Cloning and analysis of the cypemycin biosynthetic gene cluster

Jan Claesen

September 2010

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 no quotation from the thesis, nor any information derived therefrom may be published without the author's prior, written consent.

Acknowledgements

I would like to start by saying that it has been a wonderful experience to work on this project in such a productive environment as the John Innes Centre. I could not have managed this on my own, so via this way, I would like to thank all the people that have turned the past four years into a really pleasant time.

First of all, I would like to thank my supervisor Prof. Mervyn Bibb for his support, optimism, guidance and for giving me the freedom to explore this project. For many useful discussions on biochemical subjects and lantibiotic genetics, I would like to thank my supervisory committee Dr. Stephen Bornemann and Dr. Arjan Narbad. A big thank you goes to Dr. Gerhard Saalbach and Dr. Mike Naldrett for the mass spectrometric analysis of many samples and the help in the interpretation of the resulting data. I am also very grateful to Prof. Wilfred van der Donk, Dr. Yuki Goto and Dr. Bo Li for their collaboration on the venezuelin project. For day-to-day advice and guidance, I would like to thank everyone in Molecular Microbiology. I am especially grateful to Nick Bird and Sean O'Rourke for making me feel at home in the lab from day one and teaching me the tricks of *Streptomyces* biology. I would like to thank the 'lantibiotics crew' Robert Bell, Lucy Foulston, Emma Sherwood and Diane Hatziioanou for many useful discussions. For helping me out in the lab, I would also like to thank Maureen Bibb, Govind Chandra, Juan-Pa Gomez-Escribano, Jennifer Parker, Richard Little, Steve Pullan, Nick Tucker, Andy Hesketh, Kay Fowler and Anyarat Thanapipatsiri. A special thanks goes to the staff at the Bateson and Chatt media kitchens and the stores facility. Their support has made it possible for me to focus on the science.

Finally, I would like to thank my family and friends for their support. Especially my parents and brother for believing in me and showing interest in my work. Last but not least, I would like to thank Marij for experiencing this project together with me, keeping me motivated on bad days and sharing the joys of successful experiments.

Abstract

Lantibiotics are post-translationally modified peptide antibiotics produced by Gram-positive bacteria that contain characteristic 'lanthionine' residues and sometimes other unusual modifications. Cypemycin was identified as an anti-leukemia compound produced by *Streptomyces sp.* OH-4156. It contains two unique modifications (*N*,*N*-dimethylalanine and *allo*-isoleucine) and shares two further modifications with the lantibiotic family, namely 2,3-didehydrobutyrine and a *S*-[(Z)-2-aminovinyl]-D-cysteine. Streptomycetes are filamentous soil-dwelling Gram-positive bacteria that have a complex developmental cycle and are well known as producers of many antibiotics, immunosuppressants and anti-tumour compounds.

The cypemycin biosynthetic gene cluster was identified, expressed in a heterologous host and a reduced gene set constructed that identified the genes sufficient for production. Mutational analysis of the individual genes within this set revealed that even the previously described modifications are carried out by unusual enzymes or via a modification pathway unrelated to lantibiotic biosynthesis. *In vitro* enzyme assays unambiguously confirmed the involvement of the two biosynthetic enzymes that are responsible for the modification of the N- and C-terminus of cypemycin.

Bioinformatic analysis revealed the widespread occurrence of cypemycin-like gene clusters within the bacterial kingdom and in the *Archaea*. Cypemycin is the founding member of an unusual class of post-translationally modified peptides, the linaridins. Genetic analysis of a linaridin cluster in *Streptomyces griseus* resulted in the identification of grisemycin.

Analysis of cryptic lantipeptide gene clusters (clusters for which the product has not yet been identified) in the genome of *Streptomyces venezuelae* identified a novel type of lanthionine-synthetase. Various genetic approaches were employed to identify the products from these clusters. Our collaborators reconstituted the activity of the unusual synthetase *in vitro*, revealing novel mechanistic and evolutionary insights.

General abbreviations

ABC	ATP-binding cassette
Act	actinorhodin
AIP	auto-inducing peptide
Apra	apramycin
ATP	adenosine triphosphate
AviCys	*S*-[(Z)-2-aminovinyl]-D-cysteine
BSA	bovine serum albumin
Carb	carbenicillin
Cl-Trp	5-chlorotryptophan
Cm	chloramphenicol
DAB	deoxyactagardine B
Dha	2,3-didehydroalanine
Dhb	(*Z*)-2,3-didehydrobutyrine
DMSO	dimethylsulfoxide
DNA gluc	Difco Nutrient agar with supplemented glucose
EDTA	ethylenediaminetetraacetic acid
EF-Tu	elongation factor Tu
FAB-MS	fast atom bombardment mass spectrometry
GlcNAc	*N*-acetylglucosamine
HFCD	homo-oligomeric flavin-containing cysteine-decarboxylases
Hyg	hygromycin B
IC50	half maximum inhibitory concentration
IPTG	isopropyl β-D-1-thiogalactopyranoside
Kan	kanamycin
Lab	(2*S*,4*S*,8*R*)-labionin
L-allo-Ile	L-*allo*-Isoleucine
Lan	lanthionine
MALDI-TOF MS	matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry
Mcc	microcin
Me_2-Ala	*N*,*N*-dimethylalanine
MeAviCys	*S*-[(Z)-2-aminovinyl]-3-methyl-D-cysteine
MIC	minimal inhibitory concentration
MRSA	methicillin resistant *Staphylococcus aureus*
Nal	nalidixic acid
NMR	nuclear magnetic resonance
NRP	non-ribosomal peptide
NRPS	non-ribosomal peptide synthase
ORF	open reading frame
PAPA	*p*-aminophenylalanine
PCR	polymerase chain reaction
PE	phosphatadylethanolamine
PFGE	pulsed-field gel electrophoresis
PG	peptidoglycan
PPC	4'-phosphopanthothenoylcysteine
QS	quorum sensing
Q-TOF	quadrupole time-of-flight

RBS	ribosome binding site
Red	undecylprodigiosin
SAM	*S*-adenosyl methionine
SDS	sodium dodecyl sulphate
SNA	soft nutrient agar
sp.	species
Spec	spectinomycin
SRP	signal recognition particle
Strep	streptomycin
TEMED	*N,N,N′,N′*-tetramethylethylenediamine
Tet	tetracycline
Thio	thiostrepton
Vio	viomycin
VRE	vancomycin resistant enterococci
WT	wild-type

Table of contents

Chapter I - Introduction 1

I.1. Post-translationally modified peptide natural products of bacterial origin 1
I.2. Bacteriocins that only require leader cleavage 4
I.3. Circular bacteriocins 5
I.4. Lantipeptides 7
I.4.1. Introduction 7
I.4.2. Lantibiotic gene clusters 8
I.4.2.a. Structural gene, *lanA* 9
I.4.2.b. *lanB* and *lanC* 10
I.4.2.c. *lanM* 11
I.4.2.d. *lanP* 11
I.4.2.e. *lanD* 12
I.4.2.f. Other modifications 13
I.4.2.g. Regulatory genes 13
I.4.2.h. Transport genes 15
I.4.2.i. Immunity genes 15
I.4.3. The influence of recent developments in actinomycete research on classification of lantipeptides 16
I.4.3.a. Class I - LanBC (microbisporicin and planomonosporin) 21
I.4.3.b. Class II - LanM (cinnamycin and the duramycins, actagardine and michiganin A) 23
I.4.3.c. Class III - Lantionine synthetase containing a Ser/Thr kinase domain 25
I.4.4. Mechanisms of action 28
I.4.5. Engineering lantipeptides 30
I.4.6. Genome mining to identify lantipeptide compounds 32
I.4.6.a. Biochemical approaches to genome mining 33
I.4.6.b. Molecular genetic approaches to genome mining 34
I.5. Microcins 35
I.5.1. Microcin B17 contains thiazole and oxazole heterocycles 35

I.5.2. Lassopeptides 36
I.5.3. Siderophore-peptides, a marriage between the ribosomal and non-ribosomal worlds 37
I.5.4. Microcin C7 38
I.6. Modified peptides produced by cyanobacteria 38
I.6.1. Cyanobactins 38
I.6.2. Microviridins 39
I.7. Other five-membered heterocycle-containing peptides 40
I.7.1. Linear toxins 40
I.7.2. Thiopeptides 41
I.8. Linaridins 42
I.9. Modified quorum sensing peptides 43
I.9.1. Cyclic (thio)lactones 43
I.9.2. *Bacillus* pheromones 44
I.9.3. Pep1357C 44
I.10. Aims of this project 45

Chapter II – Materials and methods 46

II.1. Bacterial plasmids and strains 46
II.2. Culture media and antibiotics 53
II.2.1. Antibiotics 53
II.2.2. Culture media 53
II.3. Growth conditions and genetic manipulations 55
II.3.1. Growth and storage of *E. coli* 55
II.3.2. Growth and storage of *Streptomyces* 55
II.3.3. Plasmid isolation from *E. coli* 56
II.3.4. Cosmid isolation from *E. coli* 56
II.3.5. Genomic DNA extractions from *Streptomyces* 57
II.3.6. Digestion of DNA with restriction enzymes 57
II.3.7. Agarose gel electrophoresis 58
II.3.8. Extraction of DNA fragments from agarose gels 58
II.3.9. Preparation and transformation of electro-competent *E. coli* 58

II.3.10. Preparation and transformation of chemically competent *E. coli* 59
II.3.11. Ligation of DNA 59
II.3.12. Conjugation of DNA into *Streptomyces* 59
II.3.13. Construction of a *Streptomyces sp*. OH-4156 cosmid library 60
II.3.14. Generation of protein fusion constructs 60
II.3.15. Cypemycin bio-assays 61
II.4. Polymerase chain reaction (PCR) methods and DNA sequencing 61
II.4.1. General PCR 61
II.4.2. *E. coli* and *Streptomyces* colony PCR 62
II.4.3. Purification of PCR products 62
II.4.4. DNA sequencing 62
II.5. DNA hybridisation methods 70
II.5.1. Non-radioactive Southern hybridisation 70
II.5.2. Radioactive Southern hybridisation 72
II.6. PCR targeting 72
II.6.1. PCR amplification of disruption cassette 72
II.6.2. PCR targeting of a cosmid 74
II.6.3. Transfer of mutant cosmids into *Streptomyces* 74
II.7. Protein methods 75
II.7.1. Purification of His-tagged proteins 75
II.7.2. Protein quantification 75
II.7.3. SDS-polyacrylamide gel electrophoresis (SDS-PAGE) 75
II.8. Mass spectrometry 76
II.8.1. MALDI-ToF and Q-ToF MS 76
II.8.2. Thiol alkylation 77

Chapter III – Identification of the cypemycin gene cluster 78

III.1. Introduction 78
III.2. Characterisation of cypemycin and its producer strain 78
III.2.1. Antibiotic sensitivities of *S. sp*. OH-4156 79
III.2.2. Cypemycin detection by bio-assay and MALDI-ToF MS 80

III.2.3. Structure verification by Q-ToF 83
III.3. Generation and screening of a genomic cosmid library 84
III.3.1. PCR based screening 84
III.3.2. Southern hybridisation-based screening 86
III.3.3. Ligation-mediated-PCR 88
III.4. Genome scanning of *S. sp.* OH-4156 89
III.5. Identification of the cosmids bearing the cypemycin biosynthetic cluster 93
III.5.1. Generation of specific *cypA* probe and library screen 93
III.5.2. Heterologous expression of the positive cosmids in *S. venezuelae* 94
III.6. The cypemycin biosynthetic cluster 97
III.7. Discussion 99
III.8. Bullet point summary 101

Chapter IV – Mutational analysis of cypemycin biosynthesis 102

IV.1. Introduction 102
IV.2. Generation of minimal gene set 102
IV.3. Mutational analysis of the cypemycin cluster 106
IV.4. Cypemycin is a member of a larger family of post-translationally modified peptides, the linaridins 113
IV.5. Generation of cypemycin variants 116
IV.6. Discussion 119
IV.7. Bullet point summary 123

Chapter V – Biochemical analysis of cypemycin biosynthesis 124

V.1. Introduction 124

V.2. Substrate peptides for the enzymatic reactions 125
V.3. Formation of the AviCys residue 128
V.4. Methylation of the N-terminal alanine residue 134
V.5. Formation of the Dhb residues 136
V.6. Discussion 137
V.7. Bullet point summary 139

Chapter VI – Characterisation of grisemycin, a linaridin produced by *S. griseus* 140

VI.1. Introduction 140
VI.2. Identification of grisemycin 141
VI.2.1. Bio-informatical analysis of the grisemycin gene cluster 141
VI.2.2. Grisemycin identification and structural verification 142
VI.3. Cloning and heterologous expression of the cluster and generation of a Δ*grmA* mutant 146
VI.3.1. Cloning of the gene cluster and heterologous expression in M1146 146
VI.3.2. Generation of a *grmA* deletion mutant 148
VI.4. Discussion 148
VI.5. Bullet point summary 151

Chapter VII – Cryptic lantipeptide gene clusters from streptomycetes 152

VII.1. Introduction 152
VII.2. Bioinformatic analysis of cryptic lantipeptide gene clusters in sequenced streptomycetes 154
VII.3. Cryptic gene clusters in *S. venezuelae* 162
VII.3.1. Deletion of the modification gene for each cluster 162
VII.3.2. Micro-array expression data for the different clusters 163

VII.3.3. Heterologous expression of the clusters 166
VII.3.4. Chemical induction of antibiotic production 167
VII.3.5. Constitutive expression of the modification enzymes in each cluster 170
VII.4. Discussion 174
VII.5. Bullet point summary 177

Chapter VIII – General discussion 178

References 182

Index to figures

Figure I.1. Schematic representation of post-translationally modified bacterial peptides 3

Figure I.2. Schematic representation of the different types of lanthionine synthetases 19

Figure I.3. Schematic representation of lantipeptides produced by actinomycetes 20

Figure III.1. MALDI-ToF spectrum of cypemycin 82

Figure III.2. Q-ToF spectrum of cypemycin 83

Figure III.3. Design of degenerate primers and probes for *cypA* 85

Figure III.4. Hybridisation of a nylon membrane containing *S. sp.* OH-4156 cosmid library clones with α-^{32}P-dCTP labelled Cyp probe 87

Figure III.5. The principle of Ligation-Mediated-PCR (LM-PCR) 88

Figure III.6. Sequence of the cypemycin prepropeptide deducted by genome scanning 89

Figure III.7. Comparison of *S. sp*. OH-4156 genome scanning data and the *S. venezuelae* genome 91

Figure III.8. Analysis of the *S. venezuelae* chromosomal region harbouring the chloramphenicol biosynthetic gene cluster 92

Figure III.9. Hybridisation of a nylon membrane spotted with 3072 *S. sp*. OH-4156 genomic cosmid library clones with a 480 bp ^{32}P labelled *cypA* probe 93

Figure III.10. Heterologous expression of the cypemycin biosynthetic gene cluster in *S. venezuelae* (MALDI-ToF) 95

Figure III.11. Heterologous expression of the cypemycin biosynthetic gene cluster in *S. venezuelae* (bio-assay) 96

Figure III.12. Schematic representation of the cypemycin biosynthetic gene cluster 98

Figure IV.1. Strategy for the construction of a minimal *cyp* gene set 104

Figure IV.2. *M. luteus* bio-assays for the minimal gene set construct and

the different scar mutants in the heterologous host M1146 105
Figure IV.3. MALDI-ToF spectra for scar mutants in M1146 107-108
Figure IV.4. MALDI-ToF spectra of the apramycin-marked mutants in *S. sp*. OH-4156 109-110
Figure IV.5. Structural verification of the truncated cypemycin derivatives 111
Figure IV.6. Analysis of the complementation of the Δ*cypL* mutant M1424 confirms production of both cypemycin and the non-decarboxylated form 112
Figure IV.7. The biosynthetic gene clusters of the linaridin family of peptides 114
Figure IV.8. Schematic representation of the three cypemycin variants 117
Figure IV.9. Mass spectrometric analysis of M1462 (*cypA* S-T) 118
Figure IV.10. CypH and/or CypL are involved in the dehydratation of the residues from Dhb and catalyse a similar dethiolation reaction that converts Cys to Dha 120

Figure V.1. SDS-PAGE analysis of the 6His-MBP-proCypA fusion protein purification 127
Figure V.2. Alignment of CypD with other HFCD-family decarboxylases 129
Figure V.3. Characterisation of the CypD cofactor 131
Figure V.4. Substrate preparation for the CypD decarboxylation assay 132
Figure V.5. CypD catalyses the *in vitro* decarboxylation of preproCypA 133
Figure V.6. CypM catalyses the *in vitro* dimethylation of the N-terminal Ala in non-methylated cypemycin 135

Figure VI.1. Comparison of the biosynthetic gene clusters for cypemycin and grisemycin production 141
Figure VI.2. MALDI-ToF MS spectrum of grisemycin 144
Figure VI.3. Q-ToF spectrum of grisemycin 145
Figure VI.4. MALDI-ToF MS spectra for M1457 and M1458 147
Figure VI.5. Alignment of the partial nucleotide sequences of *cypH* and *grmH* 150

Figure VII.1. Schematic representation of gene clusters predicted to contain functional biosynthetic enzymes 157-158
Figure VII.2. Multiple sequence alignment for related LanA peptides 159
Figure VII.3. Gene expression profiles for selected secondary metabolite gene clusters from *S. venezuelae* 165
Figure VII.4. Chemical induction of the heterologously expressed *S. venezuelae* lantipeptide clusters 169
Figure VII.5. Act overproduction phenotype of *S. coelicolor* M1456 on SFM 172
Figure VII.6. Identification of a possible product from the cryptic SMD01252 gene cluster 173

Index to tables

Table II.1.	Plasmids and cosmids used and constructed in this study	46-49
Table II.2.	Strains used or constructed in this study	50-52
Table II.3.	Concentration of antibiotics used in this study	53
Table II.4.	Additional fermentation media for actinomycetes	54
Table II.5.	Primers used in this study	64-69
Table III.1.	Minimal inhibitory concentration (MIC) values for *S. sp.* OH-4156 when grown on solid medium	80
Table IV.1.	Putative structural peptides	115
Table VII.1.	Lantibiotic synthetases of the LanB, LanC, LanM and LanL types identified in streptomycete genomes	155

Chapter I – Introduction

I.1. Post-translationally modified peptide natural products of bacterial origin

Peptide natural products with various chemical structures and biological functions are produced by bacteria of all major phylogenetic lineages. These compounds can either be genetically encoded and ribosomally translated or they can be assembled on large multimodular enzymes called non-ribosomal peptide synthetases (NRPSs). The latter synthesise peptides via a mechanism similar to ribosomal translation, by linking amino acid monomers via a peptide bond to a growing chain. However, the sequence of the final product is not determined by messenger RNA (mRNA), but by the substrate specificity of the enzyme's successive modules, allowing the incorporation of non-proteinogenic amino acids. Peptide natural products of NRPS origin are not a topic of this introductory chapter, but have been described in several recent review articles (Fischbach and Walsh 2006; Koglin and Walsh 2009).

Although the ribosomal translation machinery can only use the 20 genetically encoded amino acids, Nature has invented several mechanisms to expand on this repertoire via the introduction of post-translational modifications. This introductory chapter presents an overview of the myriad of different post-translationally modified peptides that are produced by both Gram-positive and Gram-negative bacteria. These peptides are typically produced as short preproproteins consisting of N-terminal leader sequences and C-terminal proproteins that are post-translationally modified. Proteolytic cleavage of the leader peptide is generally required for liberation of the functional mature peptide product. Several 'exceptions' to this general 'rule' have been described, including C-terminal extensions or even multiple peptide products that are processed from a single structural peptide. The genetic determinants involved in the biosynthesis of peptide natural products are usually found clustered in the bacterial genome or on a plasmid. Most of the compounds under discussion were originally characterised because of their antibacterial activity, but structurally related compounds with other biological functions will also be mentioned. Introduction of post-translational modifications are speculated to contribute to peptide structural rigidity, function and stability. Since there is no

unambiguous definition of a 'peptide', a size range up to 100 amino acids will be employed as an arbitrary cut-off value for inclusion of a compound in this chapter.

Post-translationally modified peptides of bacterial origin are often intuitively associated with bacteriocins such as lantibiotics and microcins. The term 'bacteriocin' was used for the first time by Jacob *et al.* (1953) to refer to peptide compounds of the colicin type. Tagg *et al.* (1976) later expanded upon this to include peptides produced by Gram-positive bacteria and defined a bacteriocin as a 'proteinaceous compound that kills closely related bacteria'. A bacteriocin classification was proposed by Klaenhammer (1993) and as more compounds were characterised structurally and genetically, this classification was expanded and became the subject of lively debate (Jack *et al.* 1995; van Belkum and Stiles 2000; Diep and Nes 2002; Cotter *et al.* 2005; Franz *et al.* 2007; Zouhir *et al.* 2010). Unfortunately, there are almost as many different classification schemes as there are reviews on the subject. This overview is not intended as an attempt at classification and includes, but is not restricted to, modified bacteriocins. Instead it focuses on the wonderful world of post-translationally modified peptide structures (Figure I.1.), their biological function and underlying biosynthesis.

Non-lantibiotic bacteriocins are not subject to extensive post-translational modification (Cotter *et al.* 2005). Since most of these compounds are produced as a precursor peptide that at least requires post-translational cleavage of the leader sequence, they will be discussed briefly in Chapters I.2. and I.3. Several leaderless bacteriocins have been reported, including enterocin I (Floriano *et al.* 1998), aureocin A53 (Netz *et al.* 2002), mutacin BHT-B (Hyink *et al.* 2005) and lacticin Q (Fujita *et al.* 2007). These peptides contain a formylated N-terminal Met that is the result of the bacterial translation mechanism rather than a post-translational modification and thus they will not be discussed in this literature review. Chapter I.4. focuses in some detail on lantipeptides since they are central to this PhD thesis (Chapter VII). Other important families that will be discussed include microcins (Chapter I.5.), peptides produced by cyanobacteria (Chapter I.6.) and other five-membered heterocycle-containing peptides (Chapter I.7.). Cypemycin, the main subject of this thesis (Chapter III – V), was originally classified as a lantibiotic, but was identified in this work as the first representative of the linaridin family (Chapter I.8.). Finally, several types of modified peptides involved in quorum sensing will be discussed (Chapter I.9.).

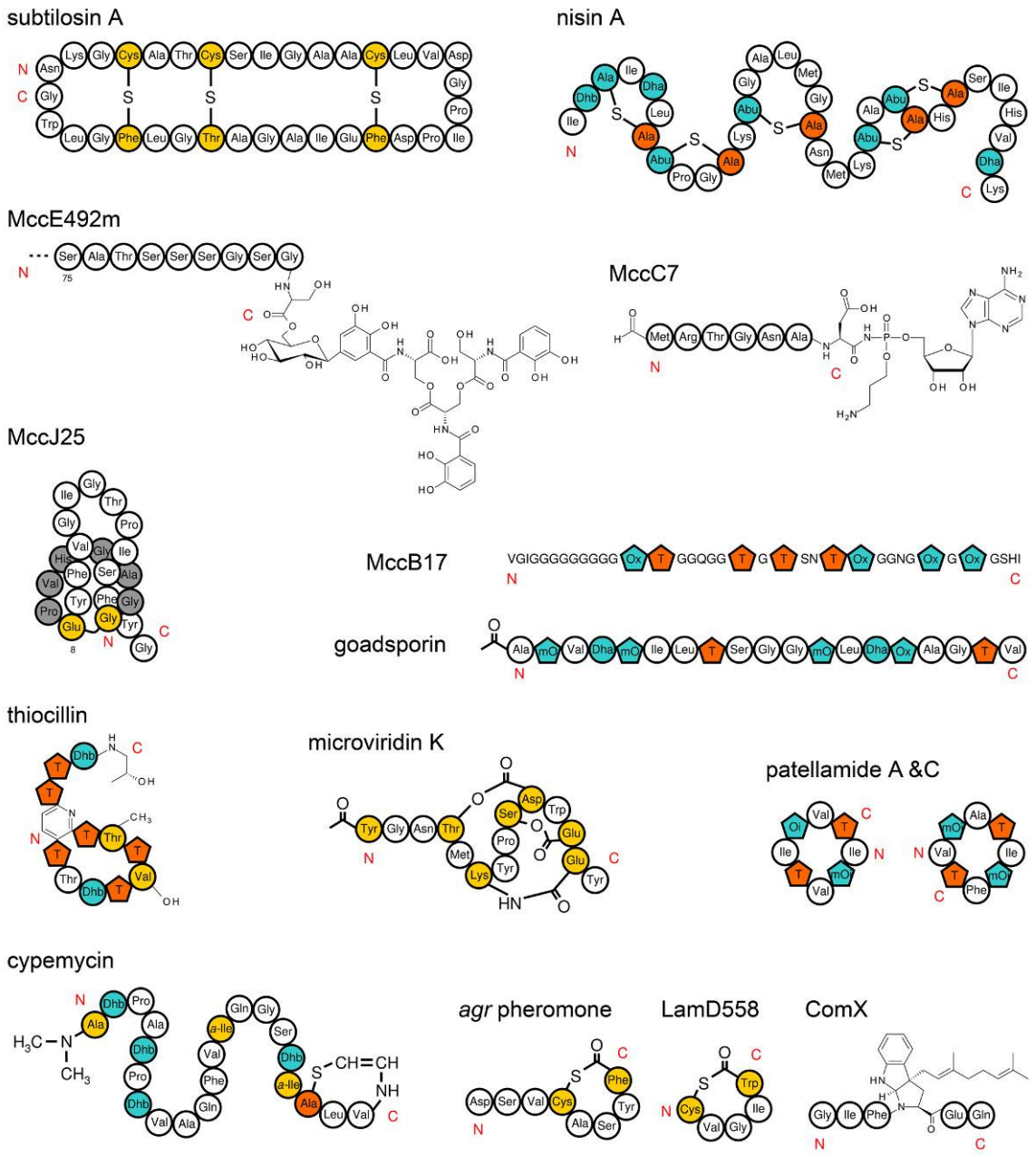


Figure I.1. Schematic representation of post-translationally modified bacterial peptides. For information on the specific modifications, producing organisms and references, refer to the text. The N-terminal 74 amino acids of MccE492m are not shown. Residues resulting from modification of Ser and Thr are depicted in blue, and modified Cys are in orange. Unusual post-translationally modified amino acids are depicted in yellow. The N-terminal amide ring in MccJ25 is shown in grey. Five-membered heterocycles are represented by pentagons. The N- and C-terminal ends of the peptides are indicated by N and C, respectively. Abu: aminobutyric acid; Ala-S-Ala: lanthionine; Abu-S-Ala: methyllanthionine; Dha: didehydroalanine; Dhb: didehydrobutyrine; Ox: oxazole; Oi: oxazoline; mO: methyloxazole; mOi: methyloxazoline; T: thiazole.

