigma factor (σ^{PspX} -PspW) and a helix-turn-helix DNA binding protein (PspR). Chapter 8 described the deletion analysis of *pspX*, *pspW* and *pspR* which revealed that all play a role in planosporicin production, while RT-PCR and qPCR experiments revealed that all are important for *psp* gene expression.

ECF sigma factors are known to play a role in the regulation of genes involved in antimicrobial resistance as part of the general cell envelope stress response. For example in *B. subtilis*, σ^{W} and σ^{X} respond to cell envelope stress to regulate genes that include those conferring antibiotic resistance (Cao and Helmann 2004; Butcher and Helmann 2006). More recently, it has also been shown that ECF sigma factors contribute to antibiotic production. Sublancin 168 is an antimicrobial produced by *B. subtilis* 168 strains that are lysogenic for the SP β phage (Luo and Helmann 2009). As well as sublancin resistance in nonlysogens being controlled by σ^{W} , production in lysogens is dependent on σ^{X} and σ^{M} . However the activation of sublancin production occurs indirectly through the upregulation of the transition state regulator Abh (Luo and Helmann 2009). The first example of the direct regulation of an antibiotic gene cluster was the regulation of microbisporicin biosynthesis by MibX (Foulston and Bibb 2011).

Deletion of the ECF sigma factor *pspX* created a strain in which no planosporicin production was detected. This strain was complemented through the *in trans* expression of *pspX* from its own promoter to restore production, although it did not reach WT levels. In contrast, deletion of the anti-sigma factor *pspW* created a strain in which planosporicin production occurred earlier and at more prolific levels than the WT. On this basis, it appears σ^{PspX} is a positive regulator and PspW is a negative regulator of the *psp* gene cluster. The reduction in planosporicin produced from the complemented *pspX* mutant might reflect an altered level of gene expression at the presumed ΦC31 integration site compared to the *psp* cluster. Consequently the expression of *pspX* would be uncoupled from the expression of *pspW* potentially unbalancing the ratio of these proteins, which may well have an adverse effect on production levels. Similar results have been observed previously in the *mibX* mutant of *M. corallina*. Interestingly, *mibX* was better complemented on agar medium than in liquid (Foulston 2010). It was proposed that the agar prevents microbisporicin from diffusing far away, resulting in an increased local concentration around the *M. corallina* mycelium, triggering increased microbisporicin production in a feed-forward mechanism (Foulston and Bibb 2011).

Deletion of the transcriptional regulator *pspR* created another strain in which no planosporicin production was detected. Similar results were also observed in the *mibR* mutant of *M. corallina*, and this phenotype was complemented back to WT through the *in*

trans expression of *mibR* from its own promoter (Foulston and Bibb 2011). Thus it would appear that PspR plays a role as another positive regulator of the *psp* gene cluster.

PspR and MibR are not the only LuxR-family proteins that regulate an antibiotic biosynthetic gene cluster without the need for a cognate signal generator. The *Serratia marcescens* ATCC 39006 *carR* gene, encodes a LuxR-like protein which functions independently of a specific pheromone signal resulting in constitutive activation of carbapenem synthesis in *S. marcescens* (Cox *et al.* 1998). It has been proposed that these unpaired LuxR-family proteins could allow bacteria to respond to signals produced by different neighbouring species, which would not be observed when they are grown in laboratory monocultures (Subramoni and Venturi 2009). Thus PspR may interpret signals from other species or even kingdoms to regulate planosporicin production.

9.3.4 Unknown genes

Bioinformatic analysis revealed PspV to be a hypothetical protein with no defined function. The deletion of its closest homologue, *mibV*, from *M. corallina*, prevented chlorination of the tryptophan at position four of microbisporicin (Foulston and Bibb 2010). As planosporicin has a valine at this position and no chlorinated residues elsewhere on the peptide, PspV is likely to play a broader role. This is consistent with the observation that PspV-like proteins are encoded in both type AI (MibV) and type AII/B (Sven_0538/0540) lantibiotic gene clusters. One possibility is as a scaffold or chaperone to aid the binding of modification and processing machinery to the prepropeptide.

Likewise, PspJ and PspQ are also hypothetical proteins. Time limitations prevented deletion analysis of these genes so conclusions can only be drawn from bioinformatics analysis. PspJ is predicted to contain TM helices, while PspQ has a predicted signal peptide sequence containing the conserved lipobox motif of lipoproteins. Homologues to PspQ occur in the microbisporicin gene cluster of *M. corallina* (MibQ) and *Microbispora* sp. ATCC PTA-5024 (MlbQ). Heterologous expression of *mibQ*, *mlbQ* or *mlbQ* with *mlbEF* in *S. coelicolor* did not confer increased resistance to microbisporicin (J.P. Gomez-Escribano, personal communication). Strikingly, homologues to *pspJ* and *pspQ* always occur in a gene cluster that also contains *pspX*, *pspW*, *pspY* and *pspZ* homologues. This cluster is not always associated with a lantibiotic gene cluster, implying a more generic role. It could be that the PspJYZQ homologues play a role in the mechanism of ECF sigma factor regulation, whether that sigma factor has a pathway-specific role in one gene cluster or has a pleiotropic role in the transcriptional activation of a diverse regulon. The fact that the ECF sigma factor is a core component, implies these genes may regulate some aspect of cell wall physiology in response to cell wall stress.

9.4 The model of *psp* regulation

One can imagine that there is a low basal level of *pspXW* transcription, giving a constitutive presence of σ^{PspX} -PspW at the cell membrane. PspR appears to be a master regulator which induces low level planosporicin production to initiate the system through an unidentified promoter triggered by an unknown signal, possibly nutrient limitation. Interaction of planosporicin with PspW (or perhaps low level peptidoglycan biosynthesis inhibition that is perceived by PspW) disrupts the binding interaction with σ^{PspX} . The release of σ^{PspX} from PspW would allow formation of the RNA polymerase holoenzyme to drive the high level expression of the entire *psp* cluster in a feed-forward mechanism (Figure 9.2).

In contrast the majority of lantibiotic systems are regulated via a two-component system. For example nisin is able to autoregulate its own transcription, acting as a peptide pheromone for quorum sensing (Kleerebezem 2004). The response regulator NisR and the histidine kinase NisK are constitutively expressed (de Ruyter *et al.* 1996). Nisin induces the autophosphorylation of NisK, with subsequent phosphotransfer to NisR. The conformational change in NisR confers the ability to act as a transcriptional activator resulting in expression of nisin biosynthetic genes due to a nisin-responsive element in the promoters of *nisABTCIP* and *nisFEG* (Kleerebezem 2004). A parallel can thus be drawn between this system and that of planosporicin in which σ^{PspX} -PspW are analogous to NisRK in sensing increasing levels of planosporicin at the cell membrane. Both mechanisms provide a means for the lantibiotic to reach a threshold sub-inhibitory concentration, leading to high level production.

In addition to pathway-specific regulation by PspX and PspR, the *psp* cluster may also be regulated by a global regulator such as ppGpp. In streptomycetes this stringent factor is induced in response to amino acid starvation to favour transcription of genes for adaptation and survival over those for growth, essentially acting as a switch (Hesketh *et al.* 2007). The bioinformatics analysis in Chapter 5 revealed a TTA codon in *pspX*. In streptomycetes this codon is known to rely on the *bldA* encoded Leu-tRNA for translation (Leskiw *et al.* 1991). Although rare in high-GC organisms, the TTA codon is often present in pathway-specific regulatory genes for antibiotic biosynthesis (Leskiw *et al.* 1991). This implies that the initiation of planosporicin synthesis via PspX may be the point at which *bldA* exerts developmental control of production. Thus the TTA codon may link *P. alba* development to planosporicin production. However planosporicin production was observed on media in which *P. alba* does not sporulate, and a lack of planosporicin production was observed on media in which *P. alba* sporulates. Therefore the decision to sporulate and the decision to produce planosporicin occur through separate hubs fed by multiple signals and so can occur independently.

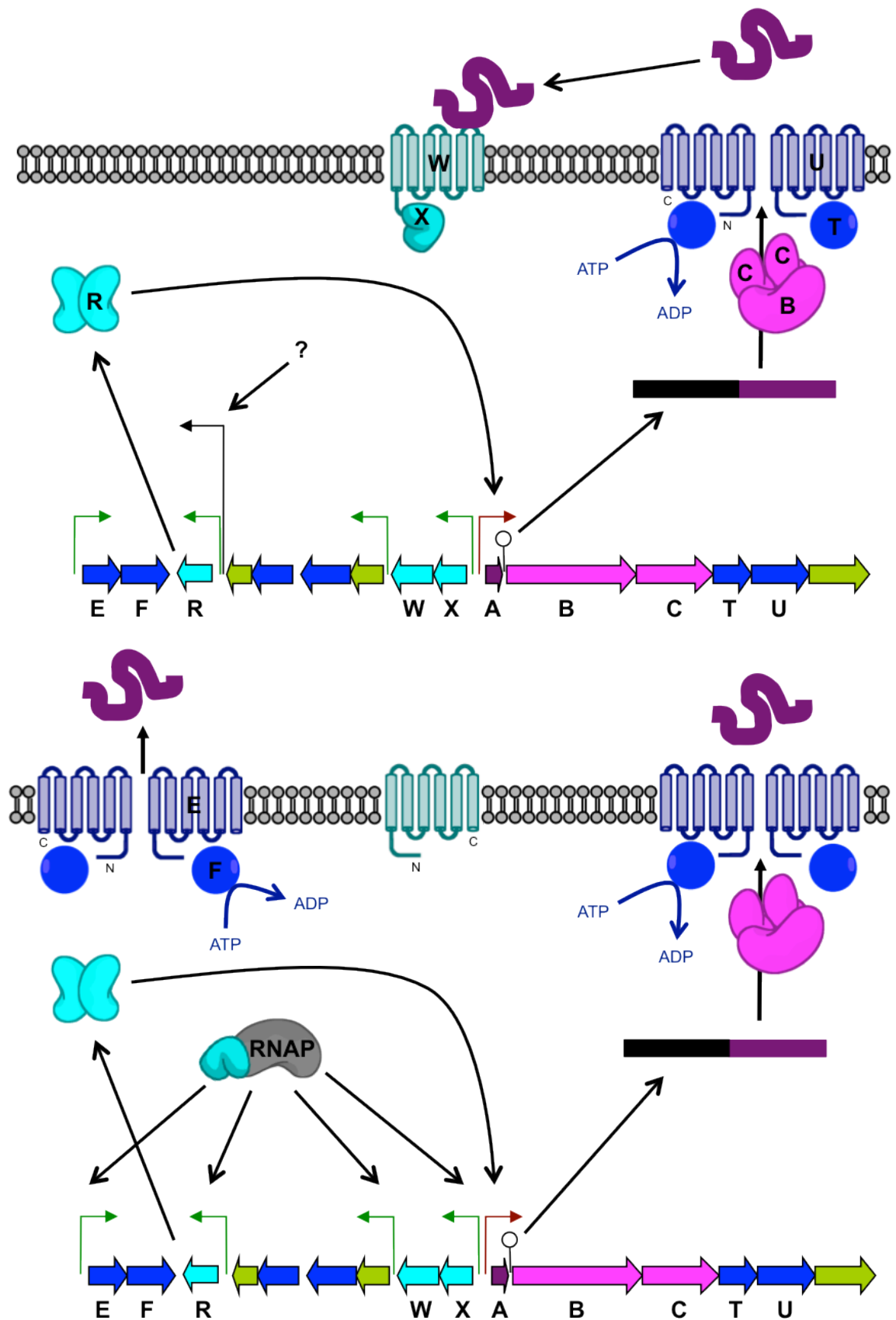


Figure 9.2 : A schematic depicting the model of regulation of planosporicin biosynthesis. The upper panel depicts the induction of transcription by the master regulator, PspR. The lower panel depicts the response to sub-inhibitory concentrations of planosporicin mediated through the σ^{PspX} -PspW complex leading to high level planosporicin production.

9.5 Future work

There remain a number of deletion mutants that would be desirable for a full understanding of the planosporicin biosynthetic gene cluster. Most notably the mechanism of immunity has not been investigated, and deletion of the putative immunity transporter PspEF and the possible lipoprotein PspQ would indicate if they are likely to play a role in self-protection. In parallel, the similarity between planosporicin and microbisporicin implies that the *M. corallina* $\Delta mibEF$ deletion mutant might be complemented through the expression of the *pspEF* genes. Cross-immunity has been observed for the producing strains of epigidin 280 and Pep5, indicating that the immunity peptides can function to protect the producer from two lantibiotics (Heidrich *et al.* 1998).