I.2. Bacteriocins that only require leader cleavage

Among bacteriocins that only require proteolytic cleavage of the leader peptide, both single peptides and two-peptide systems have been described. Single-peptide bacteriocins that resemble pediocin PA-1 are among the best characterised examples because of their anti-listerial activity (Eijsink *et al.* 1998). These peptides are cationic and partly amphiphilic or hydrophobic, containing a conserved 'pediocin box' sequence motif and two Cys residues that are linked by a disulfide bridge. The latter is predicted not to be required for activity since reduction and subsequent derivatisation with iodoacetamide did not result in loss of activity of sakacins A and P (Holck *et al.* 1992). Instead, the disulphide bridge is proposed to contribute to the rigidity and heat stability of pediocin-like bacteriocins.

Pediocin-like bacteriocins exert their mechanism of action by permeabilisation of the cytoplasmic membrane in sensitive bacteria, resulting in a disruption of the proton motive force and a depletion of the ATP pool (Chikindas *et al.* 1993; Montville and Chen 1998). The membrane-associated mannose phosphotransferase system permease has been implicated as a determinant for the specificity of these bacteriocins for their target strains (Ramnath *et al.* 2000). Single-peptide bacteriocins are produced as precursor peptides that are secreted and processed either by specialised ABC transporters or via the *sec*-dependent pathway. The former contain an N-terminal leader peptide with the characteristic double-Gly motif that is recognised by the N-terminal proteolytic domain of the ABC transporter (Havarstein *et al.* 1995). Production is usually controlled by a so-called 'three-component system', involving a bacteriocin-like pheromone, a histidine kinase that can sense the secreted pheromone and a response regulator that can activate transcription of target promoters upon phosphorylation by the histidine kinase (Brurberg *et al.* 1997). The bacteriocin-like pheromone also has a double-Gly leader peptide and is processed and secreted by the same transporter as the bacteriocin, but their antibacterial activity is negligible compared to their associated bacteriocins (Eijsink *et al.* 1996). A mechanism for self-immunity is provided by a small protein of 88 to 115 amino acids encoded by a gene in the bacteriocin cluster, often in the same operon as the structural gene to ensure co-regulation (Eijsink *et al.* 1998). In the case of the pediocin-like bacteriocins, the immunity proteins are located inside the cell, but could be loosely associated with the inside of the cell membrane (Quadri *et al.* 1995). A recent

study by Diep *et al.* (2007) revealed that immunity proteins bind to the mannose phosphotransferase permease, but only if the bacteriocin is already bound to this target. Thus, the immunity proteins are proposed to protect permease blockage and membrane leakage by recognition and binding of the permease-bacteriocin complex.

Two-peptide bacteriocins consist of two very different peptides that are required in roughly equal amounts to function as one antimicrobial entity. Lactococcin G was the first two-peptide bacteriocin to be described (Nissen-Meyer *et al.* 1992). Two-peptide bacteriocins also kill susceptible cells by permeabilisation of the cytoplasmic membrane. With a few exceptions, the individual peptides do not display antimicrobial activity (Anderssen *et al.* 1998). The two-peptide bacteriocin biosynthetic gene clusters typically only contain one immunity protein, but the mechanism by which this works remains to be determined.

I.3. Circular bacteriocins

Circular bacteriocins are a family of peptides that have a peptide linkage between their N- and C-terminal ends, resulting in a circular backbone. Compounds belonging to this family have been described in the literature both as ‘cyclic’ and as ‘circular’, but Craik *et al.* (2003) proposed to use the latter name to distinguish these head-to-tail cyclised peptides from others containing different cross-links that make their structure only partially cyclic. About nine different circular bacteriocins have been described to date with a size range between 58 and 78 amino acid residues (reviewed in Maqueda *et al.* 2008). Peptidase homologues have been proposed as candidates for the leader cleavage and cyclisation reactions (Kawulka *et al.* 2003), but the underlying mechanisms and order of these events remain speculative. A remarkable feature of several circular bacteriocins is that their propeptide sequence is preceded by a very short leader sequence. Circularin A (Kemperman *et al.* 2003), carnocyclin A (Martin-Visscher *et al.* 2008), uberolysin (Wirawan *et al.* 2007) and subtilosin A (Zheng *et al.* 1999) have leader sequences of three, four, six and eight amino acids, respectively.

Enterocin AS-48, the circular bacteriocin produced by *Enterococcus faecalis*, was the first compound to be identified from this family (Gálvez *et al.* 1986). The circular nature of enterocin AS-48 was determined by chemical degradation (Samyn *et al.* 1994) and

confirmed by the identification of its structural gene (Martinez-Bueno *et al.* 1994). The latter is a particularly challenging task in peptides from this family since it is impossible to tell from the mature structure where the original N- and C-terminal ends were linked together. The peptide exerts its antimicrobial activity by voltage-independent interaction with the cytoplasmic membrane of susceptible bacteria, followed by membrane insertion that probably leads to pore formation resulting in ion leakage and a collapse of membrane potential (Gonzalez *et al.* 2000). Structural analysis of carnocyclin A by nuclear magnetic resonance (NMR) revealed a fold that is surprisingly similar to the structure of enterocin AS-48, despite a low degree of sequence identity (Martin-Visscher *et al.* 2009). This study suggests that many circular bacteriocins might have a common structural motif, regardless of their amino acid sequence.

The circular bacteriocin subtilosin A (Figure I.1.), which is produced by *Bacillus subtilis*, is thus far the only circular bacteriocin described that contains an additional type of post-translational modification. The compound contains three linkages that involve the thiol groups of its three Cys residues and the α-carbon of two Phe and a Thr, resulting in the loss of a hydrogen from each of these six residues (Kawulka *et al.* 2003). Similar post-translational modifications have been observed in thuricin CD, a two-peptide bacteriocin produced by *Bacillus thuringiensis* with potent activity against *Clostridium difficile* (Rea *et al.* 2010). The mature peptides Trnα and Trnβ are not circular, but each peptide contains three linkages between the thiol of Cys5, Cys9 and Cys13 and the α-carbon of the amino acid at position 28, 25 and 21, respectively. Sequencing and bioinformatic analysis of the thuricin gene cluster identified two radical S-adenosylmethionine (SAM) superfamily proteins that share only low sequence identity (19 % and 17 %) with the radical SAM enzyme AlbA encoded by the subtilosin A gene cluster (Zheng *et al.* 2000; Rea *et al.* 2010). These enzymes are proposed to be involved in the introduction of the Cys to α-carbon linkages.

I.4. Lantipeptides

I.4.1. Introduction

Lantibiotics are ribosomally synthesized peptide antibiotics that are extensively post-translationally modified. They are produced by a variety of Gram-positive bacteria and their mechanisms of action include disruption of cell wall biosynthesis, inhibition of spore outgrowth, membrane permeabilisation and enzyme inhibition. Schnell *et al.* (1988) introduced the name lantibiotics as an abbreviation for 'lanthionine-containing antibiotic peptides', because of the occurrence of the unusual amino acid lanthionine (Lan) in their structures. The Lan residue consists of a thioether crosslink between the *β*-carbons of two alanine moieties and results from the intramolecular nucleophilic addition of a Cys residue onto 2,3-didehydroalanine (Dha). This unusual amino acid is introduced into the lantibiotic peptide by dehydration of Ser. Thr residues are also prone to dehydration, yielding (*Z*)-2,3-didehydrobutyrine (Dhb), which can form a methyl-lanthionine (MeLan) residue upon cyclisation with Cys. Most lantibiotics have been identified because of their antibacterial activity. However, several compounds are known that contain (Me)Lan residues, but do not display (significant) antibiotic activity. These are called lanthionine-containing peptides or 'lantipeptides' (Goto *et al.* 2010) and they will be discussed in more detail in Chapter I.4.3.

A wide range of additional amino acid modifications has been reported for lantipeptides, including *S*-[(Z)-2-aminovinyl]-D-cysteine (AviCys), (2*S*,8*S*)-lysinoalanine, β-hydroxy-aspartate, D-alanine, 2-oxopropionyl, 5-chlorotryptophan and dihydroxylated proline. Consequently, a significant amount of research is directed at unveiling the novel enzymatic mechanisms responsible for these unusual amino acids. Better understanding of the biosynthetic machinery and of substrate specificity will provide powerful tools for both *in vivo* and *in vitro* protein engineering, eventually potentially leading to intelligent drug design (Pag and Sahl 2002).

Nisin (Figure I.1.) has been studied intensively because of its widespread use in the food industry (Delves-Broughton *et al.* 1996), and is often considered as the "model

lantibiotic". However, approximately 55 different lantibiotics have been described in the literature and the biosynthesis of many has been thoroughly characterised.

Structural characterisation of lantibiotic peptides has proven problematic - dehydrated residues at the N-terminus of the peptide or exposed during Edman degradation spontaneously deaminate resulting in a sequence block (Jung 1991). A milestone in the structural determination of lantibiotics was a method developed by Meyer *et al.* (1994) that allowed direct sequencing of dehydroamino acids and thioethers after treatment with thiol compounds under alkaline conditions. More recently, a novel derivatisation technique was developed by Martin *et al.* (2004) to characterise the two-peptide lantibiotic lacticin 3147. This technique involves simultaneous hydrogenation of the dehydroamino acids and the desulphurisation of the thioethers by using nickel boride. If sodium borodeuteride is used, two deuterium ions will be incorporated into derivatives of Dha and Dhb, whereas only one deuterium will be present in residues resulting from Lan or MeLan residues. The derivate is amenable to Edman degradation and the location of the deuterated residues can be determined by mass spectrometry. A significant advantage of this methodology is that only small amounts of lantibiotic (micrograms) are required for analysis. The technique allows rapid differentiation between Dha/Dhb and (Me)Lan residues, but does not reveal the exact bridging pattern of the lantibiotic. Hence, a combination of this method and NMR spectroscopy is required to determine the precise structure.

I.4.2. Lantibiotic gene clusters

Thus far, genes encoding the biosynthetic machinery responsible for lantibiotic production are clustered together in the genome of the producing organism. In most cases, these gene clusters are located on the chromosome, but they can also reside on large plasmids, e.g. epidermin (Schnell *et al.* 1988), or on transposable elements, e.g. nisin (Horn *et al.* 1991). Each type of lantibiotic biosynthetic gene has been given a generic '*lan*' designation. Prepropeptides are encoded by structural genes termed '*lanA*' that are found in the clusters with genes required for Dha, Dhb and (Me)Lan formation (*lanB*, *lanC*, *lanM*, *ramC*, *lanL*, *labKC*), proteolytic processing (*lanP*), transport (*lanT*, *lanH*), immunity (*lanI*, *lanFEG*), regulation (*lanR*, *lanK*, *lanQ*) and additional modification (*lanD*, *lanJ*). In the following sections, the products of each of these gene types will be

described in more detail and *in vitro* reconstitution of specific enzyme activities will be discussed.

I.4.2.a. Structural gene, *lanA*

Epidermin was the first lantibiotic for which the structural gene was identified, thereby confirming its ribosomal origin and by extension that of other lantibiotics (Schnell *et al.* 1988). Epidermin was also the first lantibiotic for which the biosynthetic gene cluster was cloned and heterologously expressed in a closely related strain (Schnell *et al.* 1992).

The lantibiotic structural genes encode a prepropeptide (LanA), consisting of an N-terminal leader sequence of 23 to 59 amino acids, and the structural region or propeptide, which is modified to yield the mature lantibiotic. During the last step of biosynthesis - secretion and subsequent cleavage of the leader sequence from the mature peptide - the lantibiotic is activated. The leader peptide generally has no similarity to Sec or Tat signal sequences but may play a role in immunity (by keeping the antibiotic inactive in the cell), in transport, and/or as a recognition sequence for the biosynthetic enzymes (Chatterjee *et al.* 2005; Oman and van der Donk 2010). It has been demonstrated for Pep5 that although the leader peptide contains Ser and Thr residues, only residues in the propeptide undergo modification (Weil *et al.* 1990). The leader peptides of all characterised lantibiotics are also devoid of Cys, in contrast to their Cys-rich propeptides.

Analogous to ‘unmodified’ two-peptide bacteriocins, two-peptide lantibiotics have also been described where a combination of both peptides is required for significant activity. Each prepropeptide of a two-peptide lantibiotic is genetically encoded and requires its own dedicated LanM enzyme for modification. The only exception to date is cytolysin, where only one LanM modifies both structural peptides (Gilmore *et al.* 1994). The gene clusters of some lantibiotics (for example ruminococcin, macedocin and Bsa) contain two or more homologous structural genes with a high degree of amino acid identity (Marcille *et al.* 2002; Papadelli *et al.* 2007; Daly *et al.* 2010). The reason for this is not clear, but in the case of mutacin I and III, inactivation of one *lanA* completely abolished antimicrobial activity whereas inactivation of the homologous *lanA* did not (Qi *et al.* 1999; Qi *et al.* 2000).

I.4.2.b. *lanB* and *lanC*

The products of the *lanB* and *lanC* genes are responsible for the introduction of the Dha and Dhb residues into lantibiotic propeptides and the subsequent intramolecular nucleophilic attack of a Cys residue, resulting in the formation of a Lan or MeLan bridge.

lanB genes encode large proteins (of approximately 1000 amino acids in size) which are overall of a hydrophilic nature. However, the LanB proteins also contain hydrophobic domains, suggesting an interaction with the cell membrane, as has been shown for NisB and SpaB (Engelke *et al.* 1992). Koponen *et al.* (2002) demonstrated the role of NisB and NisC in nisin biosynthesis. An unmodified His-tagged nisin precursor was isolated from a Δ*nisB* mutant, indicating that NisB was involved in dehydration of the Ser and Thr residues. The His-tagged prepropeptide isolated from the Δ*nisC* mutant was dehydrated, but no cyclisation reactions had occurred, making the NisC protein a good candidate for the cyclisation reaction. In another study, the activity of NisB was demonstrated by expressing a non-lantibiotic peptide as a fusion construct with the C-terminal part of the NisA leader peptide (Kuipers *et al.* 2004). The peptide was expressed in a non-nisin producing *Lactococcus lactis* strain harbouring the *nisABT* genes on a plasmid. The peptide was dehydrated by NisB and secreted by the NisT transporter, demonstrating the rather broad substrate specificity of the NisB dehydratase. No *in vitro* reconstitution of an active LanB-family enzyme has been reported to date.

The *lanC* genes encode for cyclases ranging from 398 to 455 amino acids in size. In *lanC* deletion strains, only dehydrated peptides are formed (Meyer *et al.* 1995; Koponen *et al.* 2002). This illustrates the essential role of LanCs in the cyclisation reaction. NisC and SpaC are zinc binding proteins containing His and Cys residues, which are conserved in both LanC and LanM proteins (Okeley *et al.* 2003). Structural analysis of NisC showed that these residues are involved in binding the zinc ligand (Li *et al.* 2006). This study also reports the *in vitro* reconstitution of NisC activity, providing the first biochemical characterisation of a LanC cyclase. A site-directed mutagenesis study of the enzyme revealed essential catalytic residues and confirmed the involvement of the metal-binding residues (Li and van der Donk 2007).

I.4.2.c. *lanM*

Lantipeptide gene clusters that do not contain *lanB* or *lanC* genes have a single gene instead that encodes an enzyme that catalyses both the dehydration and cyclisation reactions during maturation of the prepropeptide. Two major types of these multifunctional enzymes have been reported, namely LanMs and proteins belonging the RamC-LanL-LabKC family. Members of the latter family contain an internal domain homologous to Ser/Thr kinases and will be discussed in more detail in Chapter I.4.3. *lanM* genes encode for a protein of approximately 900-1000 amino acids in size, containing a C-terminal part with 20-27% sequence identity to LanC proteins (Okeley *et al.* 2003). This domain contains the conserved residues involved in catalysis and the metal-binding site mentioned above. No homology can be detected between the N-terminal part of LanM and LanB proteins, indicating the *lanM* genes did not simply originate from a gene fusion between a *lanB* and a *lanC* gene (Siezen *et al.* 1996). Experiments involving the *in vitro* biosynthesis of lacticin 481, catalysed by LctM, provided proof that LanM proteins are involved in both dehydration and cyclisation (Xie *et al.* 2004). This study also encompasses the first *in vitro* production of a lantibiotic. Purified His-tagged LctM was incubated with a His-tagged LctA prepropeptide and after protease treatment to remove the leader sequence, the bio-active lantibiotic was formed. Both a Mg^{2+} ion and ATP are required for LctM activity.

I.4.2.d. *lanP*

After or upon secretion of the lantibiotic, a protease is required for cleavage of the leader peptide resulting in activation of the mature peptide. In some gene clusters, a *lanP* gene, encoding a dedicated subtilisin-type Ser protease is present. LanP proteins vary in size from 266 to 682 amino acids and several possess an N-terminal signal sequence which implies they can be directed out of the cell and act extracellular. NisP contains a C-terminal extension that could act as a membrane anchor, suggesting that the protease may become attached to the outside of the cytoplasmic membrane (van der Meer *et al.* 1993).

When the lantibiotic gene cluster does not harbour a *lanP* gene, one or several extracellular proteases are capable of activating the lantibiotic by removing the leader

sequence, as has been shown for subtilin (Corvey *et al.* 2003). Alternatively, in LanM-type lantibiotic gene clusters, the cleavage is performed by a chimeric ABC-transporter LanT(P) that contains an N-terminal protease domain. This implies that LanT(P) transporters are capable of removing the leader peptide while translocating the lantibiotic, as has been shown for MrsT (mersacidin), LctT (lacticin 481), CylT (cytolysin) and ScnT (SA-FF22). Lactocin S is an exception as it is processed and translocated by seprate LasP and LasT proteins (Skaugen *et al.* 1997). The *in vitro* leader peptidase activity of the N-terminal domain of the lacticin 481 transporter LctT has been reported (Furgerson Ihnken *et al.* 2008). Lantibiotic clusters from actinomycete origin (cinnamycin, actagardine A, deoxyactagardine B, michiganin A and microbisporicin) contain no designated protease gene, suggesting that the leader sequence is cleaved by an enzyme encoded elsewhere in the genome (Widdick *et al.* 2003; Holtsmark *et al.* 2006; Boakes *et al.* 2009; Boakes *et al.* 2010; Foulston and Bibb 2010).

I.4.2.e. *lanD*

The additional enzyme LanD is involved in further modification of epidermin and mersacidin. The 181 amino acid, FMN-containing EpiD functions as a peptidyl-Cys decarboxylase, catalysing the oxidative decarboxylation of the C-terminal Cys residue to AviCys (Kupke *et al.* 1994). The reaction mechanism involves the oxidation of the Cys22 thiol to a thioaldehyde, which is hypothesised to tautomerise to a more stable ene-thiolate via decarboxylation of the C-terminus. This is followed by thioether ring formation by attacking an internal Dha residue at position 19 (Blaesse *et al.* 2000). MrsD is a 194 amino acid FAD-containing enzyme catalysing the analogous reaction to form *S*-aminovinyl-3-methyl-D-cysteine (AviMeCys) in the lantibiotic mersacidin (Majer *et al.* 2002; Blaesse *et al.* 2003). The structures of epidermin and mersacidin reveal that requirements for the position of the dehydrated residue, involved in formation of the Avi(Me)Cys, are quite relaxed. In epidermin, the AviCys bridge spans two other amino acids, whereas in mersacidin four amino acids are spanned by AviMeCys.

I.4.2.f. Other modifications

Both lactocin S and lacticin 3147 contain D-Ala residues that result from the modification of an L-Ser in the propeptide (Skaugen *et al.* 1994; Ryan *et al.* 1999). These modified residues are introduced in a two-step process initiated by dehydration of L-Ser to nonchiral Dha by a lanthionine synthetase, followed by a stereospecific hydrogenation to form D-Ala, a reaction catalysed by LanJ enzymes (Cotter *et al.* 2005). Site-directed mutagenesis of the individual three Ser codons to Ala resulted in mature lacticin 3147 peptides that contain the L- rather than the D-isoform. These changes were found to have a negative effect on production and activity of the lantibiotic (Cotter *et al.* 2005).

Despite the fact that almost all Cys residues are involved in (Me)Lan formation, several compounds with a disulfide bridge or 'cystine' have been described, namely sublancin 168, plwα, bovicin HJ50, thermophilin 1277 and the labyrinthopeptins (Paik *et al.* 1998; Holo *et al.* 2001; Xiao *et al.* 2004; Kabuki *et al.* 2009; Meindl *et al.* 2010). A thiol-disulfide oxidoreductase, BdbD was implied in the introduction of the cystine residues in sublancin 168 (Dorenbos *et al.* 2002). Deletion of a thiol-disulfide oxidoreductase downstream of the bovicin HJ50 gene cluster did not alter the mass of the compound, indicating the disulfide bridge remained intact (Liu *et al.* 2009). Heterologous expression in a *L. lactis* strain lacking thiol-disulfide reductases, resulted in production of bovicin HJ50 without the disulfide bridge, but the cystine appeared not to be required for antimicrobial activity.

I.4.2.g. Regulatory genes

The biosynthesis of several lantibiotics appears to be regulated at a transcriptional level in a cell density dependent manner. This phenomenon, named quorum sensing (QS), is under control of a two-component regulatory system, encoded by the genes *lanR* and *lanK*. The lantibiotics themselves act as peptide pheromones capable of triggering a signal transduction mechanism, resulting in expression of biosynthetic genes which contain specific lantibiotic-responsive elements (*lan*-boxes) in their promoters (Kleerebezem *et al.* 2004).

Lantibiotic production is generally initiated during mid- to late-logarithmic growth phase, peaking at the onset of stationary phase. The QS regulated production allows environmental concentrations of the lantibiotic to rapidly reach the level required for killing competitors of the producer strain. The lantibiotic's resulting high extracellular concentration ensures optimal effectiveness since the target strain is unlikely to be able to initiate or develop a fast defence reaction (Kleerebezem 2004).

The two-component system comprises two different proteins, a sensor kinase (LanK) which is embedded in the cytoplasmatic membrane, and an intracellular response regulator (LanR). For nisin (Ra *et al.* 1996), subtilin (Stein *et al.* 2002; Kleerebezem *et al.* 2004), salivaricin A (Upton *et al.* 2001) and mersacidin (Schmitz *et al.* 2006), it was shown that the lantibiotic itself functions as a signalling molecule, which is predicted to be recognised by the extracellular domain of LanK. Upon recognition, the sensor kinase autophosphorylates a conserved histidine residue in its intracellular C-terminal domain. This phosphate moiety, in turn, is transferred to a conserved asparagine residue in the N-terminal part of the response regulator. The C-terminal domain of the LanR protein undergoes a conformational change and gains the ability to act as a transcriptional regulator by binding to the operator sequence(s) of the target gene(s).

While there is a wide diversity in the QS regulation of lantibiotic gene clusters, all studied examples are variants of the general mechanism described above. Biosynthesis of cytolysin L_L/L_S is regulated by two different regulator proteins, CylR1 and CylR2 (Haas *et al.* 2002). Lacticin 3147 and mutacin II do not have a dedicated sensor kinase in their gene clusters, however, they do have an orphan LanR response regulator (Pag and Sahl 2002). In the case of lacticin 3147, LtnR was implicated in regulation of the immunity determinants in the biosynthetic gene cluster.

A putative stem-loop structure has been identified downstream of the structural genes for mersacidin, cinnamycin, actagardine A and deoxyactagardine B (Altena *et al.* 2000; Widdick *et al.* 2003; Boakes *et al.* 2009; Boakes *et al.* 2010). This structural DNA element could function as a transcriptional attenuator that allows limited read-through to the downstream biosynthetic genes, probably serving to maintain an appropriate stoichiometry between the precursor peptide and the modification machinery.

I.4.2.h. Transport genes

In LanBC gene clusters, a *lanT* gene encodes a transport protein of 500 to 600 amino acids in size. The LanT transporter belongs to the group A of ATP-binding-cassette (ABC)-transport family and is responsible for the translocation of the modified lantibiotic precursor across the cytoplasmic membrane. The protein consists of two functional domains, an N-terminal hydrophobic domain containing six transmembrane helices and a C-terminal domain containing two conserved ATP-binding motifs. Translocation is driven by ATP hydrolysis as a source of energy and for most LanT proteins, the active transporter is formed by dimerisation at the cytoplasmic membrane. Inactivation of *nisT* has shown that the NisT transporter is essential for export of nisin (Skaugen *et al.* 1997). In contrast, disruption of *pepT* resulted in reduced production of Pep5 (Bierbaum *et al.* 1994) and in the epicidin 280 gene cluster, no *lanT* homologue was found (Heidrich *et al.* 1998). These findings indicate that other cellular transporters might be involved in the secretion of some lantibiotics. Most lantibiotics that are activated by cleavage after a 'double glycine' motif harbour a dual function LanT transporter in their gene cluster, as discussed in section I.4.2.d. Epicidin 280 is the only characterised lantibiotic gene cluster to date that does not contain a dedicated ABC transporter (Heidrich *et al.* 1998; Chatterjee *et al.* 2005).

I.4.2.i. Immunity genes

An efficient auto-immune system is of vital importance for lantibiotic-producing strains. Therefore, most lantibiotic gene clusters contain specific genes that are involved in self-protection. Two mechanisms for auto-immunity have been described in lantibiotic producing strains, namely the production of a specific immunity peptide LanI and the expression of a specialised transporter LanFEG. Whereas some lantibiotic gene clusters contain only one of these mechanisms, others contain both systems.

LanI immunity peptides are small proteins ranging from 57 to 245 amino acids in size. NisI and SpaI are both lipoproteins containing an N-terminal Sec signal sequence and a membrane anchor immediately following the cleavage site (Qiao *et al.* 1995). The maturation of SpaI through the lipoprotein biosynthesis pathway was investigated and site-directed mutagenesis of the candidate lipobox identified variants that blocked the

protein's acylation (Halami *et al.* 2010). PepI also seems to be associated with the extracellular side of the cytoplasmic membrane and its immunity phenotype appears to be coupled to Pep5 production at a transcriptional level (Reis *et al.* 1994). The molecular immunity mechanism of the LanI peptides remains to be determined, but they have been proposed to bind the associated lantibiotic, thereby preventing its penetration and pore-formation in the host cell membrane (Stein *et al.* 2003; Stein *et al.* 2005).

The second mechanism for auto-immunity involves a transporter from the ABC family, which is encoded by the *lanFEG* genes. This transporter consists of membrane-spanning subunits LanE and LanG, and the LanF protein which contains the ATP-binding site. The LanFEG transporters are proposed to provide auto-immunity by extruding the lantibiotic into the surrounding medium and thus keeping its concentration in the cell at a low level. This way, the cell membrane of the producer is protected from the activity of its own lantibiotic, as shown for epidermin (Otto *et al.* 1998). A different kind of transporter, encoded by *lanH*, is involved in both secretion and immunity in gallidermin biosynthesis (Hille *et al.* 2001). A recent study on the immunity mechanism of lacticin 3147 showed that producers of the closely related lantibiotic staphylococcin C55 displayed cross-immunity to lacticin 3147 (Draper *et al.* 2009). Functionally equivalent LtnI and LtnFE homologues were also produced by *Bacillus licheniformis* and *Enterococcus faecum*. Although resistance to lantibiotics in food applications has not presented any problems despite several decades of use, these observations raise concerns that need to be taken into account when lantibiotics are employed for clinical application (Draper *et al.* 2009).

I.4.3. The influence of recent developments in actinomycete research on classification of lantipeptides

As is the case for bacteriocin classification, dividing lantibiotics into distinct subclasses is the subject of some controversy regarding the classification criteria. Several recent studies, mostly in actinomycetes, have identified lanthionine-containing peptides that fulfil a physiological role for the producer organism other than antimicrobial activity (Kodani *et al.* 2004; Kodani *et al.* 2005; Goto *et al.* 2010). Thus, these compounds are related to but not considered to be true 'lanthionine containing antibiotics', hence the name 'lantipeptides' was proposed (Goto *et al.* 2010). We would like to encourage the

use of the term lantipeptide to describe 'lanthionine containing peptides' for which it is unknown whether they possess antimicrobial activity. This includes situations where antibacterial activity has been suspected, but not proven, for example when discussing cryptic lantipeptide gene clusters that are identified by genome mining. Once the product is isolated and shown to display antibacterial activity, it can then be designated a 'lantibiotic'.

A previously suggested classification scheme employed by Pag and Sahl (2002) was expanded by Willey and van der Donk (2007) to include the lantipeptides. This scheme uses the lanthionine synthetases as the main criterion for classification (Figure I.2.). Moreover, it is compatible with a further subclassification proposed by Cotter *et al.* (2005) that groups lantibiotic families according to their structural peptides. The three major groups that were proposed encompass the compounds that are modified by LanBC enzymes (Class I), LanM enzymes (Class II) and those that are modified by lanthionine synthetases that contain a Ser/Thr kinase domain (Class III). No generic 'Lan'-name has been designated to the Class III enzymes (RamC, LanL and LabKC) for historic reasons and because of the subtle differences between the different biosynthetic enzymes, as discussed below.

Class I and II contain the lantibiotics and most compounds in Class III are lantipeptides. Class I lantibiotics have a molecular weight ranging from 2164 Da to 3764 Da and either carry no net charge or up to seven positive charges. Class II lantibiotics are generally smaller than their type A counterparts. They have a compact, globular structure and a molecular weight ranging from 1959 Da to 2041 Da and carry no net charge or a net negative charge. Both types are strongly amphiphilic, but unlike the flexible Class I lantibiotics, the Class II peptides are more conformationally constrained because of overlapping bridging patterns. Two-peptide lantibiotics are also included in Class II since they are modified by LanM enzymes. No two-peptide lantibiotic produced by an actinomycete has been reported to date. Interestingly, both globular and more elongated peptides are produced by the Class III type of modification enzyme (Figure I.3.). Sequence inspection of several lantibiotic leader peptides has revealed two different consensus motifs (Nes and Tagg 1996). Class I lantibiotics contain a 'FNLD' motif between positions -20 and -15 and usually a Pro residue at position -2. These lantibiotics are post-translationally modified by LanB and LanC enzymes. Class II lantibiotics on the

other hand, contain a typical ‘GG’ or ‘GA’ cleavage site, several Asp and Glu residues and are modified by a LanM enzyme. No conserved sequence motif has yet been identified for Class III lantipeptides.

The classification of lantipeptides will further be exemplified in this section with an overview of the compounds produced by organisms belonging to the order of the Actinomycetales (Figure I.3.). Actinomycetes are Gram-positive bacteria with a high G+C content in their DNA. One of the most famous antibacterial, anti-fungal and anti-tumor compound producing genera within this order are the streptomycetes, which constitute a wide variety of soil-dwelling filamentous bacteria with a complex life cycle. The life cycle starts with a spore that germinates to yield an initial filamentous outgrowth that divides and branches to give a substrate mycelium. After this, an aerial mycelium is erected that will later in the developmental cycle septate to form chains of spores, hence completing the life cycle. Other actinomycetes are also well known for their production of secondary metabolites and several recent screens were set up aimed at the identification of novel compounds from these so-called ‘rare or uncommon actinomycetes’. Actinomycete lantipeptides contain a very wide range of post-translational modifications in addition to the commonly occurring dehydrated amino acids, (Me)Lan bridges and AviCys, making them an interesting object of study for enzymologists. Examples of these are mentioned below.

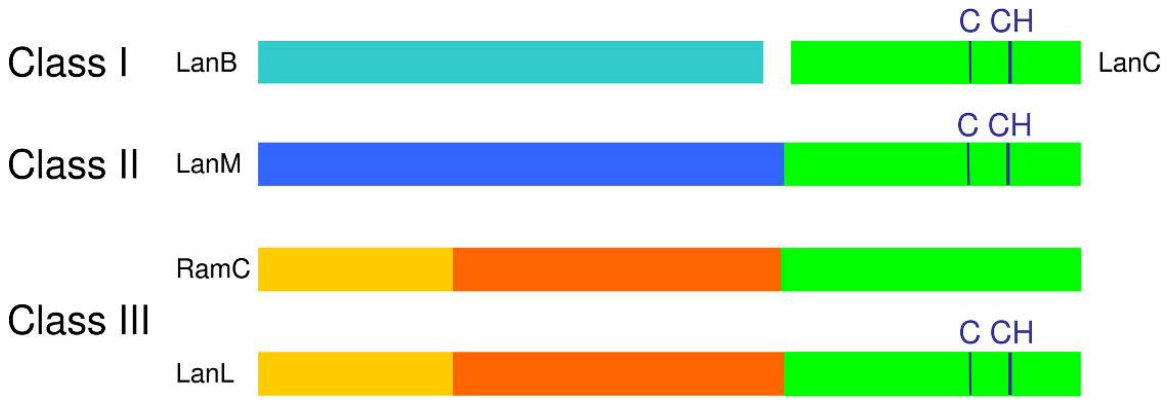


Figure I.2. Schematic representation of the different types of lanthionine synthetases. Homologous domains are depicted in the same colour. Class I lantibiotics are modified by a LanB dehydratase (light blue) and a LanC cyclase (green). The dehydration and cyclisation of Class II lantibiotics is performed by the bifunctional LanM enzyme, of which the C-terminal domain is homologous to LanC enzymes, but the N-terminal domain (dark blue) does not resemble LanBs. Class III enzymes consist of an N-terminal domain with homology to phosphoSer/Thr lyases (yellow), a middle domain resembling Ser/Thr kinases (orange) and a C-terminal LanC-like cyclase domain that lacks the conserved Zn^{2+} binding and catalytic residues. The LanC-like domain in LanL enzymes does contain these conserved residues.

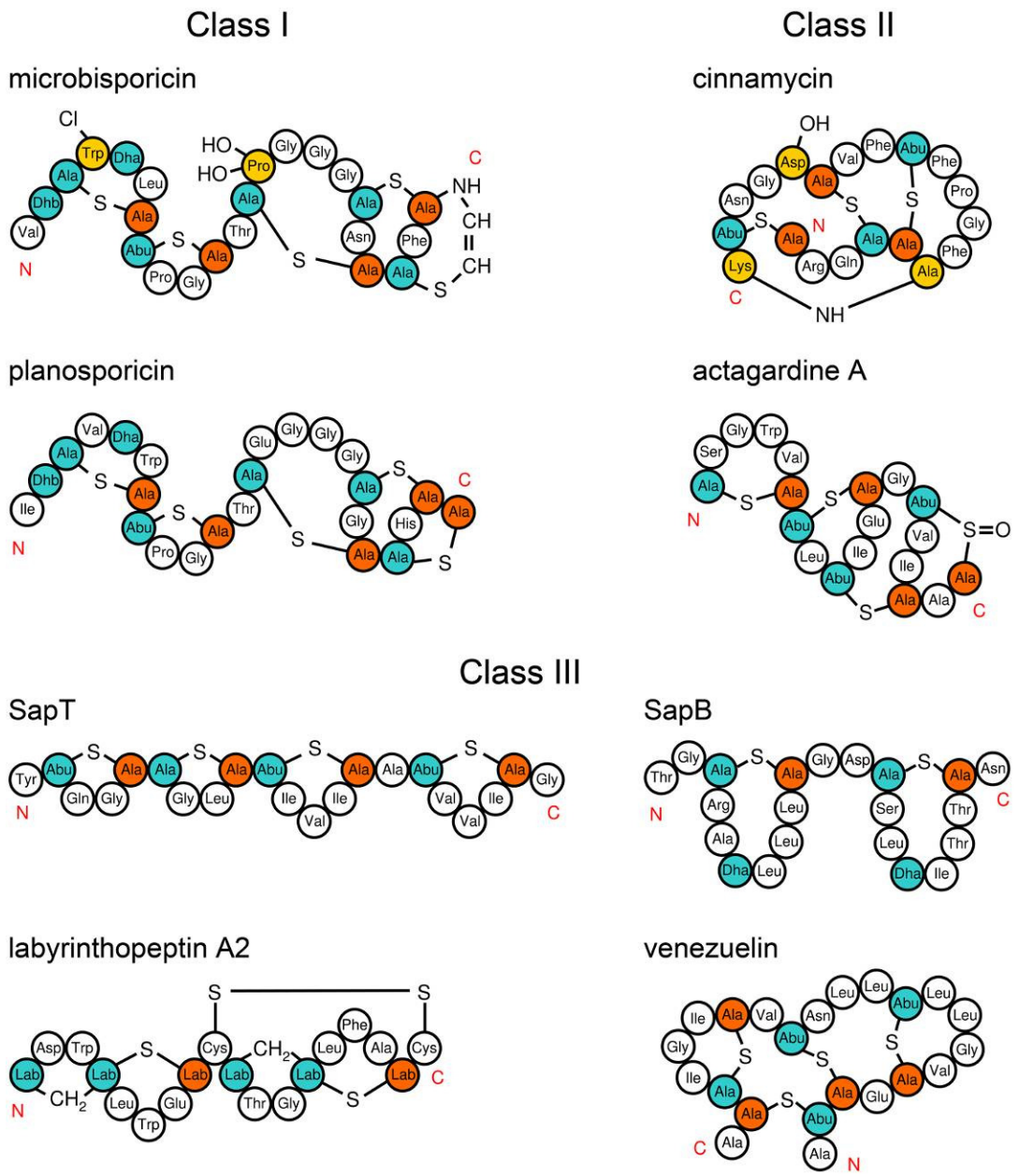


Figure I.3. Schematic representation of lantipeptides produced by actinomycetes. Compounds are classified based on the scheme proposed by Willey and van der Donk (2007), for more information refer to the text. The colour coding and abbreviations are as in Figure I.1.

I.4.3.a. Class I – LanBC (microbisporicin and planomonosporin)

Microbisporicin (also known as 107891 or NAI-107; Figure I.3.) is the collective name for four structurally related lantibiotics produced by *Microbispora corallina* (Lazzarini *et al.* 2005; Castiglione *et al.* 2008). The compound was identified in a large screen for peptidoglycan (PG) biosynthesis inhibitors and is the most potent lantibiotic characterised to date. Apart from the low MICs for important Gram-positive pathogens such as *Staphylococcus aureus*, *Enterococcus faecalis*, *Clostridium difficile* and *Propionibacterium acnes*, the inhibition spectrum also includes Gram-negative pathogens such as *Neisseria gonorrhoeae* and *Moraxella catarrhalis*, which is an uncommon feature for lantibiotics. Microbisporicin contains two post-translational modifications that have not previously been reported in lantibiotics, namely a 5-chlorotryptophan (Cl-Trp) and mono- or dihydroxylated proline (Castiglione *et al.* 2008). The sequence of the N-terminal eleven amino acids of microbisporicin resembles that of nisin, in which it was shown to be required for binding the pyrophosphate moiety of lipid II (Breukink *et al.* 1999; Hasper *et al.* 2006). Microbisporicin blocks peptidoglycan biosynthesis resulting in the accumulation of UDP-linked PG precursors in the cytoplasm of *Bacillus megaterium* (Castiglione *et al.* 2008). Recently, the microbisporcin biosynthetic gene cluster was sequenced and cloned, providing the first representative of a Class I gene cluster of actinomycete origin (Foulston and Bibb 2010). Mutational analysis identified the product of *mibH* as the flavin-dependent Trp halogenase that is responsible for the introduction of the Cl-Trp residue. The non-chlorinated form of microbisporcin could have enhanced activity, judging from the comparison of the zones of inhibition against *M. luteus* of the Δ*mibH* mutant to the WT. This study also revealed that the gene cluster is under regulatory control of a pathway-specific extracytoplasmic function σ factor-anti-σ factor complex (MibX-MibW). The mechanism for self-immunity is likely through MibEF-mediated extrusion from the cell membrane, since deletion of *mibEF* resulted in greatly reduced microbisporicin export, possibly reflecting reduced gene expression (Foulston and Bibb 2010).

Planosporicin (also known as 97518; Figure I.3.) was previously classified as a globular Class II mersacidin-like lantibiotic, based on its initial structural determination (Castiglione *et al.* 2007). However, more recent chemical data led to a structure revision involving a rearrangement two (Me)Lan bridges and a shift of the two C-terminal amino

acids to positions nine and ten in the new structure. This data is backed up by the determination of the structural gene sequence by genome scanning (Maffioli *et al.* 2009). Like microbisporicin, planosporicin exerts its mechanism of action by blocking PG biosynthesis (Castiglione *et al.* 2007).

I.4.3.b. Class II – LanM (cinnamycin and the duramycins, actagardine and michiganin A)

No two-peptide lantipeptide of actinomycete origin has been characterised to date. The best characterised two-peptide lantibiotic is lacticin 3147 (Ryan *et al.* 1996; Ryan *et al.* 1999; Martin *et al.* 2004; Cotter *et al.* 2006; Field *et al.* 2007), but several others have been identified, including staphylococcin C55 (Navaratna *et al.* 1998; Navaratna *et al.* 1999), plantaricin W (Holo *et al.* 2001), haloduracin (McClerren *et al.* 2006; Lawton *et al.* 2007; Cooper *et al.* 2008), lichenicidin (Begley *et al.* 2009) and pneumococcin (Majchrzykiewicz *et al.* 2010). Two-peptide lantibiotics have been the subject of recent review (Breukink 2006; Lawton *et al.* 2007).

Cinnamycin (formerly also known as Ro 09-0198 and lanthiopeptin; Figure I.3.) is produced by several *Streptomyces* strains, including *Streptomyces cinnamoneus* and contains unusual amino acid modifications not commonly found in other lantibiotics, namely β-hydroxy aspartate and a lysino-alanine bridge. Another interesting feature of cinnamycin is its mode of action, which is proposed to be exerted via binding of phosphatidylethanolamine (PE) (Hosoda *et al.* 1996; Machaidze *et al.* 2002). Binding requires the PE headgroup and at least one fatty acyl chain; the latter needs to have a chain length of eight carbons or larger for optimal binding (Machaidze and Seelig 2003). Apart from their antimicrobial activity, cinnamycin and the structurally related duramycins have several other properties that are of potential pharmaceutical interest. The compounds can act as an inhibitor for phospholipase A2 (Fredenhagen *et al.* 1990), angiotensin-converting enzyme (Kido *et al.* 1983) and prostaglandin and leukotriene biosynthesis. Inhibition of *Herpes simplex* virus (HSV-1) has also been reported (Naruse *et al.* 1989). Duramycin (Moli1901) is currently in phase II clinical trials for cystic fibrosis treatment (Grasemann *et al.* 2007). *cinA w*as the first class II structural gene to be characterised and this revealed that the propeptide is preceded by a leader sequence of 58 amino acid residues, which was longer than the class I leaders known at that time (Kaletta *et al.* 1991). This study also established that the lysino-alanine bridge is formed

from Lys and Ser (Dha) precursors in the propeptide. The cinnamycin gene cluster was the first lantipeptide gene cluster from a high G+C bacterium to be characterised (Widdick *et al.* 2003). This gene cluster was unusual since it contained several genes generally not found in lantibiotic clusters, including one that encodes a *Streptomyces* antibiotic regulatory protein (SARP). The ABC-transporter encoded by *cinT* and *cinH* is thought to be involved in resistance and/or secretion. A dedicated LanP protease appears to be missing, implying the requirement for a different activation mechanism. Closer inspection of the cinnamycin leader peptide cleavage site reveals the AXA motif, which might function as a recognition site for Type I signal peptidases involved in Sec-mediated secretion. The *cinM* gene contains the rare streptomycetes TTA codon, which relies on the *bldA* encoded Leu-tRNA for translation in *S. coelicolor* (Leskiw *et al.* 1991). Widdick *et al.* (2003) constructed expression cassettes for the cinnamycin cluster and succeeded in utilising the cinnamycin machinery to produce duramycins A and B (each differing from cinnamycin by only one amino acid residue). Production of the more distantly related duramycin C (obtained from the primary structure of cinnamycin by altering six amino acids) was subsequently demonstrated (J. Cortes, personal communication). These results illustrate the relaxed substrate specificity of the modification machinery and are encouraging for lantibiotic engineering.

Actagardine A (formerly known as gardimycin; Figure I.3.) was identified from the fermentation broth of *Actinoplanes garbadinensis* ATCC 31049 (Parenti *et al.* 1976). The compound was shown to inhibit peptidoglycan biosynthesis (Somma *et al.* 1977) and its mechanism of action later attributed to lipid II binding, thereby blocking transglycosylation in susceptible bacteria (Héchard and Sahl 2002). The structure of actagardine A was determined by NMR and showed the presence of one Lan and three MeLan ring structures that impose a compact, globular structure on the compound (Zimmermann and Jung 1997). The C-terminal MeLan ring in actagardine A is oxidised to a sulfoxide. Ala(0)-actagardine A, a variant of actagardine A with an additional N-terminal Ala was identified and shown to have a slightly enhanced antibacterial activity (Vértesy *et al.* 1999). The gene cluster was recently cloned and a variant generation system was developed that allowed for alanine scanning of the compound (Boakes *et al.* 2009). Deletion of the monooxygenase *garO* from *A. garbadinensis* resulted in a strain that only produced unoxidised actagardine A, indicating that introduction of the sulfoxide is catalysed by GarO (Boakes *et al.* 2009). The gene cluster of a natural variant,

deoxyactagardine B (DAB), was reported from *Actinoplanes liguriae* NCIMB41362 (Boakes *et al.* 2010). Deoxyactagardine B lacks the sulfoxide present in actagardine A and differs in two amino acids. A monooxygenase with significant homology to *garO* was identified in the gene cluster and expression of the structural gene *ligA* in *A. garbadinensis* Δ*garA* resulted in production of the oxidised form of DAB (Boakes *et al.* 2010). This suggests that the LigO monooxygenase is inactive or not expressed rather than lack of oxidation because of the two amino acid substitutions at positions 15 and 16. The mechanism of self-immunity in the actagardine A and DAB producers is unknown, but a penicillin-binding protein and a putative transporter have been proposed for this function for actagardine A and DAB, respectively (Boakes *et al.* 2009; Boakes *et al.* 2010). It is noteworthy that these putative immunity systems are different in the two strains, indicating that neither may be involved in self-protection and that a separate immunity mechanism may exist. It is unclear why actagardine A is oxidised on ring D, since a recent study related the oxidation of Lan in nisin with greatly reduced antibacterial activity (Wilson-Stanford *et al.* 2009). The authors proposed that oxidised nisin lost its activity due to its inability to bind lipid II. The study did not address the effect of one Lan oxidation as nisin always had seven, eight or nine oxidations and contains only five bridges. Whether the loss in activity was primarily due to Lan oxidation or the result of additional Met or His oxidations is unknown. In contrast, the antibacterial activity of actagardine A is slightly higher than that of the non-oxidised compound produced by the Δ*garO* strain of *A. garbadinensis*, implying a role for the sulfoxide in activity or stability.

Michiganin A is produced by the tomato pathogen *Clavibacter michiganensis* and its structure is related to actagardine A (Holtsmark *et al.* 2006). Michiganin A is longer by two residues (an N-terminal Ser and a C-terminal Arg) and has two amino acid substitutions (Leu5Val and Ile15Val) compared to actagardine. Interestingly, the first three amino acids of the propeptide are all Ser, but only the second one becomes dehydrated and is involved in formation of the A ring (Holtsmark *et al.* 2006). Actagardine A also contains a Ser that escaped dehydration C-terminal of the one involved in Lan ring formation. The bridging pattern for michiganin A was assigned based on the structure of actagardine A. Remarkably, the Edman degradation data indicated that the Dhb residue at position eight in the michiganin A propeptide is not involved in (Me)Lan bridge formation, leaving a free sulfhydryl group on one Cys

(Holtsmark *et al.* 2006). The β peptide of the two-peptide lantibiotic plantaricin W is the only other example of a lantibiotic with a free Cys residue described to date (Holo *et al.* 2001).

I.4.3.c. Class III – Lanthionine synthetases containing a Ser/Thr kinase domain

This class was first proposed by Willey and van der Donk (2007). Despite the bioinformatic identification of putative Class III gene clusters in other bacterial phyla (Goto *et al.* 2010), all lantipeptides characterised to date are from actinomycete origin. These gene clusters are of minimal composition, in most cases just consisting of a structural gene, a lanthionine synthetase gene, (two) transporter genes and, in the case of the morphogens, a transcriptional regulatory gene. The simplicity of these gene clusters is reflected by the fact that none of the resulting compounds contain additional post-translational enzymatic modifications, just the expected dehydrated Ser/Thr or (Me)Lan bridges. The introduction of the unusual modified amino acid labionin in the labyrinthopeptins does not require additional enzyme activity. Further division into three subclasses is presented here because some enzymes contain the typical catalytic and Zn-binding residues in their cyclase domain and others introduce labionin in their associated structural peptides. In general, the lantipeptides belonging to Class III appear to have a function distinct from antibacterial activity.

SapB (Figure I.3.) was previously known as a small peptide involved in morphological differentiation of *S. coelicolor* into an aerial mycelium (Willey *et al.* 1991). Its working mechanism is that of a surfactant, which allows the hyphae to escape from the aqueous environment of the colony surface (Tillotson *et al.* 1998). Among the genes known to be involved in morphological developmental are the *bld* genes (with mutants unable to form an aerial mycelium) and the *ram* genes (mutant forms of which confer *r*apid *a*erial *m*ycelium formation in *S. lividans*) (Ma and Kendall 1994; Kelemen and Buttner 1998; O'Connor *et al.* 2002). Kodani et al. (2004) discovered that SapB is produced as a post-translationally modified product of the prepropeptide encoded by *ramS*, which lies in the *ramCSAB*-operon. Although the C-terminus of RamC has some sequence similarity to the C-terminal domain of CinM and MrsM, the protein has a unique structure which is not found in any other lantibiotic modification enzymes. The N-terminus of RamC exhibits similarity to catalytic domains of Ser/Thr kinases, a feature that is not found in LanM-like

or LanB-like enzymes. This domain is thought to be involved in phosphorylation of the hydroxyl moiety (which is a poor leaving group) in the Ser residues, thus facilitating dehydration. Interestingly, whereas four of the five Ser residues in RamS are dehydrated, none of the three Thr residues are (Kodani *et al.* 2004). This is a remarkable feature since Ser residues in lantibiotics are more likely to escape dehydration than Thr (Rink *et al.* 2005). The C-terminal domain is predicted to catalyse subsequent Lan bridge formation, despite the lack of the zinc binding and catalytic residues present in LanC-like domains (Li *et al.* 2006). RamA and RamB are components of an ABC-transporter which are proposed to be involved in export of SapB (Kodani *et al.* 2004). The product of the convergently transcribed *ramR* is a response regulator that binds the *ramC* promoter, inducing transcription of the *ramCSAB* operon (Keijser *et al.* 2002; O'Connor *et al.* 2002). The identification of SapB as a lantipeptide was the first report of such a compound involved in a structural function rather than exhibiting antibiotic activity.

A more recent study in *Streptomyces tendae* revealed another morphogenetic lantipeptide, SapT (Figure I.3.), which is able to restore formation of aerial hyphae in developmentally impaired *S. coelicolor* mutants (Kodani *et al.* 2005). Interestingly, the three Thr in SapT are involved in MeLan formation, whereas in SapB all three Thr residues escaped dehydration. Modelling of the SapT structure based on SapB predicts a similar amphiphilic nature, indicating this peptide also serves as a biosurfactant. The primary structure of both peptides is quite different, but their function appears to be conserved. The characterisation of a second distinct morphogen illustrates the broad diversity present in the world of lanthionine-containing peptides. These morphogens appear to be quite common among streptomycetes, as all genome sequences currently available contain at least one gene cluster likely to encode a Sap-like peptide. AmfS, the SapB counterpart in *S. griseus*, has been the subject of several studies but still awaits structural characterisation (Ueda *et al.* 2002; Ueda *et al.* 2005). Morphogenic peptides that do not contain labionin residues are proposed to constitute Subclass III.1.

Venezuelin (Figure I.3.) was produced biochemically in an *in vitro* reaction with purified VenL synthetase and VenA prepropeptide, hereby uncovering a novel biosynthetic route to generate lantipeptides (Goto *et al.* 2010). VenL, the first characterised LanL-type lantipeptide synthetase, was identified bioinformatically in an *S. venezuelae* cryptic gene cluster. LanL enzymes contain three distinct catalytic domains, an N-terminal

phosphoSer/Thr lyase, a Ser/Thr kinase and a C-terminal LanC-like cyclase domain that contains a characteristic zinc-binding site and active-site residues. The activity of the three domains was dissected by individual expression and this revealed that Ser/Thr residues are first phosporylated by the kinase domain, followed by an elimination catalysed by the lyase domain, which results in Dha and Dhb residues. Subsequent addition of the Cys thiols onto these dehydrated residues is performed by the cyclase domain, yielding (Me)Lan bridges. The 'artificially' constructed VenL-ΔC is the first example of an *in vitro* reconstituted peptide dehydratase, since before then only bifunctional lantipeptide synthetases had been reconstituted *in vitro* (Goto *et al.* 2010). The venezuelin ring topology was determined by site-directed mutagenisis of individual Cys residues and venezuelin's structure most closely resembles cinnamycin. Since the cyclase domain of LanL enzymes is different from the other Class III enzymes, they are classified in a separate Subclass III.2.