Another desirable deletion mutant would be for the membrane protein PspJ, a potential regulator of PspW. The model of planosporicin regulation proposed in Chapter 8 is based entirely on observed expression levels in deletion mutants of three key regulatory genes. The role of the regulators PspX, PspW and PspR could be substantiated through a number of additional experiments. The predicted interaction between PspX and PspW could be investigated through bacterial-two-hybrid experiments. This technique could also be used to see if the inactivation of PspW occurs through a direct interaction with either PspJ or PspYZ. The physiological transcription start site of each *psp* operon transcribed *in vivo* could be mapped to check whether the ECF sigma factor consensus sequences identified *in silico* have a genuine function in the regulation of the gene cluster. This technique validated the consensus sequences identified in the *mib* cluster (Foulston and Bibb 2011). Once established, run-off transcription assays using the holoenzyme of RNAP with purified PspX as a sigma factor could confirm that PspX recruits RNAP to these sites. Likewise, if active PspR can be overexpressed and purified then DNA-protein binding could be investigated through qualitative techniques such as gel-shifts and footprinting. The proposed direct interaction of PspX with the promoter region of the operons *pspXW*, *pspJYZQ*, *pspR* and *pspEF* could be investigated through Lux-reporter analysis in *S. coelicolor*. This experiment could be extended to investigate the direct interaction of PspR with the promoter region of the *pspABCTUV* operon. The production of light by *luxAB* cloned downstream of the *mibX* promoter provided evidence that MibX can positively auto-regulate its own expression (Foulston and Bibb 2011).

The striking phenotype of the *pspW* deletion mutant may form a paradigm for other secondary metabolite clusters regulated by a sigma factor-anti sigma factor complex. It is tempting to mobilise the $\Delta pspW$ *psp* cluster into *Nonomuraea* to see if this mutation has the same effect of increased planosporicin expression levels in a heterologous host. However, this deletion is unlikely to solve the lack of heterologous expression in *Streptomyces*, as RT-PCR on *Streptomyces* with the integrated *mib* cluster showed some

level of transcription does occur (Foulston 2010), implying the problem may instead be with the level of immunity. Deletion of the anti-sigma factor may provide a universal method to enhance expression of respective secondary metabolite clusters. This could partly alleviate the time-consuming optimisation of culture conditions and so may be of great use to the pharmaceutical industry.

The fact that lantibiotics are gene-encoded makes them amenable to engineering strategies, which may be employed to further enhance their activity. This initially took the form of rational design, a knowledge-based approach to alter residues to enhance bioactivity. Although this failed to result in better variants, much was learnt about the structure/function relationship of lantibiotics (Cortes *et al.* 2009). Recently we have entered what is sometimes referred to as a 'golden era' in lantibiotic bioengineering through the construction of randomised libraries (Field *et al.* 2010). Through this strategy, derivatives of nisin, mersacidin and nukacin ISK-1 have been identified with improved activity (Field *et al.* 2007; Field *et al.* 2008; Appleyard *et al.* 2009; Islam *et al.* 2009). However, success depends on the type and number of variants generated. The use of alanine scanning mutagenesis of lacticin 3147 and actagardine, and the generation of a small number of randomly altered lacticin 3147 variants failed to generate more potent forms (Cotter *et al.* 2006; Field *et al.* 2007; Boakes *et al.* 2009). Thus the development of a method to generate random planosporicin variants may reveal which features confer its potent bactericidal activity and might even result in the production of more potent compounds. The combined knowledge of microbisporicin and planosporicin gives scope for the rational design of planosporicin variants containing additional modifications observed in microbisporicin.

The overall similarity between not only the structure of microbisporicin and planosporicin, but also the gene clusters directing their biosynthesis has many implications. The fusion of different parts of *pspA* and *mibA* could lead to the creation of *mib-psp* chimeras. The relaxed substrate specificity observed in many LanB and LanC proteins may enable subsequent post-translational modification as observed in nisin-subtilin chimeras. The expression of the leader peptide DNA sequence of subtilin fused to the sequence encoding pronisin Z in *L. lactis* (a strain that produces nisin A) led to the detection of nisin A and mature nisin Z with the unmodified leader peptide of subtilin still attached (Kuipers *et al.* 1993b). The reciprocal experiment expressed nisin-subtilin chimeric prepeptides in a *B. subtilis* 168 strain in which the presubtilin gene had been deleted from the *spa* cluster. The chimeras consisted of the subtilin leader peptide fused to a structural region composed of the N and C terminus from either nisin or subtilin. Chimeras in which the N-terminus of nisin was fused to the C-terminus of subtilin resulted in production of the fully modified chimera, whereas if the chimeric structural region had a subtilin N terminus and a

nisin C terminus, a heterogeneous mixture of products was detected (Chakicherla and Hansen 1995). Clearly specific recognition between the processing machinery and the prelantibiotic peptide (particularly the leader peptide and the C-terminus of the structural region) is a prerequisite for processing to occur correctly. Exchanging the leader peptide sequence of the planosporicin prepropeptide with that of microbisporicin may reveal which structural features of the leader peptide are essential for particular functions. It may also be possible to export these *mib-*psp** chimeras, as the NisT exporter has demonstrated relaxed substrate specificity in exporting fusions of the NisA leader peptide with non-lantibiotic peptides (Kuipers *et al.* 2004).

It is interesting to speculate on the evolution of the *psp* and *mib* gene clusters. The absence of a nearby transposase implies they are not related through horizontal gene transfer, but are instead related through vertical transfer. In particular, the biosynthetic operons *pspABCTUV* and *mibABCDTUV* are very similar. The main difference is the absence of a MibD homologue in *P. alba*. Two scenarios can be envisaged. In one, the common ancestor to the *mib* and *psp* clusters did not encode a *lanD*, but during the evolution of the *mib* cluster, an additional enzyme was acquired resulting in AviCys formation at the C-terminus of microbisporicin. Alternatively, it is possible that the common ancestor may have encoded a *lanD* that was subsequently deleted during the evolution of the *psp* cluster. The loss of the AviCys would likely be prompted by a selection pressure, perhaps by another mutation which further increased potency. Therefore there was a need to abolish AviCys formation, to restore the C-terminal negative charge to reduce affinity to the membrane containing the molecular target lipid II, with a consequent reduction in potency. Given the difference in potency between planosporicin and microbisporicin, there remains much scope for the improvement of planosporicin activity and it joins other lantibiotics as one of the arsenal of potential antibiotics required to fight the ever present threat posed by infectious diseases.

Chapter 10 : References

Allgaier, H., G. Jung, R. G. Werner, U. Schneider and H. Zahner (1985). "Elucidation of the structure of epidermin, a ribosomally synthesized tetracyclic heterodetic polypeptide antibiotic." Angewandte Chemie **24**: 1051-1053.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman (1990). "Basic local alignment search tool." Journal of Molecular Biology **215**(3): 403-410.

Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller and D. J. Lipman (1997). "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs." Nucleic Acids Research **25**(17): 3389-3402.

Appelbaum, P. C. and M. R. Jacobs (2005). "Recently approved and investigational antibiotics for treatment of severe infections caused by Gram-positive bacteria." Current Opinion in Microbiology **8**(5): 510-517.

Appleyard, A. N., S. Choi, D. M. Read, A. Lightfoot, S. Boakes, A. Hoffmann, I. Chopra, G. Bierbaum, B. A. Rudd, M. J. Dawson and J. Cortes (2009). "Dissecting structural and functional diversity of the lantibiotic mersacidin." Chemistry & Biology **16**(5): 490-498.

Augustin, J., R. Rosenstein, B. Wieland, U. Schneider, N. Schnell, G. Engelke, K. D. Entian and F. Gotz (1992). "Genetic analysis of epidermin biosynthetic genes and epidermin-negative mutants of *Staphylococcus epidermidis*." European journal of biochemistry / FEBS **204**(3): 1149-1154.

Bailey, T. L. and C. Elkan (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology, Menlo Park, California, AAAI Press.

Baltz, R. H. (1998). "Genetic manipulation of antibiotic-producing *Streptomyces*." Trends in Microbiology **6**(2): 76-83.

Baltz, R. H. (2007). "Antimicrobials from Actinomycetes: Back to the Future." Microbe **2**(3): 125-131.

Baltz, R. H. (2008). "Renaissance in antibacterial discovery from actinomycetes." Current Opinion in Pharmacology **8**(5): 557-563.

Barna, J. C. and D. H. Williams (1984). "The structure and mode of action of glycopeptide antibiotics of the vancomycin group." Annual Review of Microbiology **38**: 339-357.

- Bauer, R. and L. M. Dicks (2005). "Mode of action of lipid II-targeting lantibiotics." International Journal of Food Microbiology **101**(2): 201-216.
- Begley, M., P. D. Cotter, C. Hill and R. P. Ross (2009). "Identification of a novel two-peptide lantibiotic, lichenicidin, following rational genome mining for LanM proteins." Applied and Environmental Microbiology **75**(17): 5451-5460.
- Bell, R. (2010). Analysis and manipulation of "actagardine" gene clusters from *Actinoplanes*. John Innes Centre. Norwich, University of East Anglia. **PhD**.
- Beltrametti, F., R. Rossi, E. Selva and F. Marinelli (2006). "Antibiotic production improvement in the rare actinomycete *Planobispora rosea* by selection of mutants resistant to the aminoglycosides streptomycin and gentamycin and to rifamycin." Journal of Industrial Microbiology & Biotechnology **33**(4): 283-288.
- 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. Rabinowitsch, 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 (2002). "Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2)." Nature **417**(6885): 141-147.
- Bibb, M. J. (2005). "Regulation of secondary metabolism in streptomycetes." Current Opinion in Microbiology **8**(2): 208-215.
- Bibb, M. J., P. R. Findlay and M. W. Johnson (1984). "The relationship between base composition and codon usage in bacterial genes and its use for the simple and reliable identification of protein-coding sequences." Gene **30**(1-3): 157-166.
- Bibb, M. J., G. R. Janssen and J. M. Ward (1985). "Cloning and analysis of the promoter region of the erythromycin resistance gene (*ermE*) of *Streptomyces erythraeus*." Gene **38**(1-3): 215-226.
- Bibb, M. J., V. Molle and M. J. Buttner (2000). "sigma(BldN), an extracytoplasmic function RNA polymerase sigma factor required for aerial mycelium formation in *Streptomyces coelicolor* A3(2)." Journal of Bacteriology **182**(16): 4606-4616.
- Bierman, M., R. Logan, K. O'Brien, E. T. Seno, R. N. Rao and B. E. Schoner (1992). "Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp." Gene **116**(1): 43-49.
- Blaesse, M., T. Kupke, R. Huber and S. Steinbacher (2000). "Crystal structure of the peptidyl-cysteine decarboxylase EpiD complexed with a pentapeptide substrate." The EMBO journal **19**(23): 6299-6310.

Blaesse, M., T. Kupke, R. Huber and S. Steinbacher (2003). "Structure of MrsD, an FAD-binding protein of the HFCD family." Acta crystallographica. Section D, Biological crystallography **59**(Pt 8): 1414-1421.

Boakes, S., A. N. Appleyard, J. Cortes and M. J. Dawson (2010). "Organization of the biosynthetic genes encoding deoxyactagardine B (DAB), a new lantibiotic produced by *Actinoplanes liguriae* NCIMB41362." The Journal of Antibiotics **63**(7): 351-358.

Boakes, S., J. Cortes, A. N. Appleyard, B. A. Rudd and M. J. Dawson (2009). "Organization of the genes encoding the biosynthesis of actagardine and engineering of a variant generation system." Molecular Microbiology **72**(5): 1126-1136.

Bonelli, R. R., T. Schneider, H. G. Sahl and I. Wiedemann (2006). "Insights into in vivo activities of lantibiotics from gallidermin and epidermin mode-of-action studies." Antimicrobial Agents and Chemotherapy **50**(4): 1449-1457.

Bouhss, A., B. Al-Dabbagh, M. Vincent, B. Odaert, M. Aumont-Nicaise, P. Bressolier, M. Desmadril, D. Mengin-Lecreulx, M. C. Urdaci and J. Gallay (2009). "Specific interactions of clausin, a new lantibiotic, with lipid precursors of the bacterial cell wall." Biophysical Journal **97**(5): 1390-1397.

Bressolier, P., M. A. Brugo, P. Robineau, J.-M. Schmitter, M. Sofeir, M. C. Urdaci and B. Verneuil (2007). Peptide compound with biological activity, its preparation and its applications. S.-A. S. (Milano). United States.

Breukink, E. and B. de Kruijff (2006). "Lipid II as a target for antibiotics." Nature Reviews Drug Discovery **5**(4): 321-332.

Brotz, H., G. Bierbaum, K. Leopold, P. E. Reynolds and H. G. Sahl (1998). "The lantibiotic mersacidin inhibits peptidoglycan synthesis by targeting lipid II." Antimicrobial Agents and Chemotherapy **42**(1): 154-160.

Buchman, G. W., S. Banerjee and J. N. Hansen (1988). "Structure, expression, and evolution of a gene encoding the precursor of nisin, a small protein antibiotic." The Journal of Biological Chemistry **263**(31): 16260-16266.