The labyrinthopeptins (Figure I.3.) were identified recently as a product of *Actinomadura namibiensis* (Meindl *et al.* 2010). Labyrinthopeptin A2 has activity against *Herpes simplex* and displayed potent activity in an *in vivo* mouse model system for neuropathic pain. The structure of labyrinthopeptin A2 was determined by X-ray crystallography and revealed a novel type of bridge structure, (2*S*,4*S*,8*R*)-labionin (Lab). This carbacyclic triamino acid residue results from the post-translational modification of two Ser and a Cys by the Class III kinase-cyclase LabKC. Labyrinthopeptin A2 has a globular structure and contains two Cys residues that are not involved in Lab-formation, but instead form a disulfide bond (Meindl *et al.* 2010). The activity of LabKC has been reconstituted *in vitro* and a model was proposed for Lab formation (Müller *et al.* 2010). LabKC was shown to require guanosine triphosphate (GTP) for Ser phosphorylation and dehydration. In other *in vitro* lantipeptide biosynthesis reconstitutions ATP has been used as phosphate donor, even with the closely related VenL (Xie *et al.* 2004; McClerren *et al.* 2006; Goto *et al.* 2010). This difference in nucleotide requirement and the occurance of Lab residues in the peptide products likely reflects currently unidentified differences in the LabKC biosynthetic enzyme compared to RamC and AmfT, justifying classification of the labyrinthopeptins in a distinct Subclass III.3. The laybrinthopeptin biosynthetic gene cluster closely resembles the SapB and venezuelin gene clusters, but no *ramR* homologue is present downstream of the two putative transport genes. Two structural genes are present, the first of which encodes LabA1/A3 and the other LabA2 (Meindl *et*

al. 2010). The presence of Lab (and the disulfide bridge) makes these compounds unique within Class III, for which no additional post-translational modifications have been described to date apart from the common Dha/Dhb and (Me)Lan residues.

I.4.4. Mechanisms of action

Lantibiotics generally exhibit activity against Gram-positive bacteria, whereas Gram-negative bacteria are only affected when their outer membrane is disturbed (Kordel *et al.* 1988), for example by chelating agents such as EDTA. A great variety has been described in the range of susceptible organisms for different lantibiotics. Nisin displays a broad activity and even has an inhibitory effect on germination of *Bacillus* and *Clostridium* spores. In contrast, other lantibiotics only show a narrow inhibition spectrum (e.g. salivaricin A). Most Class I lantibiotics exert their activity by pore formation in the cytoplasmic membrane. For example, nisin binds lipid II with high affinity, and formation of pores consisting of eight nisin molecules and four lipid II molecules has been demonstrated by pyrene fluorescence (Breukink *et al.* 1999; Hasper *et al.* 2004). On the other hand, Class II lantibiotics act as inhibitors of different enzyme functions, most of which are involved in cell wall biosynthesis.

The pore forming activity of Class I lantibiotics causes rapid leakage of ions, cytoplasmic solutes, amino acids and nucleotides from the cytoplasm. Moreover, this efflux causes a depolarisation of the cell membrane, resulting in an instant termination of all biosynthetic processes (Sahl *et al.* 1987). Studies with artificial membranes indicated that Class I lantibiotics adopt a helical amphiphilic conformation with their charged residues on one side and their hydrophobic residues on the other side of the helix. The hydrophobic face interacts with the phospholipid head groups of the cell membrane and causes a local disruption in the bilayer structure. Lipid II functions as a docking molecule for specific binding, making the process energetically more favourable. At the same time, the hydrophobic residues insert in the hydrophobic core of the membrane, forming small pores and thus achieving leakage of small cellular compounds (Breukink *et al.* 2003). Nisin was shown to be a dual-function antibiotic with two killing mechanisms: preventing lipid II incorporation in the peptidoglycan and simultaneously using lipid II for targeted pore formation (Hasper *et al.* 2006).

Two different mechanisms of action have been described for Class II lantibiotics, involving enzyme inhibition. A first mechanism is characteristic of mersacidin and actagardine. These lantibiotics cause a block in the transglycosylation step of cell wall biosynthesis, resulting in a growth defect and the onset of cell lysis. Other fundamental cellular biosynthetic processes (DNA, RNA and protein biosynthesis) remain unharmed (Brötz *et al.* 1997).

A group of Class II lantibiotics, containing cinnamycin, the duramycins and ancovenin, exerts antimicrobial activity against relatively few bacterial strains. Susceptible strains showed increased membrane permeability and reduced ATP-dependent calcium uptake and protein transport (Sheth *et al.* 1992). Cinnamycin has been shown to cause lysis of red blood cells and duramycin is involved in inhibition of phospholipase A2, which is involved in the synthesis of prostaglandins and leukotrienes. Both of these lantibiotics have been demonstrated to interact with phosphatidylethanolamine (Choung *et al.* 1988).

The two-peptide lantibiotics form a distinct subgroup of Class II lantibiotics. Members of this group consist of two different lantibiotic peptides that display synergistic antibacterial modes of action. Among the two-peptide lantibiotics characterised to date, lacticin 3147 is the best studied. Moreover, the structure of both peptides was the first to be elucidated by using the sodium borodeuteride derivatisation process described by Martin *et al.* (2004). The lacticin 3147 gene cluster contains two distinct *lanM* genes, encoding for LtnM1 and LtnM2. These two genes were independently disrupted demonstrating that each prepropeptide (LtnA1 or LtnA2) required a dedicated LtnM gene for modification (McAuliffe *et al.* 2000). The mode of action for lacticin 3147 is similar to that for Class I lantibiotics, however in the two-peptide lantibiotic, each of the two components possesses a single functionality. The Ltnα peptide is involved in targeting of Lipid II and upon interaction, a binding site for the Ltnβ peptide is generated. Subsequently, the Ltnβ peptide binds and initiates pore formation in the cytoplasmic membrane. Thus the two peptides operate synergistically to produce a more efficient killing mechanism than the greatly reduced activity of Ltnα and complete lack of activity of Ltnβ when used alone (Wiedemann *et al.* 2006). The Ltnα peptide of lacticin 3147 shares a conserved region with certain Lipid II-interacting Class II lantibiotics. Breukink

(2006) hypothesised that this region might be responsible for the Lipid II interaction. Other two-peptide lantibiotics are expected to work in a similar way to lacticin 3147.

I.4.5. Engineering lantipeptides

Lantibiotics are ideal candidates for engineering of variants because of their peptidic nature and their ribosomal origin. This is exemplified by many successful studies aimed at improving antibacterial activity or pharmaceutical properties and understanding the structure-activity relationships of these compounds. Changes in residues that are involved in (Me)Lan formation have mostly led to either a reduction in or complete loss of activity of the respective lantibiotic (Ottenwalder *et al.* 1995; Bierbaum *et al.* 1996; Cotter *et al.* 2006; Cooper *et al.* 2008), while some other residues are more tolerant to variation. Several reviews have been published recently that adress this emerging topic in lantibiotic research (Bierbaum and Sahl 2009; Cortés *et al.* 2009).

Generation of lantibiotic variants was pioneered in the early nineties by several groups. A site-directed mutagenesis technique was used to generate a series of variant *nisA* genes that were subsequently expressed using a lactococcal expression system in a Δ*nisA* host strain (Dodd *et al.* 1992). Kuipers *et al.* (1992) developed a system that allowed for expression of PCR-mutagenised *nisZ* variants in a nisin A producing host strain. A Glu4Ile variant of subtilin was shown to stabilise degradation of the neighbouring Dha5 57-fold, resulting in a lantibiotic with prolonged activity (Liu and Hansen 1992). Chimeras of nisin and subtilin have been successfully generated, using the biosynthetic machinery of either of these two lantibiotics (Kuipers *et al.* 1993; Chakicherla and Hansen 1995). An expression system was also constructed for Pep5 and two variants were generated, one of which was produced at wild type levels, but the other only in low quantities, indicating that not all amino acids in the propeptide are equally amenable to modification (Bierbaum *et al.* 1994). An expression system was developed to generate epidermin and gallidermin variants and applied to study the effect of a change in key amino acid residues on production, activity and stability (Ottenwalder *et al.* 1995). A Leu6Val variant exhibited enhanced antimicrobial activity and two others (Ala12Leu and Dhb14Pro) proved to be more resistant to proteolytic degradation. Interestingly, the authors noted a strong decrease in production when the Ser residues, involved in Lan formation, were replaced by Thr.

Several systems have been described that are employed to produce structurally related lantibiotics. The cinnamycin biosynthetic gene cluster has been heterologously expressed in *S. lividans* and this system has been used to generate the related duramycin and duramycin B compounds (Widdick *et al.* 2003). The lacticin 481 synthetase was shown to be capable of producing related lantibiotics, including ruminococcin A, nukacin ISK-1 and mutacin II (Patton *et al.* 2008).

The complete Ala scanning of the two-peptide lantibiotic lacticin 3147 was the first thorough systematic mutational analysis of a lantipeptide (Cotter *et al.* 2006). This study identified residue locations essential for efficient production, but also 36 out of 59 amino acids were found to be more amenable to change, as their replacement by Ala did not result in a complete loss of activity. An additional two-plasmid system was generated that allowed construction of a 1500-member random mutagenesis mutant lacticin 3147 library (Field *et al.* 2007). A similar approach was successfully applied to generate nisin variants with elevated antimicrobial activity against specific pathogens (Field *et al.* 2008).

A system for mersacidin engineering was constructed and led to the identification of two key residues in the peptide (Szekat *et al.* 2003). A more recent study described a thorough saturation mutagenesis approach that resulted in a 228 *mrsA* mutant library in which all amino acid residues that are not involved in (Me)Lan formation were replaced by the other proteogenic amino acids (Appleyard *et al.* 2009). About 35 % of the engineered variants were produced in good yield, resulting in 82 new compounds, most of which were either inactive or displayed a reduced antibacterial activity. However, several compounds had increased activity against specific Gram-positive pathogens, illustrating the potential for this approach to enhance lantibiotic activity.

A heterologous expression system was developed for the nukacin ISK-1 biosynthetic gene cluster under the control of the nisin-inducible promoter P*nisA* (Nagao *et al.* 2007). This study also described a new method for the introduction of post-translational modifications into 6His-tagged NukA via co-expression with NukM in *E. coli*.

A variant generation system has also been engineered for actagardine A and was used to do alanine scanning of the residues not involved in (Me)Lan formation (Boakes *et al.* 2009). Five variants were produced by *A. garbadinensis* (out of ten constructs) and two

of these retained some antibacterial activity. The B ring of the compound proved to be least suited for Ala replacement since none of these variants were produced. This result is in agreement with data from the Ala scanning experiment in the analogous region of lacticin 3147 (Cotter *et al.* 2006).

Although it cannot be considered 'lantibiotic engineering' per se, the use of lantibiotic-derived post-translational modifications to stabilise peptide pharmaceuticals illustrates another application of fundamental research on lantibiotic biosynthesis. These applications are enabled by the relatively broad substrate specificity of lantibiotic modification enzymes. Rink *et al.* (2005) performed a bioinformatic study on 37 lantibiotics, from which they distilled guidelines for *lanA* engineering. One important conclusion of this study was that hydrophobic neighbouring amino acids promote dehydration of Ser and Thr. This analysis also showed that Ser was more likely to escape dehydration than Thr. However, the analysis has to be approached with caution, since it does not distinguish between the different lantibiotic classes and the guidelines might vary with the substrate specificity of the respective modification enzymes (Appleyard *et al.* 2009). An engineered *L. lactis* strain containing *nisBCT* was used to dehydrate, cyclise and export several non-lantibiotic peptides of medical significance (Kluskens *et al.* 2005). The lacticin 481 synthase LctM was used to install dehydrated residues and ring structures into peptides *in vitro* (Chatterjee et al. 2006). An expression system was developed that utilised the *L. lactis* nisin biosynthetic machinery for production of a Luteinizing-Hormone-Release-Hormone (LHRH) analog containing a Lan bridge between residues four and seven (Rink *et al.* 2010). This modification conferred superior protease resistance on the LHRH analog and was compatible with subsequent *in vitro* derivatisation.

I.4.6. Genome mining to identify lantipeptide compounds

The recent surge in publicly available bacterial genome sequences has led to the identification of several cryptic gene clusters that potentially encode for interesting compounds. Several research groups have embraced the idea of lantipeptide genome mining, resulting in the identification of various compounds that would not necessarily have been found in an activity-based screening approach. A cryptic lantipeptide cluster can be of particular interest because the product may be related to potent known

compounds, because of a predicted novel lantipeptide scaffold, or even because the cluster is predicted to encode a novel biochemical mechanism of post-translational modification. Various high throughput tools will be required that will help to cope with the increasing amounts of bacterial genome data, but current techniques are already very promising. Lantibiotic detection by MALDI-ToF MS on whole cell samples can make compound screens more efficient since less handling is required (Hindré *et al.* 2003). Bioinformatic mining of genomic data for gene clusters involved in production of lantipeptides and non-modified bacteriocins can be performed with the BAGEL2 software (de Jong *et al.* 2010). This section describes both biochemical and molecular genetic approaches aimed at the identification of a product from cryptic lantipeptide gene clusters.

I.4.6.a. Biochemical approaches to genome mining

One of the first successful genome mining attempts was the identification of the two-peptide lantibiotic haloduracin from a cryptic cluster found in *Bacillus halodurans* (McClerren *et al.* 2006). Both LanM enzymes were purified from a heterologous *E. coli* host and shown to modify the purified HalA1 and HalA2 precursor peptides *in vitro*. A structure was proposed for the resulting modified peptides and they were shown to display antimicrobial activity when both of them were used together, a common feature of two-peptide lantibiotics. A further structure-activity relationship study identified the Halα B ring as expendable and that the Halα cystine is not required for activity, but protects the peptide against proteolytic degradation (Cooper *et al.* 2008). Two Ser residues were shown to escape dehydration, requiring the previous structure to be revised. *In vivo* production of haloduracin was also shown by purification from a *B. halodurans* fermentation and by mutational analysis (Lawton *et al.* 2007).

A similar biochemical approach was taken for a cryptic gene cluster in *S. venezuelae*, which led to the identification of venezuelin (Goto *et al.* 2010). This case provides an example of a gene cluster that was of interest because of the unusual modification enzyme VenL. Venezuelin does not appear to have antibacterial activity and *in vivo* production has not yet been demonstrated. The genetic analysis of venezuelin production is a subject of this thesis and is discussed in more detail in Chapter VII.

A bioinformatic search designed to find examples of bacteria that can produce multiple lantipeptides with the same LanM enzyme identified several clusters in marine planktonic cyanobacteria (Li *et al.* 2010). The genome of *Prochlorococcus* MIT9313 contains only one *lanM* homologue, but seven putative *lanA* genes in this cluster along with 22 additional *lanA* genes elsewhere in the genome. The prepropeptides display a remarkably high sequence identity in the leader region, but their propeptides (ranging from 13 to 32 amino acids in length) have very diverse sequences. An *in vitro* dehydration and cyclisation assay with purified ProcM and seventeen ProcA precursors showed that this enzyme was capable of efficient modification of this wide variety of substrates. Several of these, collectively called 'prochlorosins', were shown to be produced by *Prochlorococcus* MIT9313 *in vivo* (Li *et al.* 2010). This is a very nice example of how an organism uses a relatively simple biochemical system to generate a broad diversity of secondary metabolites.

I.4.6.b. Molecular genetic approaches to genome mining

The genome sequences of clinical isolates of *S. aureus* revealed the presence of a cryptic lantipeptide gene cluster (Daly *et al.* 2010). Reverse genetics and a mutational analysis of this cluster were performed to produce the Bsa (for bacteriocin of *S. aureus*) lantibiotic, which turned out to be identical to the previously identified, but not structurally characterised, bacteriocin staphylococcin Au-26 (Scott *et al.* 1992). Bsa is structurally related to gallidermin, but its immunity genes do not contribute significantly to gallidermin immunity.

An *in silico* screening strategy of the public databases identified 89 LanM homologs, including 61 in strains that were known as lantibiotic producers. A cryptic gene cluster in *Bacillus licheniformis* was shown to produce a two-peptide lantibiotic, lichenicidin, which exhibits antimicrobial activity against important pathogens such as MRSA, VRE and *Listeria monocytogenes* (Begley *et al.* 2009). Insertion inactivation mutants were generated for the two *lanM* genes, linking them to the production of lichenicidin (Dischinger *et al.* 2009). A lichenicidin cluster with identical structural genes was identified in a different *B. licheniformis* strain and the structure of both the Lchα and Lchβ peptides were determined by NMR (Shenkarev *et al.* 2010).

An elegant system was devised that employed the nisin biosynthetic machinery to produce the previously uncharacterised two-peptide lantibiotic pneumococcin (Majchrzykiewicz *et al.* 2010). This demonstrated that the Class I nisin biosynthetic machinery was capable of post-translational modification of Class II propeptides when fused to the nisin leader sequence. Although this was a very nice proof of principle, it was also rather adventurous, since the use of the original or more closely related biosynthetic machinery intuitively seems to have a greater chance of success for correct production of cryptic lantipeptides. Indeed, several versions of both peptides were detected, each with different number of dehydrated residues (Majchrzykiewicz *et al.* 2010). This could be indicative of inefficient modification, which in turn could have implications for correct (Me)Lan formation. The two modified pneumococcin peptides displayed antibacterial activity against *Micrococcus flavus*, but they were not found to act synergistically, as commonly observed for two-peptide lantibiotics. Another issue is the introduction of modifications that do not naturally occur in nisin, such as AviCys for example.

I.5. Microcins

Microcins are a class of low-molecular weight (< 10 kDa) ribosomal peptides produced by enterobacteria to kill off closely related competitors. Some microcins are unmodified and will not be discussed here. Others contain extensive post-translational modifications, resulting in a wide variety of structures (Figure I.1.).

I.5.1. Microcin B17 contains thiazole and oxazole heterocycles

Microcin B17 (MccB17; Figure I.1.) is a DNA gyrase inhibitor produced by *E. coli* strains that contain the pMccB17 plasmid (Baquero *et al.* 1978; Yorgey *et al.* 1994). The first four coding sequences in the seven gene MccB17 biosynthetic operon, *mcbABCD*, were identified as essential for biosynthesis (Genilloud *et al.* 1989), whereas *mcbEFG* are required for export and immunity (Garrido *et al.* 1988). The peptide is post-translationally modified by heterocyclisation of Gly, Ser and Cys residues, resulting in the formation of four thiazole and four oxazole moieties. Its biosynthesis was reconstituted *in vitro* by purification of the McbBCD synthetase and incubation with McbA precursor peptides, providing the first example of a biochemically produced modified peptide (Li *et al.* 1996).

The purified synthetase requires ATP or GTP for activity and purification of the individual subunits identified the cofactor requirements and putative functions of each enzyme (Milne *et al.* 1998; Milne *et al.* 1999). The cyclodehydratase McbB is a Zn-containing protein that catalyses the initial attack of a Ser or Cys side chain onto the upstream amide carbonyl in the peptide chain. Loss of a water molecule results in formation of thiazoline and oxazoline rings, which are desaturated by the FMN-containing protein McbC. McbD is an ATPase/GTPase that is proposed to act as a conformational switch. The combination of all three subunits was required for *in vitro* heterocyclisation.

I.5.2. Lasso peptides

Microcin J25 (MccJ25; Figure I.1.) is a 21 amino acid peptide that is produced by *E. coli* strains containing pTUC100 (Salomon and Farias 1992). Its structural gene, *mcjA*, encodes a 58 amino acid precursor with an N-terminal leader sequence (Solbiati *et al.* 1999) and a C-terminal propeptide that becomes modified into a structure resembling a knot or 'lasso' (Bayro *et al.* 2003; Wilson *et al.* 2003). A loop is formed by linkage of the carboxylate side chain of Glu8 with the N-terminal Gly1 amino group. The C-terminal tail of the peptide is threaded through this ring structure and kept tightly into place by the side chains of Phe19 and Tyr20. The activity of McjB and McjC, which are both required for MccJ25 maturation, has recently been reconstituted *in vitro* (Duquesne et al. 2007). McjC is an ATP/Mg^{2+} -dependent enzyme that shares homology with class B asparagine synthetases and β-lactam synthetases, suggesting that it likely introduces the amide bond resulting in formation of the loop. McjB is homologous to transglutaminases and is proposed to possess the proteolytic activity for leader cleavage. MccJ25 exerts its antibacterial activity through inhibition of RNA polymerase and stimulation of reactive oxygen species production (Bellomio *et al.* 2007). Truncation of the 36 amino acid MccJ25 leader sequence revealed that a residual sequence of eight was sufficient to process the peptide, albeit at very low efficiency (Cheung *et al.* 2010). Deletion of the first 16 residues only led to a five-fold reduced production.

Several members of the lasso peptide family have also been identified in actinobacteria belonging to the genera *Streptomyces*, *Rhodococcus* and *Microbispora*. MccJ25 and lariatins A and B are the only lasso peptides with antibacterial activity (Iwatsuki *et al.* 2006), but different biological functions have been described for other members. For

example, anatin acts as an atrial natriuretic factor antagonist (Weber *et al.* 1991), RES-701s is an endothelin receptor antagonist (Morishita *et al.* 1994) and propeptin inhibits prolyl endopeptidase activity (Kimura *et al.* 1997).

I.5.3. Siderophore-peptides, a marriage between the ribosomal and non-ribosomal worlds

Microcin E492 (MccE492; Figure I.1.) was first characterised as an unmodified 84 amino-acid (7886 Da) peptide produced by *Klebsiella pneumoniae* (de Lorenzo 1984). However, identification of the biosynthetic gene cluster unveiled several additional biosynthetic genes that encode for products with homology to glycosyltransferases, acyltransferases, enterobactin esterases (Lagos *et al.* 2001). Inactivation of these biosynthetic genes resulted in the apparent loss of MccE492 antibacterial activity, but the compound's primary structure appeared to remain unaltered (Corsini *et al.* 2002). Thomas *et al.* (2004) described the isolation and characterisation of the fully modified form of MccE492 (MccE492m; 8717 Da) that had a four- to eight-fold increased antibacterial activity compared to MccE492, explaining previous observations. The post-translational modification was localised at the C-terminus and its structure determined by MS and NMR. It consists of a C-glycosylated linear derivative of the siderophore enterobactin that is attached to the carboxyl of the C-terminal Ser84 by an ester linkage. Because of this post-translational modification, the name siderophore-peptides was proposed for this novel family of antibacterial peptides (Thomas *et al.* 2004). MccE492m targets the iron uptake systems of enterobacteria and its mechanism of action involves pore formation, inner membrane depolarisation and eventually cell death (Destoumieux-Garzon *et al.* 2003; Strahsburger *et al.* 2005; Destoumieux-Garzon *et al.* 2006). The two-protein complex MceIJ was recently shown to be responsible for the post-translational modification of MccE492m by ATP and Mg^{2+} -dependent formation of the glycosyl ester linkage that connects the ribosomal and non-ribosomal peptide moieties (Nolan *et al.* 2007; Nolan and Walsh 2008). Several other microcins are predicted to (or have been proven to) consist of a ribosomal and a non-ribosomal part, including H47, I47 and M (Duquesne *et al.* 2007; Vassiliadis *et al.* 2010).

I.5.4. Microcin C7

Microcin C7 (MccC7; Figure I.1.) is produced by *E. coli* strains that harbour plasmid pMccC7 that carries the genes involved in its production and immunity (Novoa *et al.* 1986; Gonzalez-Pastor *et al.* 1995). MccC7 is a heptapeptide that is post-translationally modified at the C-terminus by a non-hydrolysable phosphoramidate linkage to aminopropyl-modified AMP (Guijarro *et al.* 1995). *mmcA* is the smallest known bacterial gene with a coding sequence of 21 bp and was shown to encode the heptapeptide precursor MccA (Gonzalez-Pastor *et al.* 1994). Modification of the precursor involves three distinct steps, starting with the conversion of Asn7 to *iso*Asn7. The amide group of this residue is subsequently connected to Pα of ATP through a N-P bond and finally the phosphoramidate is aminopropylated. The enzyme MccB is homologous to adenylating enzymes and it was shown to be responsible for the formation of the N-P linkage at the expense of two ATP molecules (Roush *et al.* 2008). The target of MccC7 is aspartyl tRNA synthetase (Metlitskaya *et al.* 2006).

I.6. Modified peptides produced by cyanobacteria

I.6.1. Cyanobactins

Patellamides (Figure I.1.) are a class of cyclic peptides that were identified from the extract of marine ascidians belonging to the Didemnidae family. They consist of two cyclised pseudosymmetrical dimers that each have a 'thiazole–nonpolar amino acid–oxazoline–nonpolar amino acid' sequence. The ascidians harbour obligate symbiotic cyanobacteria and genome sequencing of one of these, *Prochloron didemni*, revealed that the patellamides are produced by the symbiont rather than the host (Schmidt *et al.* 2005). They result from post-translational modification of a genetically encoded precursor. Interestingly, a single 71 amino acid precursor peptide, PatE, contains the prosequence of both patellamide A and C (Schmidt *et al.* 2005). The N-terminal 37 amino acids are proposed to serve as a leader sequence, which is followed by the eight residue patellamide C prosequence, an eight residue spacer, the eight residue patellamide A prosequence and finally a five amino acid C-terminal extension. Functionality of the patellamide gene cluster was demonstrated by heterologous

expression in *E. coli* (Schmidt et al. 2005). A large scale analysis revealed that *Prochloron* species use hypervariable cassettes in a conserved gene cluster to generate a wide diversity of patellamide structures (Donia *et al.* 2006). Each strain contains just one pathway, but the combination of multiple strains within the ascidians results in a diverse chemical library. Several additional patellamide-like gene clusters of cyanobacterial origin were identified and their products were collectively named 'cyanobactins' (Donia *et al.* 2008). This study described the heterologous production of the prenylated antitumor compound trunkamide in *E. coli*. Investigation of the broad-substrate heterocyclases PatD and TruD reveals that sequence variability in the putative substrate-binding domain accounts for the regioselectivity of the introduced post-translational modifications (McIntosh *et al.* 2010).

Aerucyclamides and microcyclamides are cyclic hexapeptides produced by the toxic freshwater cyanobacterium *Microcystis aeruginosa* (Portmann *et al.* 2008; Portmann *et al.* 2008). Their structure contains five-membered heterocycle modifications and aerucyclamide A, B and C each contain one D-*allo*-Ile residue. Identification of the biosynthetic gene clusters revealed genes with homology to the ones involved in patellamide biosynthesis (Ziemert *et al.* 2008).