Butcher, B. G. and J. D. Helmann (2006). "Identification of *Bacillus subtilis* sigma-dependent genes that provide intrinsic resistance to antimicrobial compounds produced by bacilli." Molecular Microbiology **60**(3): 765-782.

Campbell, E. A., R. Greenwell, J. R. Anthony, S. Wang, L. Lim, K. Das, H. J. Sofia, T. J. Donohue and S. A. Darst (2007). "A conserved structural module regulates transcriptional responses to diverse stress signals in bacteria." Molecular Cell **27**(5): 793-805.

Campbell, E. A., O. Muzzin, M. Chlenov, J. L. Sun, C. A. Olson, O. Weinman, M. L. Trester-Zedlitz and S. A. Darst (2002). "Structure of the bacterial RNA polymerase promoter specificity sigma subunit." Molecular Cell **9**(3): 527-539.

Cao, M. and J. D. Helmann (2004). "The *Bacillus subtilis* extracytoplasmic-function sigmaX factor regulates modification of the cell envelope and resistance to cationic antimicrobial peptides." Journal of Bacteriology **186**(4): 1136-1146.

Castiglione, F., L. Cavaletti, D. Losi, A. Lazzarini, L. Carrano, M. Feroggio, I. Ciciliato, E. Corti, G. Candiani, F. Marinelli and E. Selva (2007). "A novel lantibiotic acting on bacterial cell wall synthesis produced by the uncommon actinomycete *Planomonospora* sp." Biochemistry **46**(20): 5884-5895.

Castiglione, F., A. Lazzarini, L. Carrano, E. Corti, I. Ciciliato, L. Gastaldo, P. Candiani, D. Losi, F. Marinelli, E. Selva and F. Parenti (2008). "Determining the structure and mode of action of microbisporicin, a potent lantibiotic active against multiresistant pathogens." Chemistry & Biology **15**(1): 22-31.

Chakicherla, A. and J. N. Hansen (1995). "Role of the leader and structural regions of prelantibiotic peptides as assessed by expressing nisin-subtilin chimeras in *Bacillus subtilis* 168, and characterization of their physical, chemical, and antimicrobial properties." Journal of Biological Chemistry **270**(40): 23533-23539.

Chan, W. C., B. W. Bycroft, M. L. Leyland, L. Y. Lian and G. C. Roberts (1993). "A novel post-translational modification of the peptide antibiotic subtilin: isolation and characterization of a natural variant from *Bacillus subtilis* ATCC 6633." The Biochemical Journal **291** (Pt 1): 23-27.

Chatterjee, C., G. C. Patton, L. Cooper, M. Paul and W. A. van der Donk (2006). "Engineering dehydro amino acids and thioethers into peptides using lactacin 481 synthetase." Chemistry & Biology **13**(10): 1109-1117.

Chatterjee, C., M. Paul, L. Xie and W. A. van der Donk (2005). "Biosynthesis and Mode of Action of Lantibiotics." Chemical Reviews **105**(2): 633-684.

Chen, P., J. Novak, M. Kirk, S. Barnes, F. Qi and P. W. Caufield (1998). "Structure-activity study of the lantibiotic mutacin II from *Streptococcus mutans* T8 by a gene replacement strategy." Applied and environmental microbiology **64**(7): 2335-2340.

Chen, P., F. X. Qi, J. Novak, R. E. Krull and P. W. Caufield (2001). "Effect of amino acid substitutions in conserved residues in the leader peptide on biosynthesis of the lantibiotic mutacin II." FEMS Microbiology Letters **195**(2): 139-144.

Chenna, R., H. Sugawara, T. Koike, R. Lopez, T. J. Gibson, D. G. Higgins and J. D. Thompson (2003). "Multiple sequence alignment with the Clustal series of programs." Nucleic Acids Research **31**(13): 3497-3500.

Cherepanov, P. P. and W. Wackernagel (1995). "Gene Disruption in *Escherichia coli* - Tcr and Km(R) Cassettes with the Option of Flp-Catalyzed Excision of the Antibiotic-Resistance Determinant." Gene **158**(1): 9-14.

Chiang, Y. M., S. L. Chang, B. R. Oakley and C. C. Wang (2011). "Recent advances in awakening silent biosynthetic gene clusters and linking orphan clusters to natural products in microorganisms." Current Opinion in Chemical Biology **15**(1): 137-143.

Claesen, J. (2010). Cloning and analysis of the cypemycin biosynthetic gene cluster. John Innes Centre. Norwich, University of East Anglia. **PhD**.

Claesen, J. and M. Bibb (2010). "Genome mining and genetic analysis of cypemycin biosynthesis reveal an unusual class of posttranslationally modified peptides." Proceedings of the National Academy of Sciences of the United States of America **107**(37): 16297-16302.

Claesen, J. and M. J. Bibb (2011). "Biosynthesis and Regulation of Grisemycin, a New Member of the Linaridin Family of Ribosomally Synthesized Peptides Produced by *Streptomyces griseus* IFO 13350." Journal of Bacteriology **193**(10): 2510-2516.

Combes, P., R. Till, S. Bee and M. C. Smith (2002). "The *Streptomyces* genome contains multiple pseudo-*attB* sites for the (phi)C31-encoded site-specific recombination system." Journal of Bacteriology **184**(20): 5746-5752.

Cooper, L. E., B. Li and W. A. van der Donk (2010). Biosynthesis and Mode of Action of Lantibiotics. Comprehensive Natural Products II Chemistry and Biology. W. Mander and H.-W. Liu, Elsevier. **5**: 217-256.

Cooper, L. E., A. L. McClerren, A. Chary and W. A. van der Donk (2008). "Structure-Activity Relationship Studies of the Two-Component Lantibiotic Haloduracin." Chemistry & Biology **59**(10): 1035-1045.

Cortes, J., A. N. Appleyard and M. J. Dawson (2009). "Chapter 22. Whole-cell generation of lantibiotic variants." Methods in Enzymology **458**: 559-574.

Corvey, C., T. Stein, S. Dusterhus, M. Karas and K. D. Entian (2003). "Activation of subtilin precursors by *Bacillus subtilis* extracellular serine proteases subtilisin (AprE), WprA, and Vpr." Biochemical and Biophysical Research Communications **304**(1): 48-54.

Cotter, P. D., L. H. Deegan, E. M. Lawton, L. A. Draper, P. M. O'Connor, C. Hill and R. P. Ross (2006). "Complete alanine scanning of the two-component lantibiotic lactacin 3147: generating a blueprint for rational drug design." Molecular Microbiology **62**(3): 735-747.

Cotter, P. D., C. Hill and R. P. Ross (2005a). "Bacteriocins: developing innate immunity for food." Nature reviews. Microbiology **3**(10): 777-788.

Cotter, P. D., P. M. O'Connor, L. A. Draper, E. M. Lawton, L. H. Deegan, C. Hill and R. P. Ross (2005b). "Posttranslational conversion of L-serines to D-alanines is vital for optimal production and activity of the lantibiotic lactacin 3147." Proceedings of the National Academy of Sciences of the United States of America **102**(51): 18584-18589.

- Court, D. L., J. A. Sawitzke and L. C. Thomason (2002). "Genetic engineering using homologous recombination." Annual Review of Genetics **36**: 361-388.
- Cox, A. R., N. R. Thomson, B. Bycroft, G. S. Stewart, P. Williams and G. P. Salmond (1998). "A pheromone-independent CarR protein controls carbapenem antibiotic synthesis in the opportunistic human pathogen *Serratia marcescens*." Microbiology **144 (Pt 1)**: 201-209.
- Cox, K. L., S. E. Fishman, J. L. Larson, R. Stanzak, P. A. Reynolds, W. K. Yeh, R. M. van Frank, V. A. Birmingham, C. L. Hershberger and E. T. Seno (1986). "The use of recombinant DNA techniques to study tylosin biosynthesis and resistance in *Streptomyces fradiae*." Journal of Natural Products **49(6)**: 971-980.
- Croucher, N. J., S. R. Harris, C. Fraser, M. A. Quail, J. Burton, M. van der Linden, L. McGee, A. von Gottberg, J. H. Song, K. S. Ko, B. Pichon, S. Baker, C. M. Parry, L. M. Lambertsen, D. Shahinas, D. R. Pillai, T. J. Mitchell, G. Dougan, A. Tomasz, K. P. Klugman, J. Parkhill, W. P. Hanage and S. D. Bentley (2011). "Rapid pneumococcal evolution in response to clinical interventions." Science **331(6016)**: 430-434.
- Da Re, S., J. Schumacher, P. Rousseau, J. Fourment, C. Ebel and D. Kahn (1999). "Phosphorylation-induced dimerization of the FixJ receiver domain." Molecular Microbiology **34(3)**: 504-511.
- Daly, K. M., M. Upton, S. K. Sandiford, L. A. Draper, P. A. Wescombe, R. W. Jack, P. M. O'Connor, A. Rossney, F. Gotz, C. Hill, P. D. Cotter, R. P. Ross and J. R. Tagg (2010). "Production of the Bsa lantibiotic by community-acquired *Staphylococcus aureus* strains." Journal of Bacteriology **192(4)**: 1131-1142.
- Datsenko, K. A. and B. L. Wanner (2000). "One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products." Proceedings of the National Academy of Sciences of the United States of America **97(12)**: 6640-6645.
- Davies, J. (2011). "How to discover new antibiotics: harvesting the parvome." Current Opinion in Chemical Biology **15(1)**: 5-10.
- de Jong, A., A. J. van Heel, J. Kok and O. P. Kuipers (2010). "BAGEL2: mining for bacteriocins in genomic data." Nucleic Acids Research **38(Web Server issue)**: W647-651.
- de Kwaadsteniet, M., K. Ten Doeschate and L. M. Dicks (2008). "Characterization of the structural gene encoding nisin F, a new lantibiotic produced by a *Lactococcus lactis* subsp. *lactis* isolate from freshwater catfish (*Clarias gariepinus*)." Applied and Environmental Microbiology **74(2)**: 547-549.
- de Ruyter, P. G., O. P. Kuipers, M. M. Beerthuyzen, I. van Alen-Boerrigter and W. M. de Vos (1996). "Functional analysis of promoters in the nisin gene cluster of *Lactococcus lactis*." Journal of Bacteriology **178(12)**: 3434-3439.

de Vos, W. M., J. W. Mulders, R. J. Siezen, J. Hugenholtz and O. P. Kuipers (1993). "Properties of nisin Z and distribution of its gene, nisZ, in *Lactococcus lactis*." Applied and Environmental Microbiology **59**(1): 213-218.

Dorenbos, R., T. Stein, J. Kabel, C. Bruand, A. Bolhuis, S. Bron, W. J. Quax and J. M. Van Dijk (2002). "Thiol-disulfide oxidoreductases are essential for the production of the lantibiotic sublancin 168." The Journal of Biological Chemistry **277**(19): 16682-16688.

Draper, L. A., K. Grainger, L. H. Deegan, P. D. Cotter, C. Hill and R. P. Ross (2009). "Cross-immunity and immune mimicry as mechanisms of resistance to the lantibiotic lactacin 3147." Molecular Microbiology **71**(4): 1043-1054.

Drugeon, H. B., M. E. Juvin, J. Caillon and A. L. Courtieu (1987). "Assessment of formulas for calculating critical concentration by the agar diffusion method." Antimicrobial Agents and Chemotherapy **31**(6): 870-875.

Dworkin, M. (1996). The Prokaryotes: Archaea. Bacteria: Firmicutes, Actinomycetes, Springer.

Emanuelsson, O., S. Brunak, G. von Heijne and H. Nielsen (2007). "Locating proteins in the cell using TargetP, SignalP and related tools." Nature Protocols **2**(4): 953-971.

Engelke, G., Z. Gutowski-Eckel, M. Hammelmann and K. D. Entian (1992). "Biosynthesis of the lantibiotic nisin: genomic organization and membrane localization of the NisB protein." Applied and Environmental Microbiology **58**(11): 3730-3743.

Enright, M. C. and H. McKenzie (1997). "*Moraxella (Branhamella) catarrhalis*--clinical and molecular aspects of a rediscovered pathogen." Journal of Medical Microbiology **46**(5): 360-371.

Erickson, J. W. and C. A. Gross (1989). "Identification of the sigma E subunit of *Escherichia coli* RNA polymerase: a second alternate sigma factor involved in high-temperature gene expression." Genes & Development **3**(9): 1462-1471.

Eustaquio, A. S., B. Gust, U. Galm, S. M. Li, K. F. Chater and L. Heide (2005). "Heterologous expression of novobiocin and clorobiocin biosynthetic gene clusters." Applied and Environmental Microbiology **71**(5): 2452-2459.

Felnagle, E. A., M. R. Rondon, A. D. Berti, H. A. Crosby and M. G. Thomas (2007). "Identification of the biosynthetic gene cluster and an additional gene for resistance to the antituberculosis drug capreomycin." Applied and Environmental Microbiology **73**(13): 4162-4170.