I.6.2. Microviridins

Microviridin K (Figure I.1.) is an N-acetylated tricyclic peptidase inhibitor produced by *Planktothrix agardhii* (Philmus *et al.* 2008). Its structure contains three cross-links involving the ω-carboxy groups of Glu or Asp residues. Glu13 forms an amide bond with the ε-amino group of Lys9, and Asp10 and Glu12 are connected by an ester bond to the hydroxyl groups of Thr4 and Ser9, respectively. Similar linkages have been identified in peptides of non-ribosomal origin, but identification of the microviridin biosynthetic gene cluster revealed the compound's ribosomal origin (Philmus *et al.* 2008; Ziemert *et al.* 2008). The activity of the enzymes involved in ester and amide bond formation, as well as the acetyltransferase, have been reconstituted *in vitro* (Philmus *et al.* 2008). A subsequent biochemical study of the ligase involved in lactonisation showed that the larger ring is formed first and that the enzyme has some substrate tolerance, which could be useful for variant generation (Philmus *et al.* 2009). A PCR-based screen for

microviridin gene clusters led to the identification of 15 additional variants, illustrating the natural variety of these compounds (Ziemert *et al.* 2010).

I.7. Other five-membered heterocycle-containing peptides

Introduction of heterocycles into ribosomally synthesised peptides appears to be a frequently used strategy in Nature to convert flexible peptide backbones into a more rigid structure (Walsh and Nolan 2008). This type of modification has already been exemplified by MccB17 and the cyanobactins (Chapter I.5.1. and I.6.1.), but at least two other families exist that contain this type of modification, as discussed in this section.

I.7.1. Linear toxins

The biosynthetic gene cluster for production of streptolysin S, a haemolytic toxin produced by *Streptococcus pyogenes*, was identified by transposon mutagenesis and gene disruption analysis (Betschel *et al.* 1998; Datta *et al.* 2005). This showed that streptolysin S is ribosomally synthesised, but its exact structure in not known. The gene products encoded by the streptolysin S gene cluster display remarkable similarities to the enzymes involved in MccB17 maturation and this was predicted to reflect its structural characteristics. *In vitro* reconstitution of the streptolysin S biosynthetic enzymes confirmed this hypothesis (Lee *et al.* 2008). A bioinformatic analysis identified several related gene clusters in prokaryotic genomes, including those of important pathogens. These clusters are also predicted to be involved in the production of thiazole- and/or oxazole-containing toxins (Lee *et al.* 2008). Further experiments identified the sites of post-translational modification in the mature peptide and substrate recognition requirements of the biosynthetic machinery (Mitchell *et al.* 2009). A related toxin, clostridiolysin S, was identified from *Clostridium botulinum* and contains a single modified methyl-oxazole moiety (Gonzalez *et al.* 2010).

I.7.2. Thiopeptides

Thiopeptides, or thiazolylpeptides, are a family encompassing over 50 antibiotics that exert their mechanism of action by interfering with bacterial protein synthesis (reviewed by Bagley *et al.* 2005). They are produced by actinomycetes, bacilli and staphylococci and 29 families have currently been described containing over 75 distinct compounds. Prior to 2009, thiopeptide research focused on total chemical synthesis of the compounds and feeding studies with isotopically labelled precursors. This was probably due to the difficulty in cloning the biosynthetic gene clusters because of the highly modified nature of these compounds and since resistance genes are not typically encoded in a genetic locus close to the biosynthetic machinery.

Although goadsporin (Figure I.1.) is not a thiopeptide, the identification of its gene cluster from *Streptomyces* sp. TP-A0584 represented the first example of this kind (Onaka *et al.* 2005), although this was not realised at the time. The primary function of goadsporin seems to involve stimulation of secondary metabolism and/or sporulation in actinomycetes, although some antibacterial activity was observed at concentrations > 1 µM (Onaka *et al.* 2001). The *god* (goadsporin) gene cluster was cloned and heterologously expressed in *S lividans* TK23 and appears to reflect a merger between the genetic determinants involved in MccB17 and lantibiotic biosynthesis (Onaka *et al.* 2005). GodE is similar to McbC, while GodF and GodE both contain regions homologous to LanB-type dehydratases. GodH shares sequence identity to acetyltransferases and is thought to catalyse the acetylation of the N-terminal alanine. The *godI* product has homology to the signal recognition particle (SRP) from *E. coli* and was shown to be the goadsporin self-resistance protein by conferring goadsporin resistance on *S. lividans* TK23 (Onaka *et al.* 2005).

The biosynthesis of thiopeptide antibiotics was eludicated in 2009 by four different groups simultaneously (Kelly *et al.* 2009; Liao *et al.* 2009; Morris *et al.* 2009; Wieland-Brown *et al.* 2009). These studies show that the antibiotics are genetically encoded post-translationally modified peptides rather than products of NRPS enzymes as previously believed. Important factors that likely played a role in this major finding include advanced genome sequencing techniques and the knowledge that other complex heterocyclic peptides, such as the cyanobactins, are processed from a ribosomal precursor (Arndt *et*

al. 2009). These recent advances will speed up investigations of producer self-immunity and attempts to carry out rational drug design with molecules containing thiopeptide backbones.

Thiopeptides are probably the most intensively post-translationnaly modified type of peptide discovered thus far, as exemplified by thiocillin (Figure I.1.), in which 13 out of 14 residues in the mature peptide are modified. Thiopeptide gene clusters typically contain a structural gene with a propeptide that is rich in Ser, Thr and Cys and at least four biosynthetic genes: a cyclodehydratase homologous to PatD, a McbC-like dehydratase and two LanB-type dehydratases. Several additional enzymes can be found that are expected to be involved in further tailoring of the thiopeptide structure, for example SAM dependent methyltransferases or cytochrome P450 oxygenases. A further investigation of the biosynthetic gene clusters will be needed to assign functions to these additional enzymes and will shed more light on tailoring processes such as the introduction of a second macrocycle and the quinaldic acid moiety in thiostrepton (Li and Kelly 2010).

I.8. Linaridins

Cypemycin (Figure I.1.) is an unusual peptide antibiotic produced by *Streptomyces sp.* OH-4156 (Komiyama et al. 1993). While the mature peptide does not contain any Lan or MeLan bridges, four Dhb residues occur in its structure. Thus, strictly speaking, cypemycin is not a lantibiotic, but regardless it has always been considered a member of this peptide family. The compound was classified based on the producer strain and chemical composition, as being part of the Class II 'cinnamycin group' by Pag and Sahl (2002). Using different criteria, cypemycin was assigned to Class I because of its linear structure and the uncertainty whether its gene cluster contained *lanBC* or *lanM* modification genes (McAuliffe *et al.* 2001; Chatterjee *et al.* 2005).

Apart from the four Dhb residues in its mature structure, cypemycin also contains the following special post-translational modifications: two L-*allo*-isoleucines (L-*allo*-Ile) residues, an AviCys and an N-terminal *N,N*-dimethylalanine (Me_2-Ala) residue (Minami *et al.* 1994). AviCys residues occur in several lantibiotics (Chatterjee *et al.* 2005), but L-*allo*-Ile and Me_2-Ala have not been described in any other post-translationally modified

peptide. The mode of action for cypemycin remains to be determined, but the compound has antibiotic activity against *Micrococcus luteus* and is cytotoxic to mouse leukaemia cells in an *in vitro* assay (Komiyama *et al.* 1993).

Cypemycin is the main subject of this PhD thesis (Chapter III – V) and the identification and characterisation of its biosynthetic gene cluster have revealed that its post-translational modifications are introduced by enzymes and mechanisms unrelated to those described for lantibiotic biosynthesis (Claesen and Bibb 2010). A bioinformatic analysis identified ten similar gene clusters in the genomes of sequenced bacteria, including *S. griseus*. The products of these gene clusters constitute the linaridin family which was named after their predicted linear structure containing dehydrated (or 'arid') residues. The compound from the cryptic linaridin gene cluster of *S. griseus* is described in Chapter VI.

I.9. Modified quorum sensing peptides

I.9.1. Cyclic (thio)lactones

Virulence and biofilm formation in *S. aureus* was found to be under the control of a QS system that uses a peptide pheromone (Ji *et al.* 1995). Structural analysis of the homologous *Staphylococcus epidermis* pheromone (Figure I.1.) revealed the presence of a thiolactone ring between the thiol group of Cys4 and the C-terminal carboxyl group of Phe8 (Otto *et al.* 1998). A typical *agr* QS system consists of a two-component His kinase and response regulator, an auto-inducing peptide (AIP) that is derived from a precursor peptide by post-translational modification and an enzyme involved in the processing of the AIP. At least 4 distinct families of *S. aureus* AIP exist and the peptide from one family can inhibit the QS system of another while stimulating gene regulation in cells belonging to its own family (Ji *et al.* 1997). *agr* pheromones are produced as a precursor peptide, encoded by *agrD*, and subsequently post-translationally processed by the transmembrane protein AgrB (Zhang *et al.* 2002). The N-terminal helix of the AgrD precursor is stabilised by integration into the cytoplasmic membrane and this is required for successful processing to form the mature AIP (Zhang *et al.* 2004). Two catalytic residues were identified in AgrB that form a putative endopeptidase catalytic centre for proteolytic cleavage of the C-terminal fragment of AgrD (Qiu *et al.* 2005).

A peptide QS system was implicated in gelatinase biosynthesis in *Enterococcus faecalis* (Nakayama *et al.* 2001). The AIP is an 11 amino acid cyclic peptide that resembles the *agr* molecules, but instead of a thiolactone, it contains a lactone ring between the hydroxyl group of Ser3 and the C-terminal carboxyl group. The propeptide FsrD is converted to the mature pheromone by Cys protease-like enzyme FsrB in a similar way to the staphylococcal *agr* system (Nakayama *et al.* 2006). The production of cyclic lactone QS pheromones has also been shown for six different strains of *Staphylococcus intermedius* (Ji et al. 2005).

An *agr*-like gene cluster was identified in *Lactobacillus plantarum* and shown to be involved in regulation of adherence to glass slides (Sturme *et al.* 2005). The putative precursor gene and the gene encoding the processing protein were overexpressed under nisin control, which led to the identification of the cyclic thiolactone pentapeptide LamD558 (Figure I.1.). This study provided the first example of a functional *agr*-like system in a non-pathogenic bacterium (Sturme *et al.* 2005).

I.9.2. *Bacillus* pheromones

Genetic competence in *Bacillus subtilis* is under the regulatory control of a ComQXPA QS system (Magnuson *et al.* 1994). ComX (Figure I.1.) is an extracellular peptide pheromone that is post-translationally isoprenylated and secreted by ComQ (Okada *et al.* 2005). Characterisation of ComX pheromones from several natural *Bacillus* isolates showed that all are modified on a conserved Trp and that the modifications on the peptide backbones vary in mass among the different pheromones (Ansaldi *et al.* 2002).

I.9.3. Pep1357C

Pep1357C is a nonapeptide identified in the supernatant of *Streptococcus thermophilus* (Ibrahim et al. 2007). MALDI-ToF MS showed that the peptide is cyclised between Lys2 and Trp6, but further structural analysis would be required to determine the nature of the linkage. A gene deletion analysis indicated that Pep1357C production is the target of a streptococcus-specific QS system (Ibrahim *et al.* 2007).

I.10. Aims of this project

The main goal of this project was to identify and characterise the cypemycin biosynthetic gene cluster from *S. sp.* OH-4156. This was addressed using the following approaches:

- Construction of a genomic cosmid library of the cypemycin producer strain and identification of the cosmids that contained the biosynthetic cluster by radio-active colony hybridisation. The probe used for hybridisation was constructed based on *S. sp.* OH-4156 genome scanning data.

- Heterologous expression of the gene cluster, the generation of a reduced gene set construct and a mutational analysis of the individual genes within this set. The resulting strains were characterised by *Micrococcus luteus* bio-assays and MALDI-ToF MS.

- Setting up *in vitro* assays for the enzymes that introduce cypemycin's post-translational modifications.

In a side project, an attempt was made to obtain products from the cryptic lantipeptide gene clusters of *Streptomyces venezuelae*. The following approaches were taken:

- Screening *S. venezuelae* under various production conditions for inhibitory activity against *M. luteus* and the production of lantipeptide-candidates by MALDI-ToF MS.

- The generation of deletion mutants and constitutively expressing strains in *S. venezuelae,* and the introduction of the gene clusters in *Streptomyces lividans* as a heterologous host.

Chapter II – Materials and methods

II.1. Bacterial plasmids and strains

Table II.1. Plasmids and cosmids used and constructed in this study.

Plasmid	Description	Reference
General plasmids		
pGEM-T	TA-cloning vector	Promega
pIJ10257	*oriT*, ϕBT1 *int-attB*, *hyg*, *ermE*p*	Hong *et al.* 2005
pDONR 207	donor vector for Gateway	Invitrogen
pHM-GWA	Gateway destination vector for MBP fusions	Busso *et al.* 2005
pIJ10700	pBS SK+ containing cassette P1-FRT-*oriT*-*hyg*-FRT-P2	Gust *et al.* 2004
pIJ10702	*bla*, *aac(3)IV*, *oriT*, ϕC31 *int-attP* (also known as pMJCos1)	Boakes *et al.* 2009
pIJ10704	pGEM-T Easy containing cassette FRT-*aac(3)IV*-*oriT*-FRT-*ermE*p*	This work
pIJ773	pBS SK+ containing cassette P1-FRT-*oriT*-*aac(3)IV*-FRT-P2	Gust *et al.* 2004
pIJ790	λ-RED (*gam*, *bet*, *exo*), *cat*, *araC*, *rep101*^{ts}	Gust *et al.* 2004
pIJ86	complementation plasmid containing *ori*ColE1 SCP2*, *aac(3)IV*, *ermE*p*	Healy *et al.* 2009
pSET152	*lacZ*α, pUC19*ori*, RP4*oriT*, ϕC31 *int-attP*, *aac(3)IV*	Flett *et al.* 1997
pUZ8002	*tra*, *neo*, RP4	Paget *et al.* 1999
SuperCosI	*neo*, *bla*	Stratagene
Chapter III and IV		
Library cosmids for heterologous expression in *S. venezuelae*		
pIJ12400	genomic library cosmid 1N16 in SupercosI	This work
pIJ12401	genomic library cosmid 3O03 in SupercosI	This work
pIJ12402	genomic library cosmid 4J11 in SupercosI	This work
pIJ12403	genomic library cosmid 6C03 in SupercosI	This work
pIJ12404	genomic library cosmid 6I23 in SupercosI	This work
pIJ12405	genomic library cosmid 7P24 in SupercosI	This work
pIJ12406	genomic library cosmid 8B19 in SupercosI	This work
pIJ12407	genomic library cosmid 8G13 in SupercosI	This work
pIJ12408	genomic library cosmid 8J18 in SupercosI	This work
pIJ12409	pIJ12400 with pIJ10702 backbone	This work
pIJ12410	pIJ12401 with pIJ10702 backbone	This work
pIJ12411	pIJ12402 with pIJ10702 backbone	This work
pIJ12412	pIJ12403 with pIJ10702 backbone	This work
pIJ12413	pIJ12404 with pIJ10702 backbone	This work
pIJ12414	pIJ12405 with pIJ10702 backbone	This work
pIJ12415	pIJ12406 with pIJ10702 backbone	This work

Plasmid	Description	Reference
pIJ12416	pIJ12407 with pIJ10702 backbone	This work
pIJ12417	pIJ12408 with pIJ10702 backbone	This work
Minimal gene set constructs		
pIJ12418	pIJ12404 targeted with 773 cassette upstream of *orf1*	This work
pIJ12419	pIJ12418 with in-frame deletion upstream of *orf1*	This work
pIJ12420	pIJ12419 targeted with 773 cassette downstream of *cypI*	This work
pIJ12421	pSET152 with minimal gene set excised from pIJ12420	This work
Gene deletion constructs		
pIJ12422	pIJ12404 Δ*orf1*::(*oriT-aac(3)IV*)	This work
pIJ12423	pIJ12404 Δ*cypA*::(*oriT-aac(3)IV*)	This work
pIJ12424	pIJ12404 Δ*cypH*::(*oriT-aac(3)IV*)	This work
pIJ12425	pIJ12404 Δ*cypL*::(*oriT-aac(3)IV*)	This work
pIJ12426	pIJ12404 Δ*cypD*::(*oriT-aac(3)IV*)	This work
pIJ12427	pIJ12404 Δ*cypM*::(*oriT-aac(3)IV*)	This work
pIJ12428	pIJ12404 Δ*cypT*::(*oriT-aac(3)IV*)	This work
pIJ12429	pIJ12404 Δ*cypP*::(*oriT-aac(3)IV*)	This work
pIJ12430	pIJ12404 Δ*cypI*::(*oriT-aac(3)IV*)	This work
pIJ12431	pIJ12422 in-frame deletion of *orf1*	This work
pIJ12432	pIJ12423 in-frame deletion of *cypA*	This work
pIJ12433	pIJ12424 in-frame deletion of *cypH*	This work
pIJ12434	pIJ12425 in-frame deletion of *cypL*	This work
pIJ12435	pIJ12426 in-frame deletion of *cypD*	This work
pIJ12436	pIJ12427 in-frame deletion of *cypM*	This work
pIJ12437	pIJ12428 in-frame deletion of *cypT*	This work
pIJ12438	pIJ12429 in-frame deletion of *cypP*	This work
pIJ12439	pIJ12430 in-frame deletion of *cypI*	This work
pIJ12440	pIJ12431 with pIJ10702 backbone	This work
pIJ12441	pIJ12432 with pIJ10702 backbone	This work
pIJ12442	pIJ12433 with pIJ10702 backbone	This work
pIJ12443	pIJ12434 with pIJ10702 backbone	This work
pIJ12444	pIJ12435 with pIJ10702 backbone	This work
pIJ12445	pIJ12436 with pIJ10702 backbone	This work
pIJ12446	pIJ12437 with pIJ10702 backbone	This work
pIJ12447	pIJ12438 with pIJ10702 backbone	This work
pIJ12448	pIJ12439 with pIJ10702 backbone	This work
Complementation constructs		
pIJ12449	pIJ10257 containing cloned *cypA*	This work
pIJ12450	pIJ10257 containing cloned *cypH*	This work
pIJ12451	pIJ10257 containing cloned *cypL*	This work
pIJ12452	pIJ10257 containing cloned *cypD*	This work
pIJ12453	pIJ10257 containing cloned *cypM*	This work
Variant constructs		
pIJ12477	pIJ86 *neo*	This work
pIJ12478	pIJ12477 *cypA* WT	This work
pIJ12479	pIJ12477 *cypA* 4S	This work

Plasmid	Description	Reference
pIJ12480	pIJ12477 *cypA* S-T	This work
pIJ12481	pIJ12477 *cypA* C-T	This work

Chapter V

Entry vector constructs

pIJ12482	pDONR containing cloned *cypA*	This work
pIJ12483	pDONR containing cloned *pro-cypA*	This work
pIJ12484	pDONR containing cloned *cypH*	This work
pIJ12485	pDONR containing cloned *cypL*	This work
pIJ12486	pDONR containing cloned *cypD*	This work
pIJ12487	pDONR containing cloned *cypM*	This work
pIJ12488	pDONR containing cloned *cypI*	This work

Expression vector constructs

pIJ12489	pHM-GWA containing cloned *cypA*	This work
pIJ12490	pHM-GWA containing cloned *pro-cypA*	This work
pIJ12491	pHM-GWA containing cloned *cypH*	This work
pIJ12492	pHM-GWA containing cloned *cypL*	This work
pIJ12493	pHM-GWA containing cloned *cypD*	This work
pIJ12494	pHM-GWA containing cloned *cypM*	This work
pIJ12495	pHM-GWA containing cloned *cypI*	This work

Chapter VI

Grisemycin-related constructs

pIJ12474	pSET152 with grisemycin minimal gene set	This work
pIJ12475	SuperCosI with grisemycin minimal gene set excised from pIJ12474	This work
pIJ12476	pIJ12475 D*grmA*::(*oriT-aac(3)IV*)	This work

Chapter VII

Constructs for *S. venezuelae* gene deletion and overexpression

pIJ12454	4-G03 ΔSMD01140::(*oriT-aac(3)IV*)	This work
pIJ12455	4-G06 ΔSMD01224::(*oriT-aac(3)IV*)	This work
pIJ12456	4H08 ΔSMD01252::(*oriT-aac(3)IV*)	This work
pIJ12457	4-G03 ΔSMD01140::(*oriT-hyg*)	This work
pIJ12458	4H08 ΔSMD01252::(*oriT-hyg*)	This work
pIJ12459	4P22 Δ*cmlPH*::(*oriT-aac(3)IV*)	This work
pIJ12460	4-G03 *ermE*p*-SMD01140	This work
pIJ12461	4-G06 *ermE*p*-SMD01224	This work
pIJ12462	4H08 *ermE*p*-SMD01252	This work
pIJ12463	4H08 ΔSMD01252.2::(*oriT-aac(3)IV*)	This work

***S. lividans* heterologous expression constructs**

pIJ12464	1-B5 with pIJ10702 backbone	This work
pIJ12465	4-G03 with pIJ10702 backbone	This work
pIJ12466	4H08 with pIJ10702 backbone	This work
pIJ12467	pIJ12464 ΔSMD01139::tet	This work
pIJ12468	pIJ12465 ΔSMD01139::tet	This work

Plasmid	Description	Reference
pIJ12469	pIJ12466 ΔSMD01252.2::tet	This work
pIJ12470	pIJ12464 ΔSMD01140::tet	This work
pIJ12471	pIJ12465 ΔSMD01140::tet	This work
pIJ12472	pIJ12466 ΔSMD01252::tet	This work
ermE* test construct		
pIJ12473	6F11 *ermE*p**-afsR*	This work

Table II.2. Strains used or constructed in this study.

Strain	Description	Reference
General strains		
M. luteus		
ATTCC4698	Indicator strain in bio-assay	Boakes *et al.* 2009
E. coli		
BL21 (DE3)	strain used for fusion protein production	Studier and Moffatt 1986
BT340	DH5α/pCP20	Cherepanov and Wackernagel 1995
BW25113	K-12 derivative: Δ*araBAD*, Δ*rhaBAD*	Datsenko and Wanner 2000
DH5α	strain used for general cloning and plasmid maintenance	Stratagene
ET12567	*dam, dcm, hsdM, hsdS, hsdR, cat, tet*	MacNeil *et al.* 1992
XL1-Blue	strain used for SupercosI library	Stratagene
Chapter III and IV		
S. venezuelae		
ATCC10712	Wild type strain; used to construct the *S. venezuelae* derivatives	Stuttard 1982
M1400	pIJ10702 in φC31 *attB* (empty vector)	This work
M1401	pIJ12409 in φC31 *attB*	This work
M1402	pIJ12410 in φC31 *attB*	This work
M1403	pIJ12411 in φC31 *attB*	This work
M1404	pIJ12412 in φC31 *attB*	This work
M1405	pIJ12413 in φC31 *attB*	This work
M1406	pIJ12414 in φC31 *attB*	This work
M1407	pIJ12415 in φC31 *attB*	This work
M1408	pIJ12416 in φC31 *attB*	This work
M1409	pIJ12417 in φC31 *attB*	This work
S. coelicolor		
M1146	M145 Δ*act* Δ*red* Δ*cpk* Δ*cda*; parental strain used to construct the *S. coelicolor* derivatives	J.P. Gomez-Escribano
M1410	pIJ10702 in φC31 *attB* (empty vector)	This work
M1411	pIJ12413 in φC31 *attB* (full cosmid insert)	This work
M1412	pIJ12421 in φC31 *attB* (minimal gene set)	This work
M1413	pIJ12440 in φC31 *attB* (Δ*orf1*)	This work
M1414	pIJ12441 in φC31 *attB* (Δ*cypA*)	This work
M1415	pIJ12442 in φC31 *attB* (Δ*cypH*)	This work
M1416	pIJ12443 in φC31 *attB* (Δ*cypL*)	This work
M1417	pIJ12444 in φC31 *attB* (Δ*cypD*)	This work

Strain	Description	Reference
M1418	pIJ12445 in ϕC31 *attB* (Δ*cypM*)	This work
M1419	pIJ12446 in ϕC31 *attB* (Δ*cypT*)	This work
M1420	pIJ12447 in ϕC31 *attB* (Δ*cypP*)	This work
M1421	pIJ12448 in ϕC31 *attB* (Δ*cypI*)	This work
M1422	pIJ12441 in ϕC31 *attB* (Δ*cypA*) and pIJ12449 in ϕBT1 *attB*	This work
M1423	pIJ12442 in ϕC31 *attB* (Δ*cypH*) and pIJ12450 in ϕBT1 *attB*	This work
M1424	pIJ12443 in ϕC31 *attB* (Δ*cypL*) and pIJ12451 in ϕBT1 *attB*	This work
M1425	pIJ12444 in ϕC31 *attB* (Δ*cypD*) and pIJ12452 in ϕBT1 *attB*	This work
M1426	pIJ12445 in ϕC31 *attB* (Δ*cypM*) and pIJ12453 in ϕBT1 *attB*	This work
M1459	M1414 (Δ*cypA*) with pIJ12477 (empty vector)	This work
M1460	M1414 (Δ*cypA*) with pIJ12478 (*cypA* WT)	This work
M1461	M1414 (Δ*cypA*) with pIJ12479 (*cypA* 4S)	This work
M1462	M1414 (Δ*cypA*) with pIJ12480 (*cypA* S-T)	This work
M1463	M1414 (Δ*cypA*) with pIJ12481 (*cypA* C-T)	This work
***S. sp.* OH-4156**		
WT	Wild type strain; used to construct the *S. sp.* OH-4156 mutants	Komiyama *et al.* 1993
M1427	Δ*orf1*::(*oriT-aac(3)IV*)	This work
M1428	Δ*cypA*::(*oriT-aac(3)IV*)	This work
M1429	Δ*cypH*::(*oriT-aac(3)IV*)	This work
M1430	Δ*cypL*::(*oriT-aac(3)IV*)	This work
M1431	Δ*cypD*::(*oriT-aac(3)IV*)	This work
M1432	Δ*cypM*::(*oriT-aac(3)IV*)	This work
M1433	Δ*cypT*::(*oriT-aac(3)IV*)	This work
M1434	Δ*cypP*::(*oriT-aac(3)IV*)	This work
M1435	Δ*cypI*::(*oriT-aac(3)IV*)	This work
Chapter VI		
S. coelicolor		
M1457	pIJ12474 in ϕC31 *attB* (grisemycin minimal gene set)	This work
S. griseus		
IFO 13350	Wild type strain; used to construct the *S. griseus* mutant	Shirling and Gottlieb 1972
M1458	Δ*grmA*::(*oriT-aac(3)IV*)	This work
Chapter VII		
S. venezuelae		
M1436	ΔSMD01140::(*oriT-aac(3)IV*)	This work
M1437	ΔSMD01224::(*oriT-aac(3)IV*)	This work
M1438	ΔSMD01252::(*oriT-aac(3)IV*)	This work
M1439	ΔSMD01140::(*oriT-aac(3)IV*) and ΔSMD01252::(*oriT-hyg*)	This work
M1440	ΔSMD01252::(*oriT-aac(3)IV*) and ΔSMD01140::(*oriT-hyg*)	This work
M1441	Δ*cmlPH*::(*oriT-aac(3)IV*)	This work
M1442	*ermE*p*-SMD01140	This work

Strain	Description	Reference
M1443	*ermE*p*-SMD01224	This work
M1444	*ermE*p*-SMD01252	This work
M1445	ΔSMD01252.2::(*oriT-aac(3)IV*)	This work
S. lividans		
TK24	Wild type strain (SLP2$^-$ SLP3$^-$ *str-6*); used to construct the *S. lividans* derivatives	Hopwood *et al.* 1983
M1446	pIJ10702 in ϕC31 *attB* (empty vector)	This work
M1447	pIJ12464 in ϕC31 *attB* (1-B5)	This work
M1448	pIJ12465 in ϕC31 *attB* (4-G03)	This work
M1449	pIJ12466 in ϕC31 *attB* (4H08)	This work
M1450	pIJ12467 in ϕC31 *attB* (1-B5 ΔSMD01139::*tet*)	This work
M1451	pIJ12468 in ϕC31 *attB* (4G03 ΔSMD01139::*tet*)	This work
M1452	pIJ12469 in ϕC31 *attB* (4H08 ΔSMD01252.2::*tet*)	This work
M1453	pIJ12470 in ϕC31 *attB* (1-B5 ΔSMD01140::*tet*)	This work
M1454	pIJ12471 in ϕC31 *attB* (4-G03 ΔSMD01140::*tet*)	This work
M1455	pIJ12472 in ϕC31 *attB* (4H08 ΔSMD01252::*tet*)	This work
S. coelicolor		
M600	Wild type strain (SCP1$^-$ SCP2$^-$); used to test the pIJ10704 construct	Chakraburtty and Bibb 1997
M1456	M600 *ermE*p**afsR*	This work

II.2. Culture media and antibiotics

II.2.1. Antibiotics

Table II.3. Concentration of antibiotics used in this study.