Field, D., B. Collins, P. D. Cotter, C. Hill and R. P. Ross (2007). "A system for the random mutagenesis of the two-peptide lantibiotic lactacin 3147: analysis of mutants producing reduced antibacterial activities." Journal of Molecular Microbiology and Biotechnology **13**(4): 226-234.

Field, D., P. M. O. Connor, P. D. Cotter, C. Hill and R. P. Ross (2008). "The generation of nisin variants with enhanced activity against specific Gram-positive pathogens." Molecular Microbiology **69**(1): 218-230.

Field, D., C. Hill, P. D. Cotter and R. P. Ross (2010). "The dawning of a 'Golden era' in lantibiotic bioengineering." Molecular Microbiology **78**(5): 1077-1087.

Finn, R. D., J. Tate, J. Mistry, P. C. Coggill, S. J. Sammut, H. R. Hotz, G. Ceric, K. Forslund, S. R. Eddy, E. L. Sonnhammer and A. Bateman (2008). "The Pfam protein families database." Nucleic Acids Research **36**(Database issue): D281-288.

Fischbach, M. A. (2009). "Antibiotics from microbes: converging to kill." Current Opinion in Microbiology **12**(5): 520-527.

Fischbach, M. A. and C. T. Walsh (2009). "Antibiotics for emerging pathogens." Science **325**(5944): 1089-1093.

Fischbach, M. A., C. T. Walsh and J. Clardy (2008). "The evolution of gene collectives: How natural selection drives chemical innovation." Proceedings of the National Academy of Sciences of the United States of America **105**(12): 4601-4608.

Flett, F., V. Mersinias and C. P. Smith (1997). "High efficiency intergeneric conjugal transfer of plasmid DNA from *Escherichia coli* to methyl DNA-restricting streptomycetes." FEMS Microbiology Letters **155**(2): 223-229.

Foulston, L. (2010). Cloning and Analysis of the Microbisporicin Lantibiotic Gene Cluster from *Microbispora corallina*. John Innes Centre. Norwich, University of East Anglia. **PhD**.

Foulston, L. and M. Bibb (2011). "Feed-forward regulation of microbisporicin biosynthesis in *Microbispora corallina*." Journal of Bacteriology **193**(12): 3064-3071.

Foulston, L. C. and M. J. Bibb (2010). "Microbisporicin gene cluster reveals unusual features of lantibiotic biosynthesis in actinomycetes." Proceedings of the National Academy of Sciences of the United States of America **107**(30): 13461-13466.

Fredenhagen, A., G. Fendrich, F. Marki, W. Marki, J. Gruner, F. Raschdorf and H. H. Peter (1990). "Duramycins B and C, two new lanthionine containing antibiotics as inhibitors of phospholipase A2. Structural revision of duramycin and cinnamycin." The Journal of Antibiotics **43**(11): 1403-1412.

Furgerson Ihnken, L. A., C. Chatterjee and W. A. van der Donk (2008). "*In vitro* reconstitution and substrate specificity of a lantibiotic protease." Biochemistry **47**(28): 7352-7363.

Geissler, S., F. Gotz and T. Kupke (1996). "Serine protease EpiP from *Staphylococcus epidermidis* catalyzes the processing of the epidermin precursor peptide." Journal of Bacteriology **178**(1): 284-288.

Gilmore, M. S., R. A. Segarra, M. C. Booth, C. P. Bogie, L. R. Hall and D. B. Clewell (1994). "Genetic structure of the *Enterococcus faecalis* plasmid pAD1-encoded cytolytic toxin system and its relationship to lantibiotic determinants." Journal of Bacteriology **176**(23): 7335-7344.

Goh, E. B., G. Yim, W. Tsui, J. McClure, M. G. Surette and J. Davies (2002). "Transcriptional modulation of bacterial gene expression by subinhibitory concentrations of antibiotics." Proceedings of the National Academy of Sciences of the United States of America **99**(26): 17025-17030.

Gomez-Escribano, J. P. and M. J. Bibb (2011). "Engineering *Streptomyces coelicolor* for heterologous expression of secondary metabolite gene clusters." Microbial biotechnology **4**(2): 207-215.

Gonzalez, B., P. Arca, B. Mayo and J. E. Suarez (1994). "Detection, purification, and partial characterization of plantaricin C, a bacteriocin produced by a *Lactobacillus plantarum* strain of dairy origin." Applied and Environmental Microbiology **60**(6): 2158-2163.

Goodfellow, M., L. J. Stanton, K. E. Simpson and D. E. Minnikin (1990). "Numerical and chemical classification of *Actinoplanes* and some related actinomycetes." Microbiology **136**(1): 19.

Gordon, D., C. Abajian and P. Green (1998). "Consed: a graphical tool for sequence finishing." Genome Research **8**(3): 195-202.

Goto, Y., B. Li, J. Claesen, Y. Shi, M. J. Bibb and W. A. van der Donk (2010). "Discovery of unique lanthionine synthetases reveals new mechanistic and evolutionary insights." PLoS Biology **8**(3): e1000339.

Gowans, J. L., N. Smith and H. W. Florey (1952). "Some properties of nisin." British Journal of Pharmacology and Chemotherapy **7**(3): 438-449.

Gregory, M. A., R. Till and M. C. Smith (2003). "Integration site for *Streptomyces* phage phiBT1 and development of site-specific integrating vectors." Journal of Bacteriology **185**(17): 5320-5323.

Gross, E., H. H. Kiltz and E. Nebelin (1973). "Subtilin, VI: the structure of subtilin." Hoppe-Seyler's Zeitschrift fur physiologische Chemie **354**(7): 810-812.

Gross, E. and J. L. Morell (1971). "The structure of nisin." Journal of the American Chemical Society **93**(18): 4634-4635.

Guder, A., T. Schmitter, I. Wiedemann, H. G. Sahl and G. Bierbaum (2002). "Role of the single regulator MrsR1 and the two-component system MrsR2/K2 in the regulation of mersacidin production and immunity." Applied and Environmental Microbiology **68**(1): 106-113.

Gust, B., G. L. Challis, K. Fowler, T. Kieser and K. F. Chater (2003). "PCR-targeted *Streptomyces* gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin." Proceedings of the National Academy of Sciences of the United States of America **100**(4): 1541-1546.

Gust, B., G. Chandra, D. Jakimowicz, T. Yuqing, C. J. Bruton and K. F. Chater (2004). "Lambda red-mediated genetic manipulation of antibiotic-producing *Streptomyces*." Advances in Applied Microbiology **54**: 107-128.

Hara, O. and T. Beppu (1982). "Mutants blocked in streptomycin production in *Streptomyces griseus* - the role of A-factor." The Journal of Antibiotics **35**(3): 349-358.

Hasper, H. E., B. de Kruijff and E. Breukink (2004). "Assembly and stability of nisin-lipid II pores." Biochemistry **43**(36): 11567-11575.

Hasper, H. E., N. E. Kramer, J. L. Smith, J. D. Hillman, C. Zachariah, O. P. Kuipers, B. de Kruijff and E. Breukink (2006). "An alternative bactericidal mechanism of action for lantibiotic peptides that target lipid II." Science **313**(5793): 1636-1637.

Heidrich, C., U. Pag, M. Josten, J. Metzger, R. W. Jack, G. Bierbaum, G. Jung and H. G. Sahl (1998). "Isolation, characterization, and heterologous expression of the novel lantibiotic epicidin 280 and analysis of its biosynthetic gene cluster." Applied and Environmental Microbiology **64**(9): 3140-3146.

Heinrich, J. and T. Wiegert (2009). "Regulated intramembrane proteolysis in the control of extracytoplasmic function sigma factors." Research in Microbiology **160**(9): 696-703.

Helfrich, M., K. D. Entian and T. Stein (2007). "Structure-function relationships of the lanthionine cyclase SpaC involved in biosynthesis of the *Bacillus subtilis* peptide antibiotic subtilin." Biochemistry **46**(11): 3224-3233.

Hesketh, A., W. J. Chen, J. Ryding, S. Chang and M. Bibb (2007). "The global role of ppGpp synthesis in morphological differentiation and antibiotic production in *Streptomyces coelicolor* A3(2)." Genome Biology **8**(8): R161.

Hesketh, A., H. Kock, S. Mootien and M. Bibb (2009). "The role of absC, a novel regulatory gene for secondary metabolism, in zinc-dependent antibiotic production in *Streptomyces coelicolor* A3(2)." Molecular microbiology **74**(6): 1427-1444.

Hesketh, A., J. Sun and M. Bibb (2001). "Induction of ppGpp synthesis in *Streptomyces coelicolor* A3(2) grown under conditions of nutritional sufficiency elicits actII-ORF4 transcription and actinorhodin biosynthesis." Molecular Microbiology **39**(1): 136-144.

Hofmann, K. and W. Stoffel (1993). TMbase - A database of membrane spanning proteins segments. Biological Chemistry, Hoppe-Seyler. **374**: 166.

Holo, H., Z. Jeknic, M. Daeschel, S. Stevanovic and I. F. Nes (2001). "Plantaricin W from *Lactobacillus plantarum* belongs to a new family of two-peptide lantibiotics." Microbiology **147**(Pt 3): 643-651.

Holtsmark, I., D. Mantzilas, V. G. Eijsink and M. B. Brurberg (2006). "Purification, characterization, and gene sequence of michiganin A, an actagardine-like lantibiotic produced by the tomato pathogen *Clavibacter michiganensis* subsp. *michiganensis*." Applied and Environmental Microbiology **72**(9): 5814-5821.

Hopwood, D. A., T. Kieser, H. M. Wright and M. J. Bibb (1983). "Plasmids, recombination and chromosome mapping in *Streptomyces lividans* 66." Journal of General Microbiology **129**(7): 2257-2269.

Horinouchi, S. and T. Beppu (1994). "A-factor as a microbial hormone that controls cellular differentiation and secondary metabolism in *Streptomyces griseus*." Molecular Microbiology **12**(6): 859-864.

Hsu, S. T. D., E. Breukink, G. Bierbaum, H. G. Sahl, B. de Kruijff, R. Kaptein, N. A. J. van Nuland and A. M. J. J. Bonvin (2003). "NMR study of mersacidin and lipid II interaction in dodecylphosphocholine micelles - Conformational changes are a key to antimicrobial activity." Journal of Biological Chemistry **278**(15): 13110-13117.

Hsu, S. T. D., E. Breukink, E. Tischenko, M. A. G. Lutters, B. de Kruijff, R. Kaptein, A. M. J. J. Bonvin and N. A. J. van Nuland (2004). "The nisin-lipid II complex reveals a pyrophosphate cage that provides a blueprint for novel antibiotics." Nature Structural & Molecular Biology **11**(10): 963-967.

Huang, J., J. Shi, V. Molle, B. Sohlberg, D. Weaver, M. J. Bibb, N. Karoonuthaisiri, C. J. Lih, C. M. Kao, M. J. Buttner and S. N. Cohen (2005). "Cross-regulation among disparate antibiotic biosynthetic pathways of *Streptomyces coelicolor*." Molecular Microbiology **58**(5): 1276-1287.

Hudson, M. E. and J. R. Nodwell (2004). "Dimerization of the RamC morphogenetic protein of *Streptomyces coelicolor*." Journal of Bacteriology **186**(5): 1330-1336.

Hugenholtz, P. and E. Stackebrandt (2004). "Reclassification of *Sphaerobacter thermophilus* from the subclass Sphaerobacteridae in the phylum Actinobacteria to the class Thermomicrobia (emended description) in the phylum Chloroflexi (emended description)." International Journal of Systematic and Evolutionary Microbiology **54**(Pt 6): 2049-2051.

Hutchings, M. I., H. J. Hong, E. Leibovitz, I. C. Sutcliffe and M. J. Buttner (2006). "The sigma(E) cell envelope stress response of *Streptomyces coelicolor* is influenced by a novel lipoprotein, CseA." Journal of Bacteriology **188**(20): 7222-7229.

Hutchings, M. I., T. Palmer, D. J. Harrington and I. C. Sutcliffe (2009). "Lipoprotein biogenesis in Gram-positive bacteria: knowing when to hold 'em, knowing when to fold 'em." Trends in Microbiology **17**(1): 13-21.

Ihnken, L. A. F., C. Chatterjee and W. A. van der Donk (2008). "*In vitro* reconstitution and substrate specificity of a lantibiotic protease." Biochemistry **47**(28): 7352-7363.

Islam, M. R., K. Shioya, J. Nagao, M. Nishie, H. Jikuya, T. Zendo, J. Nakayama and K. Sonomoto (2009). "Evaluation of essential and variable residues of nukacin ISK-1 by NNK scanning." Molecular Microbiology **72**(6): 1438-1447.

Jacob, F., A. Lwoff, A. Siminovitch and E. Wollman (1953). "[Definition of some terms relative to lysogeny]." Annales de l'Institut Pasteur **84**(1): 222-224.

Johnson, A. P., A. Pearson and G. Duckworth (2005). "Surveillance and epidemiology of MRSA bacteraemia in the UK." Journal of Antimicrobial Chemotherapy **56**(3): 455-462.

Juncker, A. S., H. Willenbrock, G. Von Heijne, S. Brunak, H. Nielsen and A. Krogh (2003). "Prediction of lipoprotein signal peptides in Gram-negative bacteria." Protein science : a publication of the Protein Society **12**(8): 1652-1662.

Jung, D., A. Rozek, M. Okon and R. E. Hancock (2004). "Structural transitions as determinants of the action of the calcium-dependent antibiotic daptomycin." Chemistry & Biology **11**(7): 949-957.

Jung, G. and H.-G. Sahl (1991). Nisin and Novel Lantibiotics, Leiden, The Netherlands.

Kabuki, T., H. Uenishi, Y. Seto, T. Yoshioka and H. Nakajima (2009). "A unique lantibiotic, thermophilin 1277, containing a disulfide bridge and two thioether bridges." Journal of Applied Microbiology **106**(3): 853-862.

Kaletta, C., K. D. Entian and G. Jung (1991). "Prepeptide sequence of cinnamycin (Ro 09-0198): the first structural gene of a duramycin-type lantibiotic." European Journal of Biochemistry / FEBS **199**(2): 411-415.

Karakousis, G., N. Ye, Z. Li, S. K. Chiu, G. Reddy and C. M. Radding (1998). "The beta protein of phage lambda binds preferentially to an intermediate in DNA renaturation." Journal of Molecular Biology **276**(4): 721-731.

Kato, J. Y., I. Miyahisa, M. Mashiko, Y. Ohnishi and S. Horinouchi (2004). "A single target is sufficient to account for the biological effects of the A-factor receptor protein of *Streptomyces griseus*." Journal of Bacteriology **186**(7): 2206-2211.

Keller, L. and M. G. Surette (2006). "Communication in bacteria: an ecological and evolutionary perspective." Nature reviews. Microbiology **4**(4): 249-258.

Kellner, R., G. Jung, T. Horner, H. Zahner, N. Schnell, K. D. Entian and F. Gotz (1988). "Gallidermin: a new lanthionine-containing polypeptide antibiotic." European journal of Biochemistry / FEBS **177**(1): 53-59.

- Kelly, W. L., L. Pan and C. Li (2009). "Thiostrepton biosynthesis: prototype for a new family of bacteriocins." Journal of the American Chemical Society **131**(12): 4327-4334.
- Kiesau, P., U. Eikmanns, Z. GutowskiEckel, S. Weber, M. Hammelmann and K. D. Entian (1997). "Evidence for a multimeric subtilin synthetase complex." Journal of Bacteriology **179**(5): 1475-1481.
- Kieser, T., M. J. Bibb, M. J. Buttner, K. F. Chater and D. A. Hopwood (2000). Practical Streptomyces Genetics, John Innes Foundation.
- Klaenhammer, T. R. (1993). "Genetics of bacteriocins produced by lactic acid bacteria." FEMS Microbiology Reviews **12**(1-3): 39-85.
- Kleerebezem, M. (2004). "Quorum sensing control of lantibiotic production; nisin and subtilin autoregulate their own biosynthesis." Peptides **25**(9): 1405-1414.
- Klein, C. and K. D. Entian (1994). "Genes Involved in Self-Protection against the Lantibiotic Subtilin Produced by *Bacillus subtilis* Atcc-6633." Applied and Environmental Microbiology **60**(8): 2793-2801.
- Klein, C., C. Kaletta, N. Schnell and K. D. Entian (1992). "Analysis of genes involved in biosynthesis of the lantibiotic subtilin." Applied and Environmental Microbiology **58**(1): 132-142.
- Klusens, L. D., A. Kuipers, R. Rink, E. de Boef, S. Fekken, A. J. M. Driessen, O. P. Kuipers and G. N. Moll (2005). "Post-translational modification of therapeutic peptides by NisB, the dehydratase of the lantibiotic nisin." Biochemistry **44**(38): 12827-12834.
- Kodani, S., M. E. Hudson, M. C. Durrant, M. J. Buttner, J. R. Nodwell and J. M. Willey (2004). "The SapB morphogen is a lantibiotic-like peptide derived from the product of the developmental gene *ramS* in *Streptomyces coelicolor*." Proceedings of the National Academy of Sciences of the United States of America **101**(31): 11448-11453.
- Kodani, S., M. A. Lodato, M. C. Durrant, F. Picart and J. M. Willey (2005). "SapT, a lanthionine-containing peptide involved in aerial hyphae formation in the streptomycetes." Molecular Microbiology **58**(5): 1368-1380.
- Komiyama, K., K. Otaguro, T. Segawa, K. Shiomi, H. Yang, Y. Takahashi, M. Hayashi, T. Otani and S. Omura (1993). "A new antibiotic, cypemycin. Taxonomy, fermentation, isolation and biological characteristics." The Journal of Antibiotics **46**(11): 1666-1671.
- Koponen, O., T. M. Takala, U. Saarela, M. Qiao and P. E. Saris (2004). "Distribution of the NisI immunity protein and enhancement of nisin activity by the lipid-free NisI." FEMS microbiology letters **231**(1): 85-90.
- Koponen, O., M. Tolonen, M. Q. Qiao, G. Wahlstrom, J. Helin and P. E. J. Saris (2002). "NisB is required for the dehydration and NisC for the lanthionine formation in the post-translational modification of nisin." Microbiology-Sgm **148**: 3561-3568.

Krogh, A., B. Larsson, G. von Heijne and E. L. Sonnhammer (2001). "Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes." Journal of Molecular Biology **305**(3): 567-580.

Kuipers, A., E. de Boef, R. Rink, S. Fekken, L. D. Kluskens, A. J. M. Driessen, K. Leenhouts, O. P. Kuipers and G. N. Moll (2004). "NisT, the transporter of the lantibiotic nisin, can transport fully modified, dehydrated, and unmodified prenisin and fusions of the leader peptide with non-lantibiotic peptides." Journal of Biological Chemistry **279**(21): 22176-22182.

Kuipers, A., J. Wierenga, R. Rink, L. D. Kluskens, A. J. M. Driessen, O. P. Kuipers and G. N. Moll (2006). "Sec-mediated transport of posttranslationally dehydrated peptides in *Lactococcus lactis*." Applied and Environmental Microbiology **72**(12): 7626-7633.

Kuipers, O. P., M. M. Beerthuyzen, P. G. de Ruyter, E. J. Luesink and W. M. de Vos (1995). "Autoregulation of nisin biosynthesis in *Lactococcus lactis* by signal transduction." Journal of Biological Chemistry **270**(45): 27299-27304.

Kuipers, O. P., M. M. Beerthuyzen, R. J. Siezen and W. M. De Vos (1993a). "Characterization of the nisin gene cluster *nisABTCIPR* of *Lactococcus lactis*. Requirement of expression of the *nisA* and *nisI* genes for development of immunity." European Journal of Biochemistry / FEBS **216**(1): 281-291.

Kuipers, O. P., H. S. Rollema, W. M. de Vos and R. J. Siezen (1993b). "Biosynthesis and secretion of a precursor of nisin Z by *Lactococcus lactis*, directed by the leader peptide of the homologous lantibiotic subtilin from *Bacillus subtilis*." FEBS Letters **330**(1): 23-27.

Kupke, T. and F. Gotz (1997). "*In vivo* reaction of affinity-tag-labelled epidermin precursor peptide with flavoenzyme EpiD." FEMS Microbiology Letters **153**(1): 25-32.

Kupke, T., C. Kempter, V. Gnau, G. Jung and F. Gotz (1994). "Mass spectroscopic analysis of a novel enzymatic reaction. Oxidative decarboxylation of the lantibiotic precursor peptide EpiA catalyzed by the flavoprotein EpiD." Journal of Biological Chemistry **269**(8): 5653-5659.

Lakey, J. H., E. J. Lea, B. A. Rudd, H. M. Wright and D. A. Hopwood (1983). "A new channel-forming antibiotic from *Streptomyces coelicolor* A3(2) which requires calcium for its activity." Journal of General Microbiology **129**(12): 3565-3573.

Lane, W. J. and S. A. Darst (2006). "The structural basis for promoter -35 element recognition by the group IV sigma factors." PLoS biology **4**(9): e269.

Laureti, L., L. Song, S. Huang, C. Corre, P. Leblond, G. L. Challis and B. Aigle (2011). "Identification of a bioactive 51-membered macrolide complex by activation of a silent polyketide synthase in *Streptomyces ambofaciens*." Proceedings of the National Academy of Sciences of the United States of America **108**(15): 6258-6263.

- Lautru, S. and G. L. Challis (2004). "Substrate recognition by nonribosomal peptide synthetase multi-enzymes." Microbiology **150**(Pt 6): 1629-1636.
- Lazzarini, A., L. Gastaldo, G. Candiani, I. Ciciliato, D. Losi, F. Marinelli, E. Selva and F. Parenti (2005). Antibiotic 107891, its factors A1 and A2, pharmaceutically acceptable salts and compositions and use thereof. V. P. Inc. **WO/2005/014628**.
- Lee, M. D. (2003). Antibiotics from *Microbispora*. E. T. US, Inc. (Mountain View, CA). US.
- Leskiw, B. K., E. J. Lawlor, J. M. Fernandez-Abalos and K. F. Chater (1991). "TTA codons in some genes prevent their expression in a class of developmental, antibiotic-negative, *Streptomyces* mutants." Proceedings of the National Academy of Sciences of the United States of America **88**(6): 2461-2465.
- Levengood, M. R., G. C. Patton and W. A. D. van der Donk (2007). "The leader peptide is not required for post-translational modification by lactacin 481 synthetase." Journal of the American Chemical Society **129**(34): 10314-+.
- Li, B., D. Sher, L. Kelly, Y. Shi, K. Huang, P. J. Knerr, I. Joewono, D. Rusch, S. W. Chisholm and W. A. van der Donk (2010). "Catalytic promiscuity in the biosynthesis of cyclic peptide secondary metabolites in planktonic marine cyanobacteria." Proceedings of the National Academy of Sciences of the United States of America **107**(23): 10430-10435.
- Li, B. and W. A. van der Donk (2007). "Identification of essential catalytic residues of the cyclase NisC involved in the biosynthesis of nisin." Journal of Biological Chemistry **282**(29): 21169-21175.
- Li, B., J. P. Yu, J. S. Brunzelle, G. N. Moll, W. A. van der Donk and S. K. Nair (2006). "Structure and mechanism of the lantibiotic cyclase involved in nisin biosynthesis." Science **311**(5766): 1464-1467.
- Li, C. and W. L. Kelly (2010). "Recent advances in thiopeptide antibiotic biosynthesis." Natural Product Reports **27**(2): 153-164.
- Li, J. W. and J. C. Vederas (2009). "Drug discovery and natural products: end of an era or an endless frontier?" Science **325**(5937): 161-165.
- Liu, G., J. Zhong, J. Q. Ni, M. L. Chen, H. J. Xiao and L. D. Huan (2009). "Characteristics of the bovicin HJ50 gene cluster in *Streptococcus bovis* HJ50." Microbiology-Sgm **155**: 584-593.
- Liu, W. and J. N. Hansen (1992). "Enhancement of the chemical and antimicrobial properties of subtilin by site-directed mutagenesis." The Journal of Biological Chemistry **267**(35): 25078-25085.
- Lonetto, M. A., K. L. Brown, K. E. Rudd and M. J. Buttner (1994). "Analysis of the *Streptomyces coelicolor sigE* gene reveals the existence of a subfamily of eubacterial RNA polymerase sigma factors involved in the regulation of extracytoplasmic functions."