Antibiotic	Concentration in media (mg/ml)
Apramycin (Apra)	50
Carbenicillin (Carb)	100
Chloramphenicol (Cm)	25
Hygromycin (Hyg)	40
Kanamycin (Kan)	50
Nalidixic acid (Nal)	20
Spectinomycin (Spec)*	200
Streptomycin (Strep)*	10
Tetracyclin (Tet)	10

*added in combination

II.2.2. Culture media

Common culture media used for *Streptomyces* and *E. coli* were prepared as previously described (Kieser *et al.* 2000). Cypemycin production medium (MarM) is described in Komiyama *et al.* (1993). Additional growth media for actinomycetes used in screening for cryptic lantipeptides are described in Table II.4. (Flavia Marinelli, personal communication).

Table II.4. Additional fermentation media for actinomycetes

Components	V	D/Seed	A	B	C	INA5	RA3	GPMY	V6	AF/MS	GYM	M8
potato starch (g/l)								20				
peptone (g/l)		5	4				2		5			
soluble starch (g/l)	24	20										20
dextrose (g/l)	1									20		
meat extract (g/l)	3	2	4						5			2
yeast extract (g/l)	5	3	2				4	5	5	2	4	2
malt extract (g/l)							10	5			10	
soy-bean meal (g/l)		2	2			15				8		
glucose (g/l)							10		20		4	10
triptose (g/l)	5											
maltose (g/l)			20									
dextrin (g/l)			10									
$CaCO_3$ (g/l)		1		0.1	0.1	5				4		3
glycerol (ml/l)				20		30	5	20				
glicine (g/l)				2.5	2.5							
hydrolyzed casein (g/l)									3			4
NaCl (g/l)				1	1	2			1.5	1		
KH_2PO_4 (g/l)				1	1							
$FeSO_4$ (g/l)				0.1	0.1							
$MgSO_4$ *$7H_2O$ (g/l)				0.1	0.1							
$MgCl_2$*$6H_20$							2					
tween 60 (ml/l)					20							
pH	7.2	7	7	7	7	7.3	7.4	7	7.1	7.3	7	7

II.3. Growth conditions and genetic manipulations

II.3.1. Growth and storage of *E. coli*

E. coli was grown at 37 °C overnight on L-agar or shaking in LB broth (30 °C for BW25113 carrying pIJ790). For selection of plasmid-containing cells, appropriate antibiotics were added at appropriate concentrations (Table II.3.). For long-term storage, overnight LB cultures of *E. coli* strains were mixed with an equal volume of 40 % glycerol and stored at -80 °C.

II.3.2. Growth and storage of *Streptomyces*

Unless stated otherwise, all of the culturing methods for *Streptomyces* were taken or adapted from those described by Kieser *et al.* (2000).

For liquid growth of *Streptomyces*, ~ 10^8 spores were heat shocked at 50 °C for 10 min in 5 ml TES buffer (10 mM Tris-HCl pH 8.0, 1 mM ethylenediaminetetraacetic acid (EDTA), 1 M NaCl). For *S. venezuelae*, the heat shock step was omitted. An equal volume of 2 x YT medium was added and the mixture incubated at 37 °C for 3-4 h. The emerging germ tubes were microscopically visible at this stage. The germinated spores were recovered by centrifugation at 3000 *g* on a benchtop centrifuge and resuspended in 500 µl of water. The cells were vortexed to disperse the clumps. R5 or TSB liquid medium was inoculated to OD_{450} of 0.03-0.05. Cultures were grown with shaking at 30 °C until late stationary phase. Mycelium was harvested by centrifugation at 3000 *g* in a benchtop centrifuge.

For the generation of spore stocks, *S. lividans*, *S. coelicolor* and *S. sp.* OH-4156 mycelium was streaked out onto SFM solid medium. Mycelium of *S. venezuelae* or *S. griseus* was streaked out onto MYM tap + TE. The plates were incubated at 30 °C for up to 7 days. Spores were harvested according to the protocol described in Kieser *et al*. (2000). Spore stocks were stored in 20 % glycerol at -20 °C or -80 °C.

II.3.3. Plasmid isolation from *E. coli*

Qiagen miniprep kits were used according to the manufacturer's instructions. Briefly, 5 ml of an overnight LB culture harbouring the plasmid of interest was centrifuged at 3000 *g* for 10 min. The cell pellet was then resuspended and underwent alkaline lysis. The lysate was then neutralised and centrifuged in a microcentrifuge at 16000 *g* to remove cell debris and precipitated protein. The supernatant was then applied to a silica membrane mounted in a microcentrifuge tube where it was washed under high salt and ethanolic buffer conditions during which time the DNA remains bound to the column. DNA was eluted from the column in elution buffer. Plasmid DNA was routinely stored at -20 °C.

II.3.4. Cosmid isolation from *E. coli*

Cosmid isolation from *E. coli* was carried out by alkaline lysis as described by Sambrook *et al.* (2001). The cell pellet from 1.5 ml of culture was resuspended by vortexing in 100 µl solution I (50 mM Tris-HCl pH 8.0, 10 mM EDTA). 200 µl solution II (200 mM NaOH, 1 % sodium dodecyl sulphate (SDS)) were added and the tubes inverted ten times. 150 µl solution III (3 M potassium acetate pH 5.5) were then added and mixed in by inverting the tube five times. The tube was then centrifuged at 16000 *g* in a microcentrifuge for 5 min at room temperature. The supernatant was mixed with 400 µl phenol/chloroform, vortexed briefly to mix and then centrifuged at 16000 *g* in a microcentrifuge for 5 min. The upper phase was transferred to a 1.5 ml tube, 600 µl of ice cold isopropanol was added and DNA precipitation was achieved by placing the tube on ice for 10 min followed by centrifuging at 16000 *g* in a microcentrifuge for 5 min. The pellet was washed with 200 µl 70 % ethanol and centrifuged at 16000 *g* in a microcentrifuge. Leaving the tube open for 20 min at room temperature allowed the pellet to dry prior to resuspending in 50 µl 10 mM Tris-HCl (pH 8.0).

II.3.5. Genomic DNA extractions from *Streptomyces*

2 µl of each *Streptomyces* spore stock was used to inoculate 10 ml SOC medium and grown overnight with shaking at 30 °C. The mycelium was recovered by centrifugation at 3000 *g* for 5 min at 4 °C in a Sorvall GS3 rotor. The mycelium was resuspended in 500 µl of lysozyme solution. The mycelium was incubated at 37 °C for 60 min. 50 µl 10 % SDS was then added and incubation continued at 37 °C for a further 15 min. The sample was then vortexed until the viscosity of the solution had decreased. 300 µl phenol/chloroform were added and vortexed briefly until completely mixed. The sample was centrifuged in a microcentrifuge at 16000 *g* for 8 min. Following centrifugation, the aqueous phase was removed to a fresh tube. To decrease viscosity, 200 µl water was added to the supernatant. The supernatant was extracted twice with phenol/chloroform. Precipitation of the DNA was achieved by the addition of 0.1 volumes 3 M sodium acetate and 1 volume of ice-cold isopropanol and leaving for 10 min on ice. The precipitated DNA was then recovered by centrifugation in a microcentrifuge at 16000 *g* for 5 min. The DNA pellet was washed with 70% ethanol before dissolving in 500 µl elution buffer at room temperature for 2 h. DNAse free RNAse was added to a final concentration of 40 µg/ml and the sample incubated at 37 °C for 30 min. The DNA sample then underwent a phenol/chloroform extraction. The DNA was precipitated and washed as before and then dissolved in 100 µl elution buffer. DNA concentration was determined by spectrophotometry using the ND-1000 spectrophotometer (NanoDrop).

II.3.6. Digestion of DNA 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 after consulting the manufacturer's literature. The reaction volume was usually 20 µl for analytical digests and 50 µl for preparative digests. Unless otherwise instructed, digests were typically carried out for 1 h at 37 °C.

II.3.7. Agarose gel electrophoresis

1 % agarose gels were prepared and run in 1% TAE buffer (40 mM Tris, 1.142 % acetic acid, 1 mM EDTA) at 100 V. Hyperladder I (Bioline) was used to provide size markers. Gels were stained in 0.5 µg/ml ethidium bromide in water for 10-15 min.

II.3.8. Extraction of DNA fragments from agarose gels

DNA fragments separated in agarose gels were excised from the gel using a clean scalpel and purified using the Qiaquick gel extraction kit, following the manufacturer's instructions. 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. The column was washed and the DNA fragment was eluted in elution buffer.

II.3.9. Preparation and transformation of electro-competent *E. coli*

A single colony was used to inoculate 10 ml LB and grown overnight with shaking at 37 °C. 100 µl of this preculture were inoculated into 10 ml SOB and grown at 37 °C for 3-4 h with shaking at 200 rpm to an OD_{600} of ~ 0.6. The cells were recovered by centrifugation at 3000 *g* for 5 min at 4 °C in a Sorvall GS3 rotor. After decanting the medium, the pellet was resuspended by gentle mixing in 10 ml ice-cold 10 % glycerol. The cells were centrifuged as before and washed in 10 % glycerol a further two times. After decanting the supernatant from the final wash, cells were resuspended in the remaining ~ 100 µl of 10 % glycerol.

50 µl electro-competent cell suspension were mixed with ~ 100 ng DNA per transformation. Electroporation was carried out in a 0.2 cm ice-cold electroporation cuvette using a GenePulser II (Bio-Rad) set to 200 Ω, 25 µF and 2,5 kV. The expected time constant is 4.5 – 4.9 ms. After electroporation, 1 ml ice cold LB was immediately added to the shocked cells which were incubated with shaking for 1 h at 37 °C. Transformants were selected by spreading the mixture onto LB agar containing the appropriate antibiotic (or on DNA agar for selection with Hyg).

II.3.10. Preparation and transformation of chemically competent *E. coli*

A single colony was used to inoculate 10 ml LB and grown overnight with shaking at 37 °C. 500 µl of this culture were used to inoculate 50 ml LB and grown at 37 °C for 3-4 h with shaking at 200 rpm to an OD_{600} of ~ 0.6. The cells were transferred to a 50 ml polypropylene Falcon tube and stored on ice for 10 min. The cells were recovered by centrifugation at 3000 *g* in a Sorvall GS3 rotor for 10 min at 4 °C. The medium was decanted and the cell pellet was resuspended in 10 ml ice-cold 0.1 M $CaCl_2$. The cells were recovered by centrifugation at 3000 *g* for 10 min at 4 °C. The pellet was then resuspended in 2 ml of ice-cold 0.1 M $CaCl_2$ and stored on ice.

For transformation, 50 µl of the competent cell suspension were transferred to a 1.5 ml microcentrifuge tube. DNA was added (no more then 100 ng in a 10 µl volume) and the suspension was mixed by gently swirling the tube. The tube was stored on ice for 30 min. The tube was then transferred to a water bath preheated to 42 °C and incubated without shaking for 45 s. 950 µl of LB (prewarmed to 37 °C) was added to the cells which were incubated with shaking at 37 °C for 1 h. Transformants were selected by plating the transformation mix on to L agar plates containing the appropriate antibiotic.

II.3.11. Ligation of DNA

Fragments to be ligated were purified from solution using a Qiagen PCR purification kit or from an agarose gel using a Qiagen gel extraction kit. Purifications were carried out according to the manufacturer's instructions. DNA was eluted in elution buffer. A ligation reaction volume of 10 µl was prepared with an insert/vector molar ratio of 3:1. Ligations were carried out overnight at 16 °C using T4 DNA ligase (Promega).

II.3.12. Conjugation of DNA into *Streptomyces*

ET12567/pUZ8002 harbouring an *oriT*-containing vector was inoculated into 10 ml LB containing Cm, Kan and the appropriate antibiotic for selection of the *oriT*-containing vector and grown overnight at 37 °C. 100 µl of overnight culture were

inoculated into 10 ml fresh LB plus antibiotics as above and grown for ~ 4 h at 37 °C to an OD_{600} of 0.4 - 0.6. The cells were washed twice with 10 ml of LB to remove antibiotics that might inhibit the growth of *Streptomyces*, and were resuspended in 1 ml of LB. While washing the *E. coli* cells, 10 µl *Streptomyces* spore stock was added to 500 µl 2 x YT Broth for each conjugation, heat-shocked at 50 °C for 10 min and then allowed to cool slowly. 0.5 ml *E. coli* cell suspension was mixed with 0.5 ml heat-shocked spores and centrifuged briefly in a microcentrifuge. 800 µl of the supernatant was poured off and the pellet was resuspended in 200 µl residual liquid. A stepwise dilution series from 10^{-1} to 10^{-5} was generated. Each dilution was plated out on MS agar + 10 mM $MgCl_2$ (without antibiotics) and incubated at 30 °C for 16-20 h. Plates were overlaid with 1 ml of water containing 0.5 mg Nal and the appropriate plasmid selection for the isolation of exconjugants. Incubation was continued at 30 °C for several days.

II.3.13. Construction of a *S. sp.* OH-4156 cosmid library

High molecular weight genomic DNA was isolated from a stationary phase culture and cleaved partially with *Sau*3AI. The DNA was size-fractionated by pulsed field gel electrophoresis and fragments between 35 and 50 kb excised and gel-purified (Qiaex II gel extraction kit, Qiagen). The DNA was ligated with *Bam*HI/*Xba*I-cleaved SuperCosI cosmid vector and encapsulated in phage particles (using the Gigapack III Gold *in vitro* packaging system (Stratagene)) which were then used to transfect *E. coli* XL1-Blue. 3072 library clones were picked into 384 well microtiter plates and arrayed onto nylon membranes. The filters were baked and bacterial debris rinsed off prior to hybridisation.

II.3.14. Generation of protein fusion constructs

Gateway primer couples (Table II.5.) were used to amplify *cypA*, *cypH*, *cypL*, *cypD*, *cypM* and *cypI*. The resulting PCR products were introduced into the pDONR207 entry vector and shuttled into the pHM-GWA destination vector (following the protocol of the Gateway cloning kit, Invitrogen). The resulting plasmids pIJ12489 - pIJ12495 (Table II.1.) were introduced into *E. coli* BL21 (DE3) by transformation.

II.3.15. Cypemycin bio-assays

Seed cultures were grown in SOC medium for three days and used to inoculate MarM medium and grown for a further three days. Equal volumes of culture medium were taken from comparably grown cultures and extracted with $CHCl_3$ unless otherwise stated. The solvent was evaporated and the resulting pellet dissolved in methanol (at 1/10 of the original $CHCl_3$ volume). 20 µl of these samples were spotted onto 6 mm Whatman paper discs. After all methanol had evaporated, the discs were placed on top of soft nutrient agar containing the indicator strain *M. luteus*. Typically, the agar was inoculated with 1/8th volume of a culture of *M. luteus* grown at 37 °C to an OD_{600} of 0.4-0.5. Bio-assay plates were incubated overnight at 30 °C and zones of inhibition were recorded after 18-24 h.

II.4. Polymerase chain reaction (PCR) methods and DNA sequencing

II.4.1. General PCR

PCR from *Streptomyces* genomic DNA was typically carried out using 10 ng DNA as template (prepared according to Section II.3.5.). PCR from an *E. coli* plasmid or cosmid DNA template was typically carried out using 1 µl of a standard preparation (Section II.3.3. and II.3.4.).

Reaction mixture:

Forward primer (100 pmoles/µl)	0.2 µl	20 pmoles
Reverse primer (100 pmoles/µl)	0.2 µl	20 pmoles
Template DNA	x µl	
Buffer (10x)	5 µl	1 x
dNTPs (40 nmoles/l)	1 µl	50 µM each
Dimethylsulfoxide (DMSO)	2.5µl	5 %
DNA polymerase (2.5 U/µl)	1 µl	2.5 Units
Water up to a total volume of	50 µl	

Cycle conditions (Touchgene Gradient, Techne):

Initial denaturation:	96 °C, 3 min	
Denaturation:	96 °C, 30 s	30 cycles
Primer annealing:	55 °C, 30 s	
Extension:	72 °C, x s	
Final extension:	72 °C, 10 min	

For all PCR reactions, an extension rate of 1 kb per min was assumed for DNA polymerase.

II.4.2. *E. coli* and *Streptomyces* colony PCR

Colony PCR was used for the rapid screening of recombinant plasmids from *E. coli* colonies during cloning or for verifying plasmid transformations in *Streptomyces*. PCR was carried out as described in Section II.4.1., except that template DNA was substituted with *E. coli* cells or *Streptomyces* mycelium dissolved in the DMSO and the initial denaturation cycle was extended for an additional 2 min.

II.4.3. Purification of PCR products

The Qiaquick PCR purification kit (Qiagen) was used to remove enzymes and unincorporated primers and dNTPs from completed PCR reactions. 1/10 of the PCR reaction mixture was submitted to agarose gel electrophoresis and stained for visualisation (Section II.3.7.). The remaining PCR mixture was diluted 5 times in the manufacturer's high salt buffer and applied to a silica gel membrane mounted in a microcentrifuge tube. The PCR products were washed and the DNA fragments eluted in elution buffer.

II.4.4. DNA sequencing

DNA sequencing was used to confirm the correct sequence of PCR-amplified DNA fragments. DNA sequencing was carried out using the BigDye 3.1 dye-terminator reaction mix (Applied Biosystems) with plasmid DNA template according to the manufacturer's instructions.

Cycle conditions (Touchgene Gradient, Techne):

Initial denaturation:	96 °C, 1 min	
Denaturation:	96 °C, 10 s	30 cycles
Primer annealing:	50 °C, 5 s	
Extension:	60 °C, 4 min	

Sequence analysis was carried out by the John Innes Centre Genome Laboratory. Sequence chromatograms were analysed using the chromatogram trace viewer software FinchTV (version 1.4; http://www.softpedia.com).

Table II.5. Primers used in this study.

Primer	Sequence	Site
General primers		
P1	attccggggatccgtcgaccaataggcgtatcacgaggc	
P2	tgtaggctggagctgcttccgccggcttccattc	
bla F	ccctgataaatgcttcaataatattgaaaaaggaag	
bla R2	aatcaatctaaagtatatatgagtaaacttggtctgacag	
UNIV5-AD1	gactcgcgaattccgacagttga	
EAGI-AD2	ggcctcaactgtcg	
M13 seq F	gtaaaacgacggccagt	
M13 seq R	caggaaacagctatgac	
pSET152 R	tcgccattcaggctgc	
pSET152 F	ctcattaggcaccccagg	
T3	aattaaccctcactaaaggg	
T7	gtaatacgactcactatagggc	
pDONR F	tcgcgttaacgctagcatggatctc	
pDONR R	gtaacatcagagattttgagacac	
10257 seq F	acgtccatgcgagtgtcc	
10257 seq R	ccaaacggcattgagcgtc	
86 seq R	tcggtcatggtcggtctc	
Chapter III and IV		
Degenerate primers and probes for *cypA*		
Cyp deg F	gcnacnccngcnacnc	
Cyp deg R	rcanacnawdagngtn	
Cyp probe 1	gcsacsccsgcsacsccsacsgtsgcscagttcgtsatccagggs	
Cyp probe 2	gcsacsccsgcsacsccsacsgtsgcscagttcgtsatccagggstcsacsatctcscts gtstgc	
***cypA* specific probe/test primers**		
cypA T1	gacggtctcttgaaggtctg	
cypA T2	cctccgccacttccatc	
Minimal gene set PCR-targeting primers and test primers		
Up F	agaattcgcggccgcataatacgactcactatagggatcattccggggatccgtcgacc	
Up R	gccccgtaccgctggggcacgggcccgggggcgggtcggtctagatgtaggctggagc tgcttc	*Xba*I
Down F	ggaggggcggcggggccgcgccgtgccgggcgtcgtcgcaatattattccggggatcc gtcgacc	*Ssp*I
Down R	caagaattcgcggccgcaattaaccctcactaaaggatctgtaggctggagctgcttc	
Up T1	tggaatgaacaatggaagtcaac	
Up T2	cggtgaagtacgggc	
Down T1	gagttcccgctgctgag	
Down T2	ccgaaaagtgccacctgac	

Primer	Sequence	Site
Individual gene deletion PCR-targeting primers and test primers		
orf1 F	caccgcgtccgggggtgccgtcatggaacggcccggctaactagtattccggggatccgtcgacc	*Spe*I
orf1 R	gacgggatgggcgggcacgcgcggagggggacgcacatggctagctgtaggctggagctgcttc	*Nhe*I
cypA F	gaccaccgacgaagggttagtgaagtgcgatctgagatgactagtattccggggatccgtcgacc	*Spe*I
cypA R	gtgacgggcctctccggcggcaggctcagtcccgcgtcagctagctgtaggctggagctgcttc	*Nhe*I
cypH F	tctacctgctgccgcgcgagcaggcctcgtacgcgctccactagtattccggggatccgtcgacc	*Spe*I
cypH R	gatcgaccggtgcgccgagccgcgcaccacctcgaacttgctagctgtaggctggagctgcttc	*Nhe*I
cypL F	ggggcgctcgccgcgcacctggtcgcgaccgtgctgcagactagtattccggggatccgtcgacc	*Spe*I
cypL R	cactcggaggtgaggatcggcttcatccgctggtccgcggctagctgtaggctggagctgcttc	*Nhe*I
cypD F	ccatgtcacgggctccatcagcgcggcgctcgtgccgtgactagtattccggggatccgtcgacc	*Spe*I
cypD R	ggcaggttgaaccccacctcggccgtctgccggttgctcgctagctgtaggctggagctgcttc	*Nhe*I
cypM F	gccaacaggatgcggaaggaagggcgttccggtgagtgaactagtattccggggatccgtcgacc	*Spe*I
cypM R	tggcgccctccgcggcgatcgccggcacttcgctcactggctagctgtaggctggagctgcttc	*Nhe*I
cypT F	cgctcgcgggagatgctggtcgttctggagcggcagtgaactagtattccggggatccgtcgacc	*Spe*I
cypT R	ggtcggcctcgaccgcggtgatgctgtccgtggtcatccgctagctgtaggctggagctgcttc	*Nhe*I
cypP F	atggtggacgcctgcgacggcctcatggagctgggatgaactagtattccggggatccgtcgacc	*Spe*I
cypP R	cccaggcggaccgggggccactcccgcggggctccgttagctagctgtaggctggagctgcttc	*Nhe*I
cypI F	caggggtcccgcgcgactccgggacggcagggcaggatgactagtattccggggatccgtcgacc	*Spe*I
cypI R	cggaggggtcgcggggccgtccccgttccggaggcctcagctagctgtaggctggagctgcttc	*Nhe*I
orf1 T1	ggcccgtacttcaccg	
orf1 T2	gtccggggtctgacag	
cypH T1	atccgtgaagaattcgaagacga	
cypH T2	gcagcacggtcgc	
cypL T1	aagttcgaggtggtgcg	
cypL T2	cccgcagccagtgg	
cypD T1	gatgaagccgatcctcacc	
cypD T2	acacccaggcgacgag	
cypM T1	cgagttcgcgccgc	
cypM T2	gttgcttccggtgaggc	
cypT T1	ggtctggaggtgacgg	
cypT T2	gtgtacgcccagtcctg	
cypP T1	cctggtcgccaccc	

Primer	Sequence	Site
cypP T2	catggtgtccatcctgcc	
cypI T1	acgacggtgtcctggag	
cypI T2	ctcagcagcgggaactc	
Primers for complementation constructs		
cypA 10257 F	aaaaacatatgactcttacgagcacgaattccgc	*Nde*I
cypA 10257 R	aaaaaaagctttcagcagaccaggcagatcg	*Hin*dIII
cypH 10257 F	aaaaacatatgctcgcgggagtgacc	*Nde*I
cypH 10257 R	aaaaaaagctttcagtaagcatccatgaaccttc	*Hin*dIII
cypL 10257 F	aaaaacatatgcttactgaccggatcaaggg	*Nde*I
cypL 10257 R	aaaaaaagctttcacgcggcggctcccg	*Hin*dIII
cypD 10257 F	aaaaacatatgaacgtggagaagttcgaggg	*Nde*I
cypD 10257 R	aaaaaaagctttcactcaccggaacgccc	*Hin*dIII
cypM 10257 F	aaaaacatatgagtgacccgagcgtgta	*Nde*I
cypM 10257 R	aaaaaaagctttcactgccgctccagaacga	*Hin*dIII
10257 seq F	acgtccatgcgagtgtcc	
10257 seq R	ccaaacggcattgagcgtc	
Primers for variant constructs		
neo F	tatatgctagccgccctctggtaaggttgg	*Nhe*I
neo R	tatatgctagcgaaccccagagtcccgctc	*Nhe*I
WT 86 F	tatatggatcccgcatgaccaccgacgaagg	*Bam*HI
WT 86 R	tatataagctttcagcagaccaggcagatcgtgc	*Hin*dIII
C-T 86 R	tatataagctttcagcagaccagggtgatcgtgctgc	*Hin*dIII
S-T 86 R	tatataagctttcagcagaccaggcagatcgtggtgccctggatcacg	*Hin*dIII
4S 86 R	tatataagctttcagcagaccaggcagatggagc	*Hin*dIII