- Proceedings of the National Academy of Sciences of the United States of America **91**(16): 7573-7577.
- Lopez, D., M. A. Fischbach, F. Chu, R. Losick and R. Kolter (2009). "Structurally diverse natural products that cause potassium leakage trigger multicellularity in *Bacillus subtilis*." Proceedings of the National Academy of Sciences of the United States of America **106**(1): 280-285.
- Losi, D., L. Cavaletti, A. Lazzarini, G. Candiani, F. Castiglione and F. Marinelli (2004). Antibiotic 97518, pharmaceutically acceptable salts and compositions, and use thereof. Italy.
- Luo, Y. and J. D. Helmann (2009). "Extracytoplasmic function sigma factors with overlapping promoter specificity regulate sublancin production in *Bacillus subtilis*." Journal of Bacteriology **191**(15): 4951-4958.
- Macneil, D. J., K. M. Gewain, C. L. Ruby, G. Dezeny, P. H. Gibbons and T. Macneil (1992). "Analysis of *Streptomyces avermitilis* Genes Required for Avermectin Biosynthesis Utilizing a Novel Integration Vector." Gene **111**(1): 61-68.
- Madan Babu, M. and K. Sankaran (2002). "DOLOP--database of bacterial lipoproteins." Bioinformatics **18**(4): 641-643.
- Maffioli, S. I., C. Brunati, D. Potenza, F. Vasile and S. Donadio (2010). Lantibiotic carboxamide derivatives with enhanced antibacterial activity. S. P. Inc.
- Maffioli, S. I., D. Potenza, F. Vasile, M. De Matteo, M. Sosio, B. Marsiglia, V. Rizzo, C. Scolastico and S. Donadio (2009). "Structure Revision of the Lantibiotic 97518." Journal of Natural Products.
- Makino, A., T. Baba, K. Fujimoto, K. Iwamoto, Y. Yano, N. Terada, S. Ohno, S. B. Sato, A. Ohta, M. Umeda, K. Matsuzaki and T. Kobayashi (2003). "Cinnamycin (Ro 09-0198) promotes cell binding and toxicity by inducing transbilayer lipid movement." The Journal of Biological Chemistry **278**(5): 3204-3209.
- Malhotra, A., E. Severinova and S. A. Darst (1996). "Crystal structure of a sigma 70 subunit fragment from *E. coli* RNA polymerase." Cell **87**(1): 127-136.
- Malpartida, F. and D. A. Hopwood (1986). "Physical and genetic characterisation of the gene cluster for the antibiotic actinorhodin in *Streptomyces coelicolor* A3(2)." Molecular & General Genetics : **MGG 205**(1): 66-73.
- Marcone, G. L., F. Beltrametti, E. Binda, L. Carrano, L. Foulston, A. Hesketh, M. Bibb and F. Marinelli (2010a). "Novel mechanism of glycopeptide resistance in the A40926 producer *Nonomuraea* sp. ATCC 39727." Antimicrobial Agents and Chemotherapy **54**(6): 2465-2472.

Marcone, G. L., L. Carrano, F. Marinelli and F. Beltrametti (2010b). "Protoplast preparation and reversion to the normal filamentous growth in antibiotic-producing uncommon actinomycetes." Journal of Antibiotics **63**(2): 83-88.

Marcone, G. L., L. Foulston, E. Binda, F. Marinelli, M. Bibb and F. Beltrametti (2010c). "Methods for the genetic manipulation of *Nonomuraea* sp. ATCC 39727." Journal of Industrial Microbiology & Biotechnology **37**(10): 1097-1103.

Marinelli, F. (2009). "Chapter 2. From microbial products to novel drugs that target a multitude of disease indications." Methods in Enzymology **458**: 29-58.

Martin, N. I., T. Sprules, M. R. Carpenter, P. D. Cotter, C. Hill, R. P. Ross and J. C. Vederas (2004). "Structural characterization of lacticin 3147, a two-peptide lantibiotic with synergistic activity." Biochemistry **43**(11): 3049-3056.

Matsushima, P. and R. H. Baltz (1996). "A gene cloning system for '*Streptomyces toyocaensis*'." Microbiology **142** (Pt 2): 261-267.

Mazza, P., P. Monciardini, L. Cavaletti, M. Sosio and S. Donadio (2003). "Diversity of *Actinoplanes* and related genera isolated from an Italian soil." Microbial Ecology **45**(4): 362-372.

McClerren, A. L., L. E. Cooper, C. Quan, P. M. Thomas, N. L. Kelleher and W. A. van der Donk (2006). "Discovery and in vitro biosynthesis of haloduracin, a two-component lantibiotic." Proceedings of the National Academy of Sciences of the United States of America **103**(46): 17243-17248.

McLeod, M. P., R. L. Warren, W. W. Hsiao, N. Araki, M. Myhre, C. Fernandes, D. Miyazawa, W. Wong, A. L. Lillquist, D. Wang, M. Dosanjh, H. Hara, A. Petrescu, R. D. Morin, G. Yang, J. M. Stott, J. E. Schein, H. Shin, D. Smailus, A. S. Siddiqui, M. A. Marra, S. J. Jones, R. Holt, F. S. Brinkman, K. Miyauchi, M. Fukuda, J. E. Davies, W. W. Mohn and L. D. Eltis (2006). "The complete genome of *Rhodococcus* sp. RHA1 provides insights into a catabolic powerhouse." Proceedings of the National Academy of Sciences of the United States of America **103**(42): 15582-15587.

Meindl, K., T. Schmiederer, K. Schneider, A. Reicke, D. Butz, S. Keller, H. Guhring, L. Vertesy, J. Wink, H. Hoffmann, M. Bronstrup, G. M. Sheldrick and R. D. Sussmuth (2010). "Labyrinthopeptins: a new class of carbacyclic lantibiotics." Angewandte Chemie **49**(6): 1151-1154.

Mertz, F. P. (1994). "*Planomonospora alba* Sp-Nov and *Planomonospora sphaerica* Sp-Nov, 2 New Species Isolated from Soil by Baiting Techniques." International Journal of Systematic Bacteriology **44**(2): 274-281.

Meyer, C., G. Bierbaum, C. Heidrich, M. Reis, J. Suling, M. I. Iglesias-Wind, C. Kempter, E. Molitor and H. G. Sahl (1995). "Nucleotide sequence of the lantibiotic Pep5 biosynthetic gene cluster and functional analysis of PepP and PepC. Evidence for a role of PepC in thioether formation." European journal of biochemistry / FEBS **232**(2): 478-489.

Meyer, H. E., M. Heber, B. Eisermann, H. Korte, J. W. Metzger and G. Jung (1994). "Sequence analysis of lantibiotics: chemical derivatization procedures allow a fast access to complete Edman degradation." Analytical Biochemistry **223**(2): 185-190.

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 (2005). "Daptomycin biosynthesis in *Streptomyces roseosporus*: cloning and analysis of the gene cluster and revision of peptide stereochemistry." Microbiology **151**(Pt 5): 1507-1523.

Miao, V. and J. Davies (2010). "Actinobacteria: the good, the bad, and the ugly." Antonie van Leeuwenhoek **98**(2): 143-150.

Miller, L. M., C. Chatterjee, W. A. van der Donk and N. L. Kelleher (2006). "The dehydratase activity of lactacin 481 synthetase is highly processive." Journal of the American Chemical Society **128**(5): 1420-1421.

Missiakas, D. and S. Raina (1998). "The extracytoplasmic function sigma factors: role and regulation." Molecular Microbiology **28**(6): 1059-1066.

Miyadoh, S. (1997). Atlas of Actinomycetes, The Society of Actinomycetes Japan.

Moller, S., M. D. Croning and R. Apweiler (2001). "Evaluation of methods for the prediction of membrane spanning regions." Bioinformatics **17**(7): 646-653.

Mulders, J. W., I. J. Boerrigter, H. S. Rollema, R. J. Siezen and W. M. de Vos (1991). "Identification and characterization of the lantibiotic nisin Z, a natural nisin variant." European Journal of Biochemistry / FEBS **201**(3): 581-584.

Muller, W. M., T. Schmiederer, P. Ensle and R. D. Sussmuth (2010). "*In vitro* biosynthesis of the prepeptide of type-III lantibiotic labyrinthopeptin A2 including formation of a C-C bond as a post-translational modification." Angewandte Chemie **49**(13): 2436-2440.

Muthaiyan, A., J. A. Silverman, R. K. Jayaswal and B. J. Wilkinson (2008). "Transcriptional profiling reveals that daptomycin induces the *Staphylococcus aureus* cell wall stress stimulon and genes responsive to membrane depolarization." Antimicrobial Agents and Chemotherapy **52**(3): 980-990.

Nakajima, Y., V. Kitpreechavanich, K.-i. Suzuki and T. Kudo (1999). "*Microbispora corallina* sp. nov., a new species of the genus *Microbispora* isolated from Thai soil." International Journal of Systematic Bacteriology **49**(4): 1761-1767.

Neis, S., G. Bierbaum, M. Josten, U. Pag, C. Kempter, G. Jung and H. G. Sahl (1997). "Effect of leader peptide mutations on biosynthesis of the lantibiotic Pep5." FEMS Microbiology Letters **149**(2): 249-255.

Nett, M., H. Ikeda and B. S. Moore (2009). "Genomic basis for natural product biosynthetic diversity in the actinomycetes." Natural Product Reports **26**(11): 1362-1384.

Newman, D. J. and G. M. Cragg (2007). "Natural products as sources of new drugs over the last 25 years." Journal of Natural Products **70**(3): 461-477.

Newman, D. J., G. M. Cragg and K. M. Snader (2000). "The influence of natural products upon drug discovery." Natural Product Reports **17**(3): 215-234.

Novakova, R., A. Rehakova, P. Kutas, L. Feckova and J. Kormanec (2011). "The role of two SARP family transcriptional regulators in regulation of the auricin gene cluster in *Streptomyces aureofaciens* CCM 3239." Microbiology **157**(Pt 6): 1629-1639.

O'Neill, A. J., K. Miller, B. Oliva and I. Chopra (2004). "Comparison of assays for detection of agents causing membrane damage in *Staphylococcus aureus*." The Journal of Antimicrobial Chemotherapy **54**(6): 1127-1129.

O'Sullivan, D. J. and J.-H. Lee (2011). Lantibiotics and uses thereof, Regents of the University of Minnesota. **US 7.960,505 B2**.

Okeley, N. M., M. Paul, J. P. Stasser, N. Blackburn and W. A. van der Donk (2003). "SpaC and NisC, the cyclases involved in subtilin and nisin biosynthesis, are zinc proteins." Biochemistry **42**(46): 13613-13624.

Okesli, A., L. E. Cooper, E. J. Fogle and W. A. van der Donk (2011). "Nine Post-translational Modifications during the Biosynthesis of Cinnamycin." Journal of the American Chemical Society **133**(34): 13753-13760.

Okuda, K., S. Yanagihara, T. Sugayama, T. Zendo, J. Nakayama and K. Sonomoto (2010). "Functional significance of the E loop, a novel motif conserved in the lantibiotic immunity ATP-binding cassette transport systems." Journal of Bacteriology **192**(11): 2801-2808.

Oman, T. J., J. M. Boettcher, H. Wang, X. N. Okalibe and W. A. van der Donk (2011). "Sublancin is not a lantibiotic but an S-linked glycopeptide." Nature Chemical Biology **7**(2): 78-80.

Onaka, H., M. Nakaho, K. Hayashi, Y. Igarashi and T. Furumai (2005). "Cloning and characterization of the goadsporin biosynthetic gene cluster from *Streptomyces* sp. TP-A0584." Microbiology **151**(Pt 12): 3923-3933.

Otto, M., A. Peschel and F. Gotz (1998). "Producer self-protection against the lantibiotic epidermin by the ABC transporter EpiFEG of *Staphylococcus epidermidis* Tu3298." FEMS Microbiology Letters **166**(2): 203-211.

Pag, U., C. Heidrich, G. Bierbaum and H. G. Sahl (1999). "Molecular analysis of expression of the lantibiotic pep5 immunity phenotype." Applied and Environmental Microbiology **65**(2): 591-598.

Paget, M. S., J. B. Bae, M. Y. Hahn, W. Li, C. Kleanthous, J. H. Roe and M. J. Buttner (2001). "Mutational analysis of RsrA, a zinc-binding anti-sigma factor with a thiol-disulphide redox switch." Molecular Microbiology **39**(4): 1036-1047.

Paget, M. S., L. Chamberlin, A. Atrih, S. J. Foster and M. J. Buttner (1999). "Evidence that the extracytoplasmic function sigma factor sigmaE is required for normal cell wall structure in *Streptomyces coelicolor* A3(2)." Journal of Bacteriology **181**(1): 204-211.

Paik, S. H., A. Chakicherla and J. N. Hansen (1998). "Identification and characterization of the structural and transporter genes for, and the chemical and biological properties of, sublancin 168, a novel lantibiotic produced by *Bacillus subtilis* 168." The Journal of Biological Chemistry **273**(36): 23134-23142.

Patton, G. C., M. Paul, L. E. Cooper, C. Chatterjee and W. A. van der Donk (2008). "The importance of the leader sequence for directing lanthionine formation in lactacin 481." Biochemistry **47**(28): 7342-7351.

Patton, G. C. and W. A. van der Donk (2005). "New developments in lantibiotic biosynthesis and mode of action." Current Opinion in Microbiology **8**(5): 543-551.

Paul, M., G. C. Patton and W. A. van der Donk (2007). "Mutants of the zinc ligands of lactacin 481 synthetase retain dehydration activity but have impaired cyclization activity." Biochemistry **46**(21): 6268-6276.

Payne, D. J., M. N. Gwynn, D. J. Holmes and D. L. Pompliano (2007). "Drugs for bad bugs: confronting the challenges of antibacterial discovery." Nature reviews. Drug discovery **6**(1): 29-40.

Peschel, A. and F. Gotz (1996). "Analysis of the *Staphylococcus epidermidis* genes epiF, -E, and -G involved in epidermin immunity." Journal of Bacteriology **178**(2): 531-536.

Pfefferle, C., U. Theobald, H. Gurtler and H. Fiedler (2000). "Improved secondary metabolite production in the genus *Streptosporangium* by optimization of the fermentation conditions." Journal of Biotechnology **80**(2): 135-142.

Qiao, M. and P. E. Saris (1996). "Evidence for a role of NisT in transport of the lantibiotic nisin produced by *Lactococcus lactis* N8." FEMS microbiology letters **144**(1): 89-93.

Ra, R., M. M. Beerthuyzen, W. M. de Vos, P. E. Saris and O. P. Kuipers (1999). "Effects of gene disruptions in the nisin gene cluster of *Lactococcus lactis* on nisin production and producer immunity." Microbiology **145** (Pt 5): 1227-1233.

Rao, D. K. and P. Kaur (2008). "The Q-loop of DrrA is involved in producing the closed conformation of the nucleotide binding domains and in transduction of conformational changes between DrrA and DrrB." Biochemistry **47**(9): 3038-3050.

Rice, L. B. (2009). "The clinical consequences of antimicrobial resistance." Current Opinion in Microbiology **12**(5): 476-481.

Rigali, S., F. Titgemeyer, S. Barends, S. Mulder, A. W. Thomae, D. A. Hopwood and G. P. van Wezel (2008). "Feast or famine: the global regulator DasR links nutrient stress to antibiotic production by *Streptomyces*." EMBO reports **9**(7): 670-675.

Rink, R., L. D. Kluskens, A. Kuipers, A. J. Driessen, O. P. Kuipers and G. N. Moll (2007a). "NisC, the cyclase of the lantibiotic nisin, can catalyze cyclization of designed nonlantibiotic peptides." Biochemistry **46**(45): 13179-13189.

Rink, R., A. Kuipers, E. de Boef, K. J. Leenhouts, A. J. M. Driessen, G. N. Moll and O. P. Kuipers (2005). "Lantibiotic structures as guidelines for the design of peptides that can be modified by lantibiotic enzymes." Biochemistry **44**(24): 8873-8882.

Rink, R., J. Wierenga, A. Kuipers, L. D. Kluskens, A. J. M. Driessen, O. P. Kuipers and G. N. Moll (2007b). "Dissection and modulation of the four distinct activities of nisin by mutagenesis of rings A and B and by C-terminal truncation." Applied and Environmental Microbiology **73**(18): 5809-5816.

Rink, R., J. Wierenga, A. Kuipers, L. D. Muskens, A. J. M. Driessen, O. P. Kuipers and G. N. Moll (2007c). "Production of dehydroamino acid-containing peptides by *Lactococcus lactis*." Applied and Environmental Microbiology **73**(6): 1792-1796.

Roes, M. and P. R. Meyers (2008). "*Nonomuraea candida* sp. nov., a new species from South African soil." Antonie van Leeuwenhoek **93**(1-2): 133-139.

Rogers, L. A. (1928). "The Inhibiting Effect of *Streptococcus lactis* on *Lactobacillus bulgaricus*." Journal of Bacteriology **16**(5): 321-325.

Ross, K. F., C. W. Ronson and J. R. Tagg (1993). "Isolation and characterization of the lantibiotic salivaricin A and its structural gene *salA* from *Streptococcus salivarius* 20P3." Applied and Environmental Microbiology **59**(7): 2014-2021.

Rudd, B. A. and D. A. Hopwood (1980). "A pigmented mycelial antibiotic in *Streptomyces coelicolor*: control by a chromosomal gene cluster." Journal of general microbiology **119**(2): 333-340.

Rutherford, K., J. Parkhill, J. Crook, T. Horsnell, P. Rice, M. A. Rajandream and B. Barrell (2000). "Artemis: sequence visualization and annotation." Bioinformatics **16**(10): 944-945.

Ryan, M. P., R. W. Jack, M. Josten, H. G. Sahl, G. Jung, R. P. Ross and C. Hill (1999). "Extensive post-translational modification, including serine to D-alanine conversion, in the two-component lantibiotic, lactacin 3147." The Journal of Biological Chemistry **274**(53): 37544-37550.

Sahl, H. G. and G. Bierbaum (1998). "Lantibiotics: biosynthesis and biological activities of uniquely modified peptides from gram-positive bacteria." Annual Review of Microbiology **52**: 41-79.

Sahl, H. G., R. W. Jack and G. Bierbaum (1995). "Biosynthesis and biological activities of lantibiotics with unique post-translational modifications." European Journal of Biochemistry / FEBS **230**(3): 827-853.

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

Schmitz, S., A. Hoffmann, C. Szekat, B. Rudd and G. Bierbaum (2006). "The lantibiotic mersacidin is an autoinducing peptide." Applied and Environmental Microbiology **72**(11): 7270-7277.

Schneewind, O., A. Fowler and K. F. Faull (1995). "Structure of the cell wall anchor of surface proteins in *Staphylococcus aureus*." Science **268**(5207): 103-106.

Schneider, E. and S. Hunke (1998). "ATP-binding-cassette (ABC) transport systems: functional and structural aspects of the ATP-hydrolyzing subunits/domains." FEMS Microbiology Reviews **22**(1): 1-20.

Schneider, T., T. Kruse, R. Wimmer, I. Wiedemann, V. Sass, U. Pag, A. Jansen, A. K. Nielsen, P. H. Mygind, D. S. Raventos, S. Neve, B. Ravn, A. M. Bonvin, L. De Maria, A. S. Andersen, L. K. Gammelgaard, H. G. Sahl and H. H. Kristensen (2010). "Plectasin, a fungal defensin, targets the bacterial cell wall precursor Lipid II." Science **328**(5982): 1168-1172.

Schneider, T. R., J. Karcher, E. Pohl, P. Lubini and G. M. Sheldrick (2000). "Ab initio structure determination of the lantibiotic mersacidin." Acta crystallographica. Section D, Biological crystallography **56**(Pt 6): 705-713.

Schobel, S., S. Zellmeier, W. Schumann and T. Wiegert (2004). "The *Bacillus subtilis* sigmaW anti-sigma factor RsiW is degraded by intramembrane proteolysis through YluC." Molecular Microbiology **52**(4): 1091-1105.

Shank, E. A. and R. Kolter (2009). "New developments in microbial interspecies signaling." Current Opinion in Microbiology **12**(2): 205-214.

Shimizu, K. and T. Masaki (2004). Peptidic antibiotic. K. M.-c. C.-k. ASAHI KASEI PHARMA CORPORATION; 9-1, Tokyo 101-8481.

Siegers, K. and K. D. Entian (1995). "Genes Involved in Immunity to the Lantibiotic Nisin Produced by *Lactococcus lactis* 6f3." Applied and Environmental Microbiology **61**(3): 1082-1089.

Siegers, K., S. Heinzmann and K. D. Entian (1996). "Biosynthesis of lantibiotic nisin - Posttranslational modification of its prepeptide occurs at a multimeric membrane-associated lanthionine synthetase complex." Journal of Biological Chemistry **271**(21): 12294-12301.

Skaugen, M. and I. F. Nes (1994). "Transposition in *Lactobacillus sake* and its abolition of lactocin S production by insertion of IS1163, a new member of the IS3 family." Applied and Environmental Microbiology **60**(8): 2818-2825.

Smith, L., H. Hasper, E. Breukink, J. Novak, J. Cerkasov, J. D. Hillman, S. Wilson-Stanford and R. S. Orugunty (2008). "Elucidation of the antimicrobial mechanism of mutacin 1140." Biochemistry **47**(10): 3308-3314.

Smith, L., J. Novak, J. Rocca, S. McClung, J. D. Hillman and A. S. Edison (2000). "Covalent structure of mutacin 1140 and a novel method for the rapid identification of lantibiotics." European Journal of Biochemistry / FEBS **267**(23): 6810-6816.

Somma, S., W. Merati and F. Parenti (1977). "Gardimycin, a new antibiotic inhibiting peptidoglycan synthesis." Antimicrobial Agents and Chemotherapy **11**(3): 396-401.

Sosio, M., S. Stinchi, F. Beltrametti, A. Lazzarini and S. Donadio (2003). "The gene cluster for the biosynthesis of the glycopeptide antibiotic A40926 by *Nonomuraea* species." Chemistry & Biology **10**(6): 541-549.

Staron, A., H. J. Sofia, S. Dietrich, L. E. Ulrich, H. Liesegang and T. Mascher (2009). "The third pillar of bacterial signal transduction: classification of the extracytoplasmic function (ECF) sigma factor protein family." Molecular Microbiology **74**(3): 557-581.

Steen, H. and M. Mann (2004). "The ABC's (and XYZ's) of peptide sequencing." Nature Reviews Molecular Cell Biology **5**(9): 699-711.

Steenbergen, J. N., J. Alder, G. M. Thorne and F. P. Tally (2005). "Daptomycin: a lipopeptide antibiotic for the treatment of serious Gram-positive infections." The Journal of antimicrobial chemotherapy **55**(3): 283-288.

Stein, T., S. Borchert, B. Conrad, J. Feesche, B. Hofemeister, J. Hofemeister and K. D. Entian (2002a). "Two different lantibiotic-like peptides originate from the ericin gene cluster of *Bacillus subtilis* A1/3." Journal of bacteriology **184**(6): 1703-1711.

Stein, T., S. Borchert, P. Kiesau, S. Heinzmann, S. Kloss, C. Klein, M. Helfrich and K. D. Entian (2002b). "Dual control of subtilin biosynthesis and immunity in *Bacillus subtilis*." Molecular Microbiology **44**(2): 403-416.

Stein, T. and K. D. Entian (2002). "Maturation of the lantibiotic subtilin: matrix-assisted laser desorption/ionization time-of-flight mass spectrometry to monitor precursors and their proteolytic processing in crude bacterial cultures." Rapid communications in mass spectrometry : RCM **16**(2): 103-110.

Stein, T., S. Heinzmann, I. Solovieva and K. D. Entian (2003). "Function of *Lactococcus lactis* nisin immunity genes *nisl* and *nisFEG* after coordinated expression in the surrogate host *Bacillus subtilis*." The Journal of Biological Chemistry **278**(1): 89-94.

- Sternberg, M., C. Y. Kim and F. J. Schwende (1975). "Lysinoalanine: presence in foods and food ingredients." Science **190**(4218): 992-994.
- Stevens, A. M., K. M. Dolan and E. P. Greenberg (1994). "Synergistic binding of the *Vibrio fischeri* LuxR transcriptional activator domain and RNA polymerase to the lux promoter region." Proceedings of the National Academy of Sciences of the United States of America **91**(26): 12619-12623.
- Stevens, K. A., B. W. Sheldon, N. A. Klapes and T. R. Klaenhammer (1991). "Nisin treatment for inactivation of *Salmonella* species and other Gram-negative bacteria." Applied and Environmental Microbiology **57**(12): 3613-3615.
- Stinchi, S., S. Azimonti, S. Donadio and M. Sosio (2003). "A gene transfer system for the glycopeptide producer *Nonomuraea* sp. ATCC39727." FEMS Microbiology Letters **225**(1): 53-57.
- Stutzman-Engwall, K. J. and C. R. Hutchinson (1989). "Multigene families for anthracycline antibiotic production in *Streptomyces peucetius*." Proceedings of the National Academy of Sciences of the United States of America **86**(9): 3135-3139.
- Subramoni, S. and V. Venturi (2009). "LuxR-family 'solos': bachelor sensors/regulators of signalling molecules." Microbiology **155**(Pt 5): 1377-1385.
- Sutcliffe, I. C. and D. J. Harrington (2002). "Pattern searches for the identification of putative lipoprotein genes in Gram-positive bacterial genomes." Microbiology **148**(Pt 7): 2065-2077.
- Suzuki, S. I., T. Okuda and S. Komatsubara (2001). "Selective isolation and distribution of the genus *Planomonospora* in soils." Canadian Journal of Microbiology **47**(3): 253-263.
- Szekat, C., R. W. Jack, D. Skutlarek, H. Farber and G. Bierbaum (2003). "Construction of an expression system for site-directed mutagenesis of the lantibiotic mersacidin." Applied and Environmental Microbiology **69**(7): 3777-3783.
- Technikova-Dobrova, Z., F. Damiano, S. M. Tredici, G. Vigliotta, R. di Summa, L. Palese, A. Abbrescia, N. Labonia, G. V. Gnoni and P. Alifano (2004). "Design of mineral medium for growth of *Actinomadura* sp. ATCC 39727, producer of the glycopeptide A40926: effects of calcium ions and nitrogen sources." Applied Microbiology and Biotechnology **65**(6): 671-677.
- Thiemann, J. E. (1970). International Journal of Systematic and Evolutionary Microbiology **30**: 345.
- Thiemann, J. E., H. Pagani and G. Beretta (1967). "A new genus of the actinoplanaceae: *Planomonospora* gen. nov." Gionale Microbiol. **15**: 27-28.
- Tiwari, K. and R. K. Gupta (2011). "Rare actinomycetes: a potential storehouse for novel antibiotics." Critical Reviews in Biotechnology.