Chapter V

Primers for fusion protein constructs

Gat *cypA* F	ggggacaagtttgtacaaaaaagcaggcttagaaaacctgtattttcagggcatgactcttacgagcacgaattccgc
Gat pro*cypA* F	ggggacaagtttgtacaaaaaagcaggcttaatcgaaggtcgtgccaccccggccacgccgac
Gat *cypA* R	ggggaccactttgtacaagaaagctgggtttcagcagaccaggcagatcg
Gat *cypH* F	ggggacaagtttgtacaaaaaagcaggcttagaaaacctgtattttcagggcgtgctcgcgggagtgacc
Gat *cypH* R	ggggaccactttgtacaagaaagctgggtttcagtaagcatccatgaaccttc
Gat *cypL* F	ggggacaagtttgtacaaaaaagcaggcttagaaaacctgtattttcagggcatgcttactgaccggatcaaggg
Gat *cypL* R	ggggaccactttgtacaagaaagctgggtttcacgcggcggctcccg

Primer	Sequence	Site
Gat *cypD* F	ggggacaagtttgtacaaaaaagcaggcttagaaaacctgtattttcagggcgtgaacgt ggagaagttcgaggg	
Gat *cypD* R	ggggaccactttgtacaagaaagctgggtttcactcaccggaacgccc	
Gat *cypM* F	ggggacaagtttgtacaaaaaagcaggcttagaaaacctgtattttcagggcgtgagtga cccgagcgtgta	
Gat *cypM* R	ggggaccactttgtacaagaaagctgggtttcactgccgctccagaacga	
Gat *cypI* F	ggggacaagtttgtacaaaaaagcaggcttagaaaacctgtattttcagggcatggaca ccatggcgaaccg	
Gat *cypI* R	ggggaccactttgtacaagaaagctgggtttcagagcttccggcgcagt	
Gat *cypH* int	cgtcggaacggtggtgtg	
Gat *cypI* int 1	gcatcgtgaaggacctcg	
Gat *cypI* int 2	ggtgaagcggtcgatgac	

Chapter VI

Grisemycin-related primers

grmA F	tttcgctcagcataggttccgaggaaggacagcgaaatgactagtattccggggatccgt cgacc	*Spe*I
grmA R	agggcggacccctccgacacgcgagtcagcccaccatcaactagttgtaggctggagc tgcttc	*Spe*I
grmA T1	aggagcgggccttgctc	
grmA T2	cgccgcgagtgtcacc	
Gris F1 T1	accctgggcgagatacc	
Gris F1 T2	caccggcgtctcatgg	
Gris F2 T1	tcagggtgtccagggtc	
Gris F2 T2	ggtcggccatgtgctg	
Gris F3 T1	acagacgttcgagtaccgtc	
Gris F3 T2	cgaagtcgcgcagcag	

Chapter VII

Primers for *S. venezuelae* gene deletion and test primers

1136 F	gcctccgggcccgccccgccccgcgccccttccggtctaactagtattccggggatccgt cgacc	*Spe*I
1136 R	ccaccgtccgccccaccacgaacgagggagccccgagtgtctagatgtaggctggag ctgcttc	*Xba*I
1137 F	ccgaccgactcccgtcgccatacggaggtttgctcgatgactagtattccggggatccgtc gacc	*Spe*I
1137 R	gggcccggaggccgggcccggccgcctcccgggcggtcatctagatgtaggctggag ctgcttc	*Xba*I
1138 F	ctgccgtgcgcgaatcggttcaagtagcctgcgctcatgactagtattccggggatccgtc gacc	*Spe*I
1138 R	cccagacatcgcgccggagggttccgggccggcccgtcatctagatgtaggctggagct gcttc	*Xba*I
1139 F	caccactcccgcggacacgaaggacgagtgaggaccatgactagtattccggggatcc gtcgacc	*Spe*I
1139 R	tcccggacacgtcacgcgcggcgtgtccgccgcccgtcatctagatgtaggctggagct gcttc	*Xba*I
1140 F	atccgcggggcggagggccgacggggcggggcggcttcaactagtattccggggatc cgtcgacc	*Spe*I

Primer	Sequence	Site
1140 R	tgggccgggtcgggtacggagcggtgcgagggggctatgtctagatgtaggctggagctgcttc	*Xba*I
1141 F	ccgaggccggcgggccggcgggcggcggcccggcctatgactagtattccggggatccgtcgacc	*Spe*I
1141 R	cccccatgtttgtgcgggcgcggggacggccgggggttatctagatgtaggctggagctgcttc	*Xba*I
1224.2 F	cgcggccgtcggctgacggcgtcgggcgcgcgtcccatgactagtattccggggatccgtcgacc	*Spe*I
1224.2 R	ctcccggggatcggtgcgttacatcggctcggggaagattctagatgtaggctggagctgcttc	*Xba*I
1252 F	tccggtcagcggacctcaccgtcgagaaggacacgcatgactagtattccggggatccgtcgacc	*Spe*I
1252 R	ttcgatgtcgtggttctccatggttctctccttcgttcatctagatgtaggctggagctgcttc	*Xba*I
1252.2 F	ggactgtcgccggctcactgaacgaaggagagaaccatgactagtattccggggatccgtcgacc	*Spe*I
1252.2 R	ttcccgggcgggccgccggccggatgcatgggtgcgtgctctagatgtaggctggagctgcttc	*Xba*I
1255 F	gcgcgccgcttcaaccgtcctgactgacgtaccggttcaactagtattccggggatccgtcgacc	*Spe*I
1255 R	atgttgcaacgaaccactccagagacggggaccaaagtgtctagatgtaggctggagctgcttc	*Xba*I
cmlPH F	acgagtagcgcatgggggactacggctccgttcctgtcaactagtattccggggatccgtcgacc	*Spe*I
cmlPH R	gtcgagttcggctggggcacgacggtgacccgatgaccgtctagatgtaggctggagctgcttc	*Xba*I
1136 T1	caccgctcctgaccgccc	
1136 T2	gaaccgcccgcgacggac	
1137 T1	gactgtgcgggtccggca	
1137 T2	acgttcctggcccagcggg	
1138 T1	gatcgacgtgccctacacc	
1138 T2	tgccggacccgcacagtc	
1139 T1	ggcacgggccacctgac	
1139 T2	ggctacttgaaccgattcg	
1140 T1	ccgagtcgctgatctgcatg	
1140 T2	cacaaacatgggggatgttc	
1141 T1	ctgcacgtcaagcacggcg	
1141 T2	gaacatccccccatgtttgtg	
1224.2 T1	ctgtggaagtacgccgtc	
1224.2 T2	gtacccacgcgtcgac	
1252 T1	gttcaagtgaatgcgggcgg	
1252 T2	ctcacacgtcgccgcgaac	
1252.2 T1	cacgcagacccccggcactg	
1252.2 T2	ggtcctccctgcggttcc	
1255 T1	gtgggcaggcggagac	
1255 T2	ctcgccgtgccattcg	
cmlPH T1	gaccagccgtaccagcc	
cmlPH T2	ccacaccctcctcaccg	

Primer	Sequence	Site
Primers for *S. coelicolor and S. venezuelae* constitutive gene expression		
afsR ermE F	gatcggcgtctcccacggctgacgtggtcggcatgaacaaatttaaattgtaggctggag ctgc	
afsR ermE R	cgggacgccgctgctccggaacccgcggtccaccgtccatatggggcctcctgttcta	
1140 *ermE* F	aggcgccccccgtaccaggatcggacgtggcccggagcggatttaaattgtaggctgg agctga	
1140 *ermE* R	gcgcggccccaccccgcccgtgtccggcctccaccgccatatggggcctcctgttcta	
1224 *ermE* F	cgtaccggcggctgaaaccgatcgcctgccgaccgcatgatttaaattgtaggctggagc tgc	
1224 *ermE* R	cacgggatctcctcgcgggatggggacgggcggccacatatggggcctcctgttcta	
1252 *ermE* F	gtgtgttggttccgtgatgcggtcgggaaccttccggtcaatttaaattgtaggctggagctg c	
1252 *ermE* R	gcagtccttccagctcgacttccgtaacccggctcgtcatatggggcctcctgttcta	
afsR ermE T1	ggggcaaagccgaagaag	
afsR ermE T2	cgattcctcctcttccgc	
1140 *ermE* T2	gtcccaccaccgctc	
1224 *ermE* T1	tcttcgtatcgggctctgtg	
1224 *ermE* T2	ggacacgtggatcttccag	
1252 *ermE* T2	gcttcgaccgcccagc	

II.5. DNA hybridisation methods

II.5.1. Non-radioactive Southern hybridisation

4 µg of *Streptomyces* genomic DNA were digested with the appropriate restriction enzyme at 37 °C overnight. Extraction with phenol/chloroform was carried out to remove the enzyme. The digested DNA was precipitated by the addition of 0.1 volumes 3 M sodium acetate and 1 volume of ice-cold isopropanol and leaving for 10 min on ice. The precipitated DNA was recovered by centrifugation in a microcentrifuge at 16000 *g* for 5 min. The DNA pellet was washed with 70 % ethanol before dissolving in 15 µl water and loaded on to one lane of a 1 % TAE gel. The gel was typically 20 cm in length and run at 30 V for 16-18 h. The inclusion of 5 ng of 1 kb DNA ladder (Invitrogen) was used to determine the size of the bands on the developed Southern blot. The gel was stained with 0.5 µg/ml ethidium bromide both prior to blotting and after capillary transfer to confirm efficient transfer to the membrane. The gel was rinsed in distilled water and soaked twice with shaking for 15 min in denaturation buffer (500 mM NaOH, 1.5 M NaCl) at room temperature. The gel was subsequently washed in distilled water and soaked twice with shaking for 15 min in neutralisation buffer (3 M NaCl, 500 mM Tris-HCl pH 7.5) at room temperature. DNA was transferred to a Hybond-N nylon membrane (Amersham) using capillary transfer. 20 x SSC (3 M NaCl, 300 mM trisodium citrate dihydrate) was poured into a plastic tray and a glass plate was placed across it. A sheet of 3MM paper (Whatman) was soaked in 20 x SSC and placed on the plate so that the ends of the paper were in contact with the buffer in the tray. The treated and neutralised agarose gel was placed on the paper pad. A piece of nylon membrane cut to the size of the gel was placed on top of the gel, followed by three pieces of 3MM paper of the same size, stacks of paper towels, a glass plate and finally a weight. The DNA was allowed to transfer overnight by capillary action. After transfer, the filter was removed and the transferred DNA was permanently fixed to the membrane by UV-cross-linking in a UV Stratalinker 2400 (Stratagene). The membranes were stored between sheets of 3MM paper inside a plastic bag at 4 °C until further use.

For generation of the digoxigenin (DIG)-labelled DNA probes, DNA was labelled with DIG-11-dUTP using the random primed DNA labelling method. 1 µg of cosmid DNA template (or 1 kb ladder for the ladder probe) was used in a labelling reaction. 1 g DNA template was diluted in H_2O to a total volume of 16 µl. The DNA template was heat-denatured in a boiling water bath for 10 min, and quickly chilled on ice. 2 µl of hexanucleotide mixture (10 x) and 2 µl dNTP labelling mixture (10 x) were added to the tube. 1 µl Klenow enzyme was added for a final concentration of 100 U/ml and mixed. The reaction mixture was incubated at 37 °C for 20 h. The probe was purified using a Qiagen PCR purification kit according to the manufacturer's instructions to remove unincorporated label and enzyme. The probe was eluted from the Qiagen column in 50 µl of elution buffer.

For hybridisation of the DNA probes with membrane-bound DNA, the membrane was placed in a hybridisation tube containing 20 ml prehybridisation solution (5 x SSC, 0.1 % *N*-lauroylsarcosine, 0.2 % SDS, 1 % blocking reagent) per 100 cm^2 of membrane surface area. Prehybridisation was carried out at 67 °C for at least 2 h. The DIG-labelled DNA probe was denatured by heating in a boiling water bath for 10 min and chilled directly on ice. After discarding the prehybridisation solution, the hybridisation solution (= prehybridisation solution containing the DIG-labelled probe at 25 ng/ml) was added. The hybridisation was carried out overnight at 67 °C. The hybridisation solution was discarded and the membrane was washed twice, 15 min per wash, in 50 ml of preheated wash solution I (0.5 x SSC, 0.1 % SDS) at 67 °C. This was followed by three washes, 20 min per wash with 50 ml of preheated wash solution II (0.1 x SSC, 0.1 % SDS) at 67 °C.

For detection of the membrane-bound DIG-labelled probe, the membrane was transferred to a freshly washed dish and equilibrated in maleic acid buffer (0.1 M maleic acid, 0.15 M NaCl, NaOH to pH 7.5) for 1 min at room temperature. The membrane was blocked by gently agitating it in 30 ml of blocking solution (1 % (w/v) blocking reagent in maleic acid buffer) for 60 min at room temperature. The blocking solution was discarded and replaced with 30 ml of blocking solution containing the anti-DIG antibody-conjugate diluted 1:10,000. The membrane was incubated for 30 min with gentle agitation. The antibody solution was discarded and the membrane was washed twice, 15 min per wash in maleic acid buffer and

then equilibrated for 2 min in 30 ml of detection buffer (100 mM Tris-HCl pH 9.5, 100 mM NaCl). Diluting CSPD* (Roche) 1:100 in detection buffer generated the working chemiluminescent substrate. The membrane was placed between two sheets of plastic and 0.5-1 ml of working chemiluminescent substrate was added to the membrane (per 100 cm^2 of surface area) ensuring it was dispersed evenly across the surface of the membrane. The membrane was incubated for 5 min at room temperature. The membrane was then sealed in a fresh plastic bag and incubated at 37 °C for 15 min to allow the alkaline phosphatase chemiluminescent reaction to reach a steady state. For detection of the chemiluminescent signal, the membrane was exposed to X-ray film for 30 min, adjusting the exposure time to optimise the signal.

II.5.2. Radioactive Southern hybridisation

A 480 bp PCR product corresponding to *cypA* and flanking regions was generated using primers *cypA* T1 and *cypA* T2 and genomic DNA of *S. sp.* OH-4156 as template. The fragment was labelled with α-^{32}P-dCTP (following the protocol of the Prime-It Random Primer Labeling Kit, Stratagene). Library membranes (Section II.3.13.) were hybridized with this probe under the conditions are described in Section II.5.1., identifying cosmids containing *cypA.*

II.6. PCR targeting

PCR targeting (also known as the Redirect technology) was used to construct mutagenised cosmids in *E. coli* that could subsequently be used for generating mutants or heterologous expression strains after conjugation into *Streptomyces.* The method was performed according to Gust *et al.* (2004).

II.6.1. PCR amplification of disruption cassette

The pIJ773 ($Apra^R$) and pIJ10700 (Hyg^R) templates were used for the generation of deletion mutants. pIJ10704 was used as template for the replacement of native promoters with the constitutively transcribed *ermE*p* promoter. The 5.2-kb *Ssp*I fragment of pIJ10702 containing *oriT*, ϕC31 *int* and *attP* was used to target the

backbone of cosmids that needed to be integrated into the host genome for heterologous expression. Gene specific primers for the generation of each deletion mutant were designed in accordance with the criteria specified by (Gust *et al.* 2004). These primer pairs were used in association with the appropriate disruption template to generate cassettes flanked by the gene-specific 39 nt homology extensions and are listed in Table II.5.

All PCR amplifications were performed using the Expand high fidelity PCR system (Roche) according to the manufacturer's instructions.

Reaction mixture:

Primers (100 pmoles/µl)	0.5 µl each	50 pmoles each
Template DNA (100 ng/µl)	0.5 µl	50 ng
Buffer (10x)	5 µl	1 x
dNTPs (10 mM)	1 µl each	50 µM each
DMSO	2.5 µl	5%
DNA polymerase (2.5 U/µl)	1 µl	2.5 Units
Water	36 µl	

Cycle conditions (Touchgene Gradient, Techne):

Initial denaturation:	94 °C, 2 min	
Denaturation:	94 °C, 45 s	10 cycles
Primer annealing:	50 °C, 45 s	
Extension:	72 °C, 90 s	
Denaturation:	94 °C, 45 s	15 cycles
Primer annealing:	55 °C, 45 s	
Extension:	72 °C, 90 s	
Final extension:	72 °C, 5 min	

5 µl of the PCR product was used for analysis by agarose gel electrophoresis (Section II.3.7.). To remove enzymes and unincorporated primers and dNTPs, the remaining 45 µl of the PCR product were treated as in Section II.4.3. The PCR product was eluted from the column with 20 µl of elution buffer to yield a DNA concentration of approximately 100 ng/µl.

II.6.2. PCR targeting of a cosmid

Streptomyces Supercos I-based cosmids were introduced into *E. coli* BW25113/pIJ790 by electroporation (Section II.3.9.). *E. coli* BW25113 is a λ-RED recombination-proficient host strain. pIJ790 carries the antibiotic resistance marker *cat* (Cm^R) and a temperature sensitive origin of replication (i.e. it replicates at 30 °C, but not at 37 °C). Transformants from the electroporation were selected by spreading onto L agar containing Carb, Kan and Cm and incubated overnight at 30 °C.

10 ml SOB (without $MgSO_4$) containing Carb, Kan and Cm was inoculated at a concentration of 1 % with an overnight culture of *E. coli* BW25113/pIJ790 containing the cosmid of interest. 100 µl 1 M L-arabinose was added to a final concentration of 10 mM to induce the λ-RED recombination system. The culture was grown for 3-4 h at 30 °C with shaking at 200 rpm to an OD_{600} of ~ 0.6 and electrocompetent cells were prepared. 50 µl cell suspension were then mixed with 1 µl (~ 100 ng) of PCR product and electroporated (Section II.3.9.). Selection for targeted cosmids was carried out on L agar containing Carb, Kan and Apra or Hyg (on DNA medium) overnight at 37 °C. Correct targeting of cosmids was confirmed by PCR and restriction analysis.

II.6.3. Transfer of mutant cosmids into *Streptomyces*

For *Streptomyces* species that carry a methyl-specific restriction system, it was necessary to passage the targeted cosmids through the non-methylating host *E. coli* ET12567, which contains the RP4 derivative pUZ8002. The cosmid was subsequently transferred to *Streptomyces* by conjugation (Section II.3.12.) and exconjugants were selected using the appropriate antibiotics. Single colonies were replica-plated onto DNA medium containing Nal and Apra or Hyg with and without Kan, which allowed for the identification of double crossover exconjugants (Kan^S and $Apra^R$). These were picked from the DNA medium and streaked for single colonies on SFM containing Nal and Apra or Hyg. The Kan sensitivity was reconfirmed by replica plating onto DNA medium containing Nal with and without

Kan. Purified Kan^S strains were verified by PCR and for replacement mutants as well by Southern hybridisation analysis (Section II.5.1.).

II.7. Protein Methods

II.7.1. Purification of His-tagged proteins

Recombinant 6His-MBP fusion proteins were purified from 3 liters of L medium after induction with 0.5 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) at 30 °C for 4 h. Cell pellets were resuspended in 40 ml lysis buffer (50 mM Tris HCl pH 7.9, 10 % glycerol, 1 mM phenylmethanesulfonyl fluoride), lysed with a French press, and the cell lysate clarified by centrifugation (20 min at 40,000 *g*). The cleared lysate was applied to a 1 ml Ni^{2+}-loaded Hi-Trap Chelating HD column (GE Healthcare), washed with 25 mM K_2HPO_4 pH 7.9, 200 mM NaCl, 20 mM imidazole and eluted with an imidazole gradient. The eluted proteins were dialysed into storage buffer (50 mM Tris HCl pH 7.9, 50 mM NaCl, 1 mM tris(2-carboxyethyl)phosphine (TCEP)), aliquoted and stored at -80 ºC.

II.7.2. Protein quantification

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. Unknown samples were diluted to fit within the range of standards. The DC protein assay (Bio-Rad) was used according to the manufacturer's instructions, with an equivalent amount of protein buffer in each BSA standard, sample and blank. The standard curve generated from the standards was used to calculate the protein concentrations in the experimental samples.

II.7.3. SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

Polyacrylamide gels were assembled in their casting stands and prepared using the Bio-Rad system. Resolving gels of between 10-15 % acrylamide were prepared, containing 0.325 M Tris-HCl pH 8.8, 0.1 % (w/v) SDS, 0.01 % (v/v)

N,N,N',N'-tetramethylethylenediamine (TEMED), 0.1 % (v/v) ammonium persulphate. The mixture was poured into the plates to 2 cm below the top and covered with 0.1 % SDS. After polymerisation, the 0.1 % SDS was poured off and a stacking gel, comprising 3 % acrylamide, 0.125 M Tris-HCl pH 6.8, 0.1 % (w/v) SDS, 0.01 % (v/v) TEMED, 0.1 % (w/v) ammonium persulphate, was poured on top and a comb inserted. After polymerisation, gels were removed from their casting holders, fitted to electrodes and placed into the tank supplied which was filled with 1 x SDS running buffer (25 mM Tris, 200 mM glycine, 0.1 % SDS). Samples were prepared for loading by dilution in SDS loading buffer (62.5 mM Tris-HCl, 10 % glycerol, 2 % SDS, 0.01 mg/ml bromophenol blue, 5 % β-mercaptoethanol) and boiled for 10 min after which they were placed immediately on ice. Gels were typically run at 150 V until the bromophenol blue had passed through the resolving gel. To visualise protein bands, the gel was removed from between the glass plates, rinsed in de-ionised water and immersed in InstantBlue protein gel stain (Expedeon). Gels were incubated with gentle shaking at room temperature to allow protein bands to appear.

II.8. Mass spectrometry

II.8.1. MALDI-ToF and Q-ToF MS

$CHCl_3$ was evaporated from 1 ml of production culture extract and the resulting pellet dissolved in 50 µl 5% formic acid. Samples (ca. 0.8 µl) were spotted onto a PAC plate (Prespotted AnchorChip MALDI target plate, Bruker Daltonics, Bremen, Germany) and washed briefly with 8 µl 5% formic acid. After drying, the samples were analyzed by Matrix-Assisted Laser Desorption Ionization-Time-of-Flight MS on a Bruker Ultraflex TOF/TOF. The instrument was calibrated using pre-spotted standards (ca. 200 laser shots). Samples were analyzed using a laser power of approx. 25% and spectra were summed from ca. 20 x 20 laser shots. For Q-ToF MS analysis, the peptide was directly infused into a QToF II (Waters, Manchester, UK), and analysed with MassLynx 4.1. The sample was diluted into 30% methanol/30% Acetonitrile/1% acetic acid and applied with a GlassTip (New Objective, Woburn, MA 01801, USA) by nano-electrospray. Full MS scan analysis

was performed with standard settings, and fragmentation was achieved by increasing the collision energy up to 40.

II.8.2. Thiol alkylation

Alkylation of the free thiol group in the cypemycin intermediate produced by M1517 (Δ*cypD*) was adapted from Sechi and Chait (1998). Briefly, culture medium from a three day old production culture was incubated at 60 °C for 30 min under reducing conditions in the presence of 5 mM *tris*(2-carboxyethyl)phosphine (TCEP). The mixture was allowed to cool to 25 °C. Iodoacetamide was added to a final concentration of 10 mM and the mixture incubated in the dark at 25 °C for 30 min. $CHCl_3$ extraction was performed as described before and the sample analyzed by MALDI-ToF and Q-ToF MS.

Chapter III – Identification of the cypemycin gene cluster

III.1. Introduction

Cypemycin is a linear peptide antibiotic that is produced by *Streptomyces sp.* OH-4156 (Komiyama *et al.* 1993)(Chapter I.8.). It displays potent cytocidal activity (half maximum inhibitory concentration (IC_{50}) = 1.3 µg/ml) against P388 mouse leukemia cells in an *in vitro* assay and it inhibits growth of *Micrococcus luteus* (minimal inhibitory concentration (MIC) = 0.2 µg/ml) (Komiyama *et al.* 1993). The structure of cypemycin was determined by fast atom bombardment mass spectrometry (FAB-MS), nuclear magnetic resonance (NMR) and amino acid analysis (Minami *et al.* 1994). This revealed the occurrence of four different non-proteogenic amino acids, two of which are unique to cypemycin. The first unique residue is a dimethylalanine (Me_2-Ala), located at the N-terminus. Cypemycin also contains two L-allo-isoleucines (L-*allo*-Ile), the second modification that is unique to cypemycin (reviewed in McIntosh *et al.* 2009). The two remaining modifications are also found in lantibiotics. There are four dehydrobutyrine (Dhb), residues that result from the dehydration of Thr in the propeptides of lantibiotics (Chatterjee *et al.* 2005), and at the C-terminus, a *S*-[(Z)-2-aminovinyl]-D-cysteine (AviCys). This latter modification also occurs in the lantibiotics epidermin (Allgaier *et al.* 1986), gallidermin (Kellner *et al.* 1988), mutacin 1140 (Smith *et al.* 2000) and microbisporicin (Castiglione *et al.* 2008; Foulston & Bibb, 2010). Mersacidin contains the related *S*-[(Z)-2-aminovinyl]-3-methyl-D-cysteine (MeAviCys), which is in essence a methylated version of AviCys (Schneider *et al.* 2000). At the start of this thesis, the information available on cypemycin was restricted to its structural properties, its producer organism *S. sp.* OJ-4156 and the fermentation conditions for production. A genetic analysis of cypemycin biosynthesis was therefore performed to obtain more insight into the production of this unusual peptide.

This chapter deals with the identification of the cypemycin gene cluster. First, some general experiments will be discussed that characterise the strain and its compound. A cosmid library was generated from genomic DNA of *S. sp.* OH-4156 as a first step towards the identification of the cluster. Next, experiments based on

the predicted degenerate nucleotide sequence of the precursor gene *cypA* will be described that were aimed at identifying cosmids from the genomic library containing the cluster. These attempts failed and a genome scanning approach was adopted that revealed a striking similarity between the genomes of *S. sp.* OH-4156 and *S. venezuelae*. The data acquired from genome scanning provided us with sufficient DNA sequence to generate a specific probe for the cypemycin structural gene. This probe was used in a Southern hybridisation experiment to identify cosmids containing *cypA* and subsequent heterologous expression identified those that contained all necessary genes for cypemycin production. A candidate cosmid was sequenced and this revealed that the cypemycin biosynthetic gene cluster encodes novel enzymes and that cypemycin is not, in fact, a lantibiotic.