Trefzer, A., S. Pelzer, J. Schimana, S. Stockert, C. Bihlmaier, H. P. Fiedler, K. Welzel, A. Vente and A. Bechthold (2002). "Biosynthetic gene cluster of simocyclinone, a natural multihybrid antibiotic." Antimicrobial Agents and Chemotherapy **46**(5): 1174-1182.

Turner, D. L., L. Brennan, H. E. Meyer, C. Lohaus, C. Siethoff, H. S. Costa, B. Gonzalez, H. Santos and J. E. Suarez (1999). "Solution structure of plantaricin C, a novel lantibiotic." European Journal of Biochemistry / FEBS **264**(3): 833-839.

Ueda, K., A. Yamashita, J. Ishikawa, M. Shimada, T. O. Watsuji, K. Morimura, H. Ikeda, M. Hattori and T. Beppu (2004). "Genome sequence of *Symbiobacterium thermophilum*, an uncultivable bacterium that depends on microbial commensalism." Nucleic Acids Research **32**(16): 4937-4944.

Uguen, M. and P. Uguen (2002). "The LcnC homologue cannot replace LctT in lacticin 481 export." FEMS Microbiology Letters **208**(1): 99-103.

Vaara, M. (2009). "New approaches in peptide antibiotics." Current Opinion in Pharmacology **9**(5): 571-576.

Van Bambeke, F., M. P. Mingeot-Leclercq, M. J. Struelens and P. M. Tulkens (2008). "The bacterial envelope as a target for novel anti-MRSA antibiotics." Trends in Pharmacological Sciences **29**(3): 124-134.

van der Meer, J. R., J. Polman, M. M. Beerthuyzen, R. J. Siezen, O. P. Kuipers and W. M. De Vos (1993). "Characterization of the *Lactococcus lactis* nisin A operon genes *nisP*, encoding a subtilisin-like serine protease involved in precursor processing, and *nisR*, encoding a regulatory protein involved in nisin biosynthesis." Journal of Bacteriology **175**(9): 2578-2588.

van der Meer, J. R., H. S. Rollema, R. J. Siezen, M. M. Beerthuyzen, O. P. Kuipers and W. M. de Vos (1994). "Influence of amino acid substitutions in the nisin leader peptide on biosynthesis and secretion of nisin by *Lactococcus lactis*." The Journal of Biological Chemistry **269**(5): 3555-3562.

van Heijenoort, J. (2001). "Recent advances in the formation of the bacterial peptidoglycan monomer unit." Natural Product Reports **18**(5): 503-519.

van Sappareo, H. B. V., P. J. Bakkes, G. N. Moll and A. J. M. Driessen (2008). "Distinct contributions of the nisin biosynthesis enzymes NisB and NisC and transporter NisT to prenisin production by *Lactococcus lactis*." Applied and Environmental Microbiology **74**(17): 5541-5548.

van Wageningen, A. M., P. N. Kirkpatrick, D. H. Williams, B. R. Harris, J. K. Kershaw, N. J. Lennard, M. Jones, S. J. Jones and P. J. Solenberg (1998). "Sequencing and analysis of genes involved in the biosynthesis of a vancomycin group antibiotic." Chemistry & Biology **5**(3): 155-162.

- Velasquez, J. E. and W. A. van der Donk (2011). "Genome mining for ribosomally synthesized natural products." Current Opinion in Chemical Biology **15**(1): 11-21.
- Ventura, M., C. Canchaya, A. Tauch, G. Chandra, G. F. Fitzgerald, K. F. Chater and D. van Sinderen (2007). "Genomics of Actinobacteria: Tracing the evolutionary history of an ancient phylura." Microbiology and Molecular Biology Reviews **71**(3): 495-+.
- Wagner, N., C. Osswald, R. Biener and D. Schwartz (2009). "Comparative analysis of transcriptional activities of heterologous promoters in the rare actinomycete *Actinoplanes friuliensis*." Journal of Biotechnology **142**(3-4): 200-204.
- Waksman, S. A. and H. B. Woodruff (1941). "*Actinomyces antibioticus*, a New Soil Organism Antagonistic to Pathogenic and Non-pathogenic Bacteria." Journal of Bacteriology **42**(2): 231-249.
- Walsh, C. (2003). Antibiotics: actions, origins, resistance, ASM press.
- 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 (2001). "Tailoring enzymes that modify nonribosomal peptides during and after chain elongation on NRPS assembly lines." Current Opinion in Chemical Biology **5**(5): 525-534.
- Wang, L. and L. C. Vining (2003). "Control of growth, secondary metabolism and sporulation in *Streptomyces venezuelae* ISP5230 by *jadW*(1), a member of the *afsA* family of gamma-butyrolactone regulatory genes." Microbiology **149**(Pt 8): 1991-2004.
- Watve, M. G., R. Tickoo, M. M. Jog and B. D. Bhole (2001). "How many antibiotics are produced by the genus *Streptomyces*?" Archives of microbiology **176**(5): 386-390.
- Widdick, D. A., H. M. Dodd, P. Barraille, J. White, T. H. Stein, K. F. Chater, M. J. Gasson and M. J. Bibb (2003). "Cloning and engineering of the cinnamycin biosynthetic gene cluster from *Streptomyces cinnamoneus cinnamoneus* DSM 40005." Proceedings of the National Academy of Sciences of the United States of America **100**(7): 4316-4321.
- Wiedemann, I., T. Bottiger, R. R. Bonelli, T. Schneider, H. G. Sahl and B. Martinez (2006a). "Lipid II-based antimicrobial activity of the lantibiotic plantaricin C." Applied and Environmental Microbiology **72**(4): 2809-2814.
- Wiedemann, I., T. Bottiger, R. R. Bonelli, A. Wiese, S. O. Hagge, T. Gutschmann, U. Seydel, L. Deegan, C. Hill, P. Ross and H. G. Sahl (2006b). "The mode of action of the lantibiotic lactacin 3147 - a complex mechanism involving specific interaction of two peptides and the cell wall precursor lipid II." Molecular Microbiology **61**(2): 285-296.
- Wilkinson, C. J., Z. A. Hughes-Thomas, C. J. Martin, I. Bohm, T. Mironenko, M. Deacon, M. Wheatcroft, G. Wirtz, J. Staunton and P. F. Leadlay (2002). "Increasing the efficiency of heterologous promoters in actinomycetes." Journal of Molecular Microbiology and Biotechnology **4**(4): 417-426.

Willems, R. J., J. Top, M. van Santen, D. A. Robinson, T. M. Coque, F. Baquero, H. Grundmann and M. J. Bonten (2005). "Global spread of vancomycin-resistant *Enterococcus faecium* from distinct nosocomial genetic complex." Emerging Infectious Diseases **11**(6): 821-828.

Willey, J. M. and W. A. van der Donk (2007). "Lantibiotics: Peptides of Diverse Structure and Function." Annual Review of Microbiology **61**(1): 477-501.

Wilson-Stanford, S., A. Kalli, K. Hakansson, J. Kastrantas, R. S. Orugunty and L. Smith (2009). "Oxidation of Lanthionines Renders the Lantibiotic Nisin Inactive." Applied and Environmental Microbiology **75**(5): 1381-1387.

Winter, J. M., S. Behnken and C. Hertweck (2011). "Genomics-inspired discovery of natural products." Current Opinion in Chemical Biology **15**(1): 22-31.

Wirawan, R. E., N. A. Kleese, R. W. Jack and J. R. Tagg (2006). "Molecular and genetic characterization of a novel nisin variant produced by *Streptococcus uberis*." Applied and Environmental Microbiology **72**(2): 1148-1156.

Woodford, N. and D. M. Livermore (2009). "Infections caused by Gram-positive bacteria: a review of the global challenge." The Journal of Infection **59 Suppl 1**: S4-16.

Wright, L. F. and D. A. Hopwood (1976a). "Actinorhodin is a chromosomally-determined antibiotic in *Streptomyces coelicolor* A3(2)." Journal of General Microbiology **96**(2): 289-297.

Wright, L. F. and D. A. Hopwood (1976b). "Identification of the antibiotic determined by the SCP1 plasmid of *Streptomyces coelicolor* A3(2)." Journal of general microbiology **95**(1): 96-106.

Wyszynski, F. J., A. R. Hesketh, M. J. Bibb and B. G. Davis (2010). "Dissecting tunicamycin biosynthesis by genome mining: cloning and heterologous expression of a minimal gene cluster." Chemical Science **1**(5): 581-589.

Xiao, H., X. Chen, M. Chen, S. Tang, X. Zhao and L. Huan (2004). "Bovicin HJ50, a novel lantibiotic produced by *Streptococcus bovis* HJ50." Microbiology **150**(Pt 1): 103-108.

Xie, L. L., L. M. Miller, C. Chatterjee, O. Averin, N. L. Kelleher and W. A. van der Donk (2004). "Lacticin 481: In vitro reconstitution of lantibiotic synthetase activity." Science **303**(5658): 679-681.

Xie, L. L. and W. A. van der Donk (2004). "Post-translational modifications during lantibiotic biosynthesis." Current Opinion in Chemical Biology **8**(5): 498-507.

Yamazaki, H., A. Tomono, Y. Ohnishi and S. Horinouchi (2004). "DNA-binding specificity of AdpA, a transcriptional activator in the A-factor regulatory cascade in *Streptomyces griseus*." Molecular Microbiology **53**(2): 555-572.

Yanai, K., T. Murakami and M. Bibb (2006). "Amplification of the entire kanamycin biosynthetic gene cluster during empirical strain improvement of *Streptomyces kanamyceticus*." Proceedings of the National Academy of Sciences of the United States of America **103**(25): 9661-9666.

Yim, G., H. H. Wang and J. Davies (2007). "Antibiotics as signalling molecules." Philosophical transactions of the Royal Society of London. Series B, Biological sciences **362**(1483): 1195-1200.

Young, J. and I. B. Holland (1999). "ABC transporters: bacterial exporters-revisited five years on." Biochimica et Biophysica Acta **1461**(2): 177-200.

Young, M., V. Artsatbanov, H. R. Beller, G. Chandra, K. F. Chater, L. G. Dover, E. B. Goh, T. Kahan, A. S. Kaprelyants, N. Kyrpides, A. Lapidus, S. R. Lowry, A. Lykidis, J. Mahillon, V. Markowitz, K. Mavromatis, G. V. Mukamolova, A. Oren, J. S. Rokem, M. C. Smith, D. I. Young and C. L. Greenblatt (2010). "Genome sequence of the Fleming strain of *Micrococcus luteus*, a simple free-living actinobacterium." Journal of Bacteriology **192**(3): 841-860.

Yuksel, S. and J. N. Hansen (2007). "Transfer of nisin gene cluster from *Lactococcus lactis* ATCC 11454 into the chromosome of *Bacillus subtilis* 168." Applied microbiology and biotechnology **74**(3): 640-649.

Zellmeier, S., W. Schumann and T. Wiegert (2006). "Involvement of Clp protease activity in modulating the *Bacillus subtilis* sigmaW stress response." Molecular Microbiology **61**(6): 1569-1582.

Zendo, T., M. Fukao, K. Ueda, T. Higuchi, J. Nakayama and K. Sonomoto (2003). "Identification of the lantibiotic nisin Q, a new natural nisin variant produced by *Lactococcus lactis* 61-14 isolated from a river in Japan." Bioscience, Biotechnology, and Biochemistry **67**(7): 1616-1619.

Zerikly, M. and G. L. Challis (2009). "Strategies for the discovery of new natural products by genome mining." ChemBiochem : a European journal of chemical biology **10**(4): 625-633.

Zhang, X., L. Shi, S. Shu, Y. Wang, K. Zhao, N. Xu, S. Liu and P. Roepstorff (2007). "An improved method of sample preparation on AnchorChip targets for MALDI-MS and MS/MS and its application in the liver proteome project." Proteomics **7**(14): 2340-2349.

Zhang, Z., Y. Wang and J. Ruan (1998). "Reclassification of *Thermomonospora* and *Microtetraspora*." International Journal of Systematic Bacteriology **48 Pt 2**: 411-422.

Zhao, C., J. Ju, S. D. Christenson, W. C. Smith, D. Song, X. Zhou, B. Shen and Z. Deng (2006). "Utilization of the methoxymalonyl-acyl carrier protein biosynthesis locus for cloning the oxazolomycin biosynthetic gene cluster from *Streptomyces albus* JA3453." Journal of Bacteriology **188**(11): 4142-4147.

Zhou, H. and W. A. van der Donk (2002). "Biomimetic stereoselective formation of methyllanthionine." Organic letters **4**(8): 1335-1338.

Zimmermann, N. and G. Jung (1997). "The three-dimensional solution structure of the lantibiotic murein-biosynthesis-inhibitor actagardine determined by NMR." European journal of biochemistry / FEBS **246**(3): 809-819.

Zimmermann, N., J. W. Metzger and G. Jung (1995). "The Tetracyclic Lantibiotic Actagardine - H-1-Nmr and C-13-Nmr Assignments and Revised Primary Structure." European journal of biochemistry / FEBS **228**(3): 786-797.