III.2. Characterisation of cypemycin and its producer strain

III.2.1. Antibiotic sensitivities of *S. sp.* OH-4156

The taxonomy of and culture conditions for *S. sp.* OH-4156 were described in Komiyama *et al.* (1993). Since no information was available on antibiotic resistance in this strain, the minimal inhibitory concentration (MIC) was determined for apramycin (Apra), kanamycin (Kan), hygromycin B (Hyg), thiostrepton (Thio) and spectinomycin/streptomycin (Spec/Strep) on a range of different media. The results of the MIC determination are summarised in Table III.1. Apra and Kan appeared to be good candidates for selection at their standardly used (for streptomycetes) concentration of 50 µg/ml. Thio and Spec/Strep also appeared suitable for selection at concentrations of 50 and 400/10 µg/ml, respectively. *S. sp.* OH-4156 was resistant to high levels of Hyg (up to 35 µg/ml) and use of this antibiotic, which is commonly added at a final concentration of 40 µg/ml, was avoided. The most commonly used resistance markers in PCR targeting mutagenesis are Apra and Kan, and cassettes with resistance markers for Hyg and Spec/Strep are also available. The antibiotic sensitivity profile of *S. sp.* OH-4156 suggested that genetic manipulation by PCR targeting might be possible, provided that the resistance markers for these antibiotics were functional in this strain.

Table III.1. Minimal inhibitory concentration (MIC) values for *S. sp.* OH-4156 when grown on solid medium. Values shown in the table are the minimal concentration (in µg/ml) for which no growth of *S. sp.* OH-4156 was observed. Spec/Strep are commonly added in combination for selection. N/D = Not determined.

	SFM	MYM tap	DNA
Kan	35	5	5
Apra	20	5	5
Thio	20	5	N/D
Spec/Strep	100 / 2.5	200 / 5	N/D
Hyg	N/D	N/D	35

III.2.2. Cypemycin detection by bio-assay and MALDI-ToF MS

S. sp. OH-4156 was grown on common agar media for culturing *Streptomyces* species (Kieser *et al.* 2000) to determine suitable conditions for cypemycin production. Cypemycin was known to be produced in liquid culture (Komiyama *et al.* 1993) and an adapted version of the medium described in that study (called MarM in this thesis) was also used in solid agar form to grow *S. sp.* OH-4156. Production of antimicrobial activity was assayed after three and five days by overlaying the plates with soft nutrient agar (SNA) containing *M. luteus* as an indicator strain (Chapter II.3.15.). The antimicrobial activity of *S. lividans* TK24 was assessed under the same conditions to identify media that would be suitable for screening for heterologous cypemycin production. Neither *S. sp.* OH-4156 nor *S. lividans* developed morphologically on solid MarM medium, even after prolonged incubation for up to two weeks. Moreover, the cypemycin producer spread out across the medium surface far beyond the location of the initial inoculation, making overlay assays very challenging. However, a zone of inhibition was observed around *S. sp.* OH-4156 but not *S. lividans* on MarM (data not shown). Similar observations were made for R5 and SFM, indicating that these three media could be used for heterologous expression in *S. lividans*.

Because of difficulties with the overlay assay using the cypemycin production medium MarM, bio-assays from liquid grown cultures were also considered. *S. sp.* OH-4156 was grown for three days in a SOC seed culture which was used to inoculate MarM production medium. After another three days of growth,

comparably grown cultures were extracted with $CHCl_3$. This method proved best for cypemycin detection by bio-assay and Matrix-Assisted Laser Desorption/Ionization Time-of-Flight mass spectrometry (MALDI-ToF MS) and was used for all further experiments described in this thesis. Growth of *S. lividans* under similar conditions did not result in the production of antibacterial activity in a *M. luteus* bio-assay.

Several streptomycetes are known to produce multiple compounds with antibiotic activity and *M. luteus* is an indicator organism that is sensitive to many antibiotics (Young *et al.* 2010). To assess whether the observed antibacterial activity against *M. luteus* could reflect, at least in part, cypemycin production, the extract from the liquid MarM grown culture of *S. sp.* OH-4156 was analysed by MALDI-ToF MS. A set of three peaks ($[M+H]^+$ = 2096 Da, $[M+Na]^+$ = 2118 Da and $[M+K]^+$ = 2134 Da) was observed in the spectrum, which is in good agreement with the calculated mono-isotopic mass of 2095.13 Da for cypemycin (Figure III.1.). Note that even though Minami *et al.* (1994) reported the correct molecular formula for cypemycin as $C_{99}H_{154}N_{24}O_{24}S$, an $[M+H]^+$ of 2095 Da is reported, which in its turn is inconsistent with the $[M+H]^+$ = 2097 Da in the preceding paper by the same group (Komiyama *et al.* 1993).

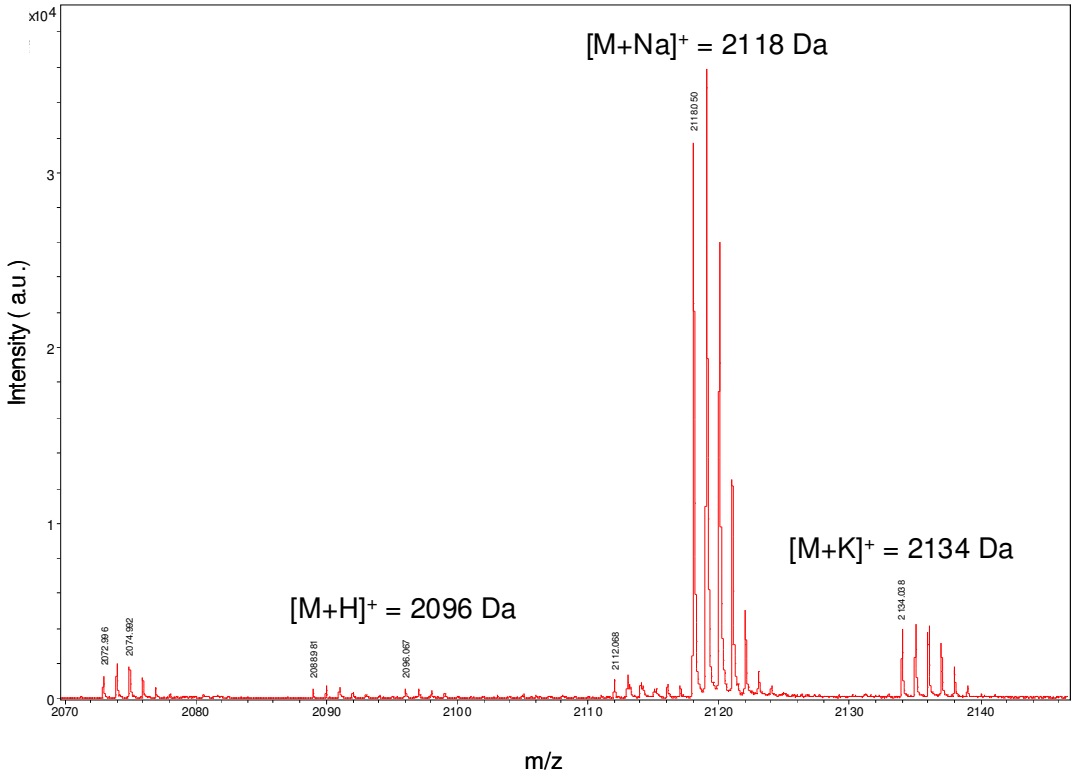


Figure III.1. MALDI-ToF MS spectrum of cypemycin ($[M+H]^+$ = 2096 Da, $[M+Na]^+$ = 2118 Da and $[M+K]^+$ = 2134 Da).

III.2.3. Structure verification by Q-ToF

Quadrupole Time-of-Flight (Q-ToF) MS was also employed to analyze the structure of cypemycin (Figure III.2.). This allowed the determination of the amino acid sequence between Pro3 and the C-terminal AviCys residue (the latter did not fragment into readily identifiable sub-fragments). Cypemycin is especially suitable for this approach as it does not contain any Me(Lan) or other bridge structures, apart from the C-terminal AviCys. Q-TOF MS analysis was also used subsequently to study the location and nature of the modifications in cypemycin and its intermediates without the need for pure compound or large quantities.

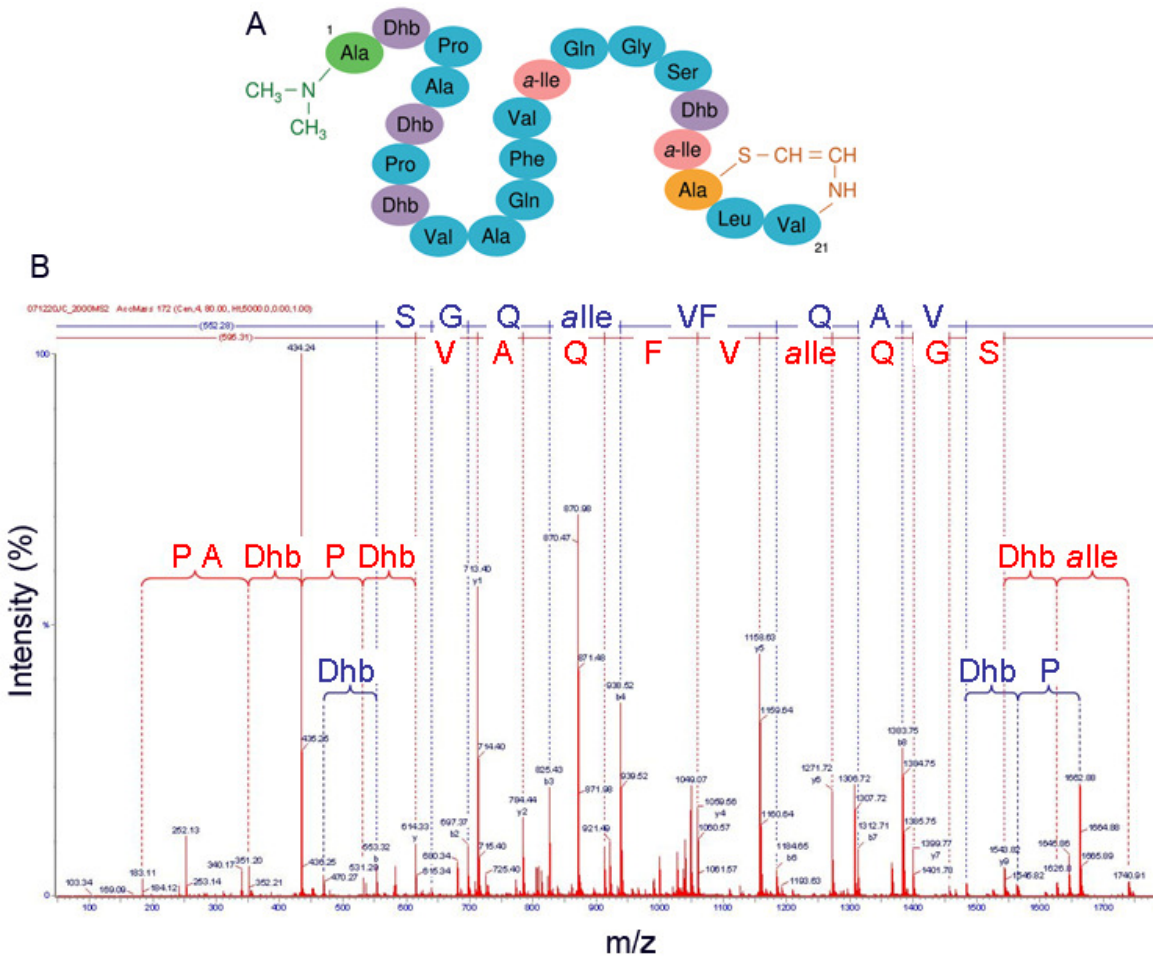


Figure III.2. Q-ToF spectrum of cypemycin. Fragmentation of the modified linear peptide allows determination of the amino acid sequence, and the nature and locations of modified residues. The amino acid sequence between Pro3 and the AviCys residue (which does not fragment into easily interpretable masses) can be readily discerned and is shown above the spectrum.

III.3. Generation and screening of a genomic cosmid library

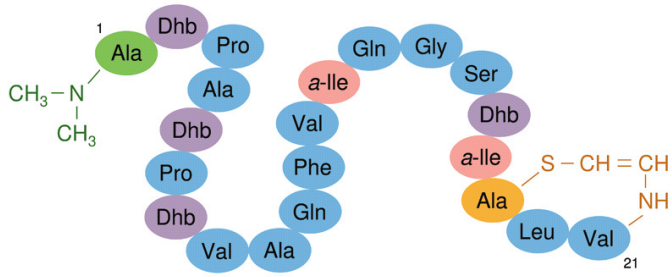
To clone the cypemycin biosynthetic gene cluster, a cosmid library was generated from genomic DNA of *S. sp.* OH-4156 in SuperCosI (Chapter II.3.13). Various methods were employed to either identify the cosmids from this library that harboured the cypemycin biosynthetic gene cluster or to obtain a specific sequence for the region containing the structural gene *cypA*.

III.3.1. PCR based screening

All amino acids in cypemycin are in the L-configuration (Minami *et al.* 1994), presumably reflecting the post-translational modification of a ribosomal precursor peptide. A propeptide sequence was predicted based on the mature peptide and assuming that the Dhb residues result from Thr dehydration and that AviCys is formed from a Ser at position 19 and a C-terminal Cys, as observed in lantibiotics (Chatterjee *et al.* 2005). The other two modifications in cypemycin have not been described previously, so the two L-*allo*-Ile residues were assumed to result from modification of Ile, and the N-terminal Me_2-Ala from methylation of an Ala in the prosequence (Figure III.3.). This predicted pro-CypA sequence was reverse translated to give a degenerate nucleotide sequence that was used to design 16 nt primers Cyp deg F and Cyp deg R. These primers were used in a gradient PCR with *S. sp.* OH-4156 genomic DNA as a template to amplify a predicted 66 bp DNA fragment. Such a fragment would allow for the identification of a partial non-degenerate pro-*cypA* sequence, which could then be used to generate a specific probe for screening the cosmid library. Unfortunately, no amplification product was obtained, possibly because of the high degree of degeneracy of the 16 nt primers (1024-fold and 2048-fold respectively).

An attempt was made to reduce the degeneracy of the probe sequence by adopting the preferred GC-rich *Streptomyces* codon usage (Kieser *et al.* 2000). The resulting modified sequence was used to generate a 20 nt probe (Cyp probe 1) with only four-fold degeneracy and full length probe of 66 bp (Cyp probe 2) with 65536-fold degeneracy (Figure III.3.). Although the latter probe appears highly degenerate, it is only 64-fold more degenerate than the much shorter 16 nt primers

Cyp deg F and Cyp deg R. Cyp probe 1 and 2 were used subsequently for Southern hybridisation (Chapter III.3.2.) and ligation-mediated PCR (LM-PCR) (Chapter III.3.3.).



Predicted propeptide sequence and reverse translation:

Ala Thr Pro Ala Thr Pro Thr Val Ala Gln Phe Val Ile Gln Gly Ser Thr Ile Ser Leu Val Cys

Cyp deg F →

GCN ACN CCN GCN ACN CCN ACN GTN GCN CAR TTY GTN ATH CAR GGN WSN ACN ATH WSN YTN GTN TGY

← Cyp deg R

Degenerate probes (preferential *Streptomyces* codon usage):

GCS ACS CCS GCS ACS CCS ACS GTS GCS CAG TTC GTS ATC CAG GGS ICS ACS ATC ICS CTS GTS TGC

GCS CAG TTC GTS ATC CAG GG

Figure III.3. Design of degenerate primers and probes for *cypA*. The predicted propeptide sequence for *cypA* was reverse translated and the resulting degenerate sequence was used to design primers Cyp deg F and Cyp deg R as indicated by the blue arrows. This sequence was then adapted to the preferential *Streptomyces* codon usage to reduce degeneracy. Degenerate positions are indicated in red, codon positions for which other bases are possible are underlined. N = A, C, T or G; S = C or G; W = A or T; R = A or G; Y = C or T; H = A, C or T.

III.3.2. Southern hybridisation-based screening

Cyp probe 2 was labelled with digoxigenin (DIG) and used in a Southern hybridisation experiment to probe *S. sp.* OH-4156 genomic DNA digested with an enzyme for which recognition sites were predicted to occur frequently (*Bam*HI), intermediately (*Bgl*II) or infrequently (*Eco*RI). Despite the use of low-stringency conditions in the hybridisation, no hybridising fragments were identified with the Cyp probe 2. A possible explanation could be that detection with DIG-labelled degenerate probes of this size is not sensitive enough.

To increase the sensitivity of the Southern hybridisation, Cyp probe 2 was labelled with α-^{32}P-dCTP (Chapter II.5.2.). The resulting probe was used in a hybridisation with a membrane that was spotted with 3072 clones from the genomic cosmid library of *S. sp.* OH-4156. After incubation under low-stringency conditions, a high background signal was detected, but two double dot patterns could be identified indicating putative hybridising cosmid clones (Figure III.4.). Subsequent washes of the membrane at stepwise temperature increments succeeded in removing the background noise, but no additional pairs of dots were observed. The two candidate cosmids were taken from the library and spotted onto a nylon membrane. Unfortunately, neither cosmid hybridised with the Cyp probe 2, suggesting spurious initial hybridisation.

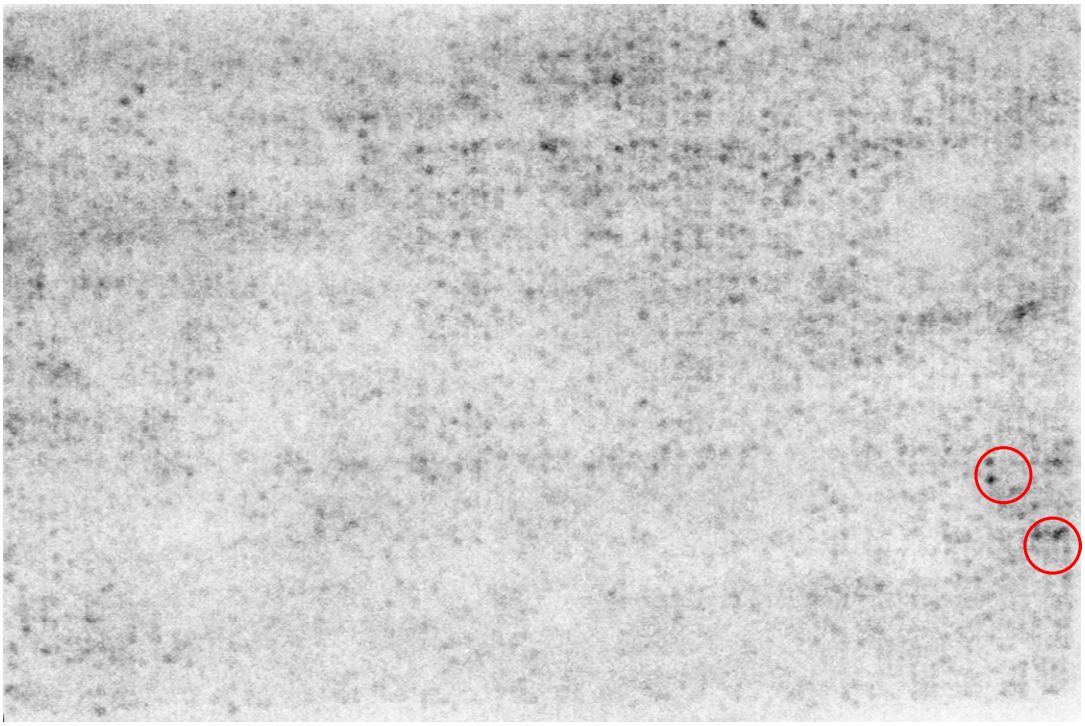


Figure III.4. Hybridisation of a nylon membrane containing *S. sp.* OH-4156 cosmid library clones with α-^{32}P-CTP labelled Cyp probe 2. Two double dot patterns indicative of putative hybridising cosmids are circled.

III.3.3. Ligation-mediated-PCR

An LM-PCR-based approach was taken to try and acquire sequence information of the *cypA* region that could be used to generate a specific probe. LM-PCR is a technique that was previously used to identify the location of transposon insertions in the genome of *S. coelicolor*. The principle is based on a PCR with one primer that is homologous to a known sequence and a second primer that anneals to a designed linker sequence (Figure III.5.). Genomic DNA is digested with a frequently cutting restriction enzyme (such as *Eag*I) and subsequently ligated to double stranded oligonucleotides with complementary overhanging ends. The resulting ligation product is used as template in the LM-PCR reaction. Only fragments that contain the complementary sequence to the internal primer will generate an amplification product.

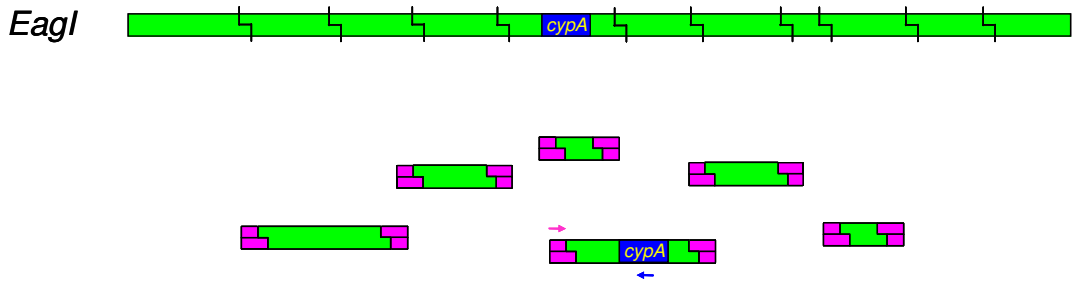


Figure III.5. The principle of Ligation-Mediated-PCR (LM-PCR). Genomic DNA is digested with a frequent cutting restriction enzyme (e.g. *Eag*I). Double stranded oligonucleotides with complementary overhanging sequences (depicted in pink) are ligated to the restriction fragments. The resulting ligation product is used as template in the LM-PCR reaction with one linker primer (pink arrow) and one primer complementary to a known sequence (blue arrow).

The LM-PCR approach was applied to *S. sp.* OH-4156 genomic DNA (and the two putative positive cosmids identified in Chapter III.3.2.) as template, using an *Eag*I linker and Cyp probe 1 as the internal primer. An advantage of LM-PCR over amplification with degenerate primers (Chapter III.3.1.) is that only one degenerate primer is used in the reaction mix. Also, the expected PCR products are larger, facilitating their detection by agarose gel-electrophoresis. The LM-PCR reaction

yielded a number of different sized amplicons that were cloned in pGEM-T Easy for sequencing. Unfortunately, none of the resulting sequences appeared to be derived from the *cypA* region of the genome, indicating that all presumably resulted from non-specific annealing of Cyp probe 1.

III.4. Genome scanning of *S. sp.* OH-4156

Since the strategies described in the previous section all proved to be unsuccessful, Solexa-based genome scanning of *S. sp.* OH-4156 was carried out. 15471 contigs were generated with a median length of 378 bp and an additive length of 8.5 Mb, typical of a *Streptomyces* genome (Bentley *et al.* 2002). A tBLASTn search of the contig database with cypemycin's predicted propeptide sequence identified a 1888 bp contig containing the cypemycin precursor gene, *cypA*, confirming that cypemycin is indeed ribosomally synthesized (Figure III.6.).

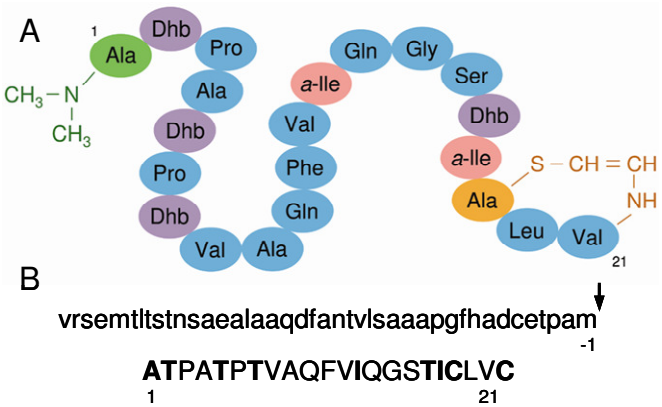


Figure III.6. Sequence of the cypemycin prepropeptide deduced by genome scanning. A) Schematic representation of fully modified cypemycin and B) the unmodified CypA preproprotein.

The 22 amino acid CypA propeptide sequence revealed the identity of the residues subjected to post-translational modification. As expected, the four Dhb residues of cypemycin result from dehydration of Thr, Me_2-Ala is formed by methylation of an alanine residue and the two L-*allo*-Ileu are introduced by side-chain isomerization of L-Ile. Surprisingly, the AviCys modification is formed from two Cys residues. This

is unprecedented - previously described AviCys residues are generated from an internal Ser and a C-terminal Cys. The first step in AviCys formation is LanD-mediated oxidative decarboxylation of the propeptide's C-terminal Cys. This yields an enethiol intermediate that can form the AviCys after addition to a Dha, created by dehydration from the internal Ser (Blaesse *et al.* 2000).

As in most other post-translationally modified peptides, the propeptide is preceded by a leader sequence that is removed during or after maturation (Oman and van der Donk 2010). The CypA propeptide is preceded by a 42 amino acid leader sequence that does not display similarity to lantibiotic leader peptides. Moreover, while lantibiotic leader sequences are always devoid of Cys (Chatterjee *et al.* 2005), the CypA leader peptide contains a Cys at position -6 relative to the cleavage site.

Additional tBLASTn searches of the Solexa data were performed to identify putative cypemycin biosynthetic genes. Using amino acid sequences of LanB and LanM dehydratases from the NCBI database, a few contigs with apparent homology to *lanM* were identified. Physical linkage between these and the *cypA* contig could not be established by PCR analysis. While tBLASTn searches with LanD decarboxylases did not identify a LanD homologue, a contig with the partial sequence of a 4'-phosphopanthothenoylcysteine (PPC) decarboxylase was found. PPC decarboxylases belong to the homo-oligomeric flavin-containing Cys-decarboxylase (HFCD) family that also encompasses LanD proteins (Kupke *et al.* 2000).

BLAST searches revealed that contigs from the *S. sp.* OH-4156 genome scanning closely resembled the genome sequence of *S. venezuelae*. When the *S. sp.* OH-4156 contigs were mapped onto their orthologous sequences in the *S. venezuelae* genome (>95 % identity for contigs of 100 bp or longer), only a few gaps remained for which no corresponding contigs were identified (Govind Chandra, Figure III.7.). Interestingly, most of the known or predicted secondary metabolite gene clusters from *S. venezuelae* fall into these gaps. Conversely, the four putative lantipeptide clusters in *S. venezuelae* (Chapter VII.3.) all have an orthologous counterpart in *S. sp.* OH-4156.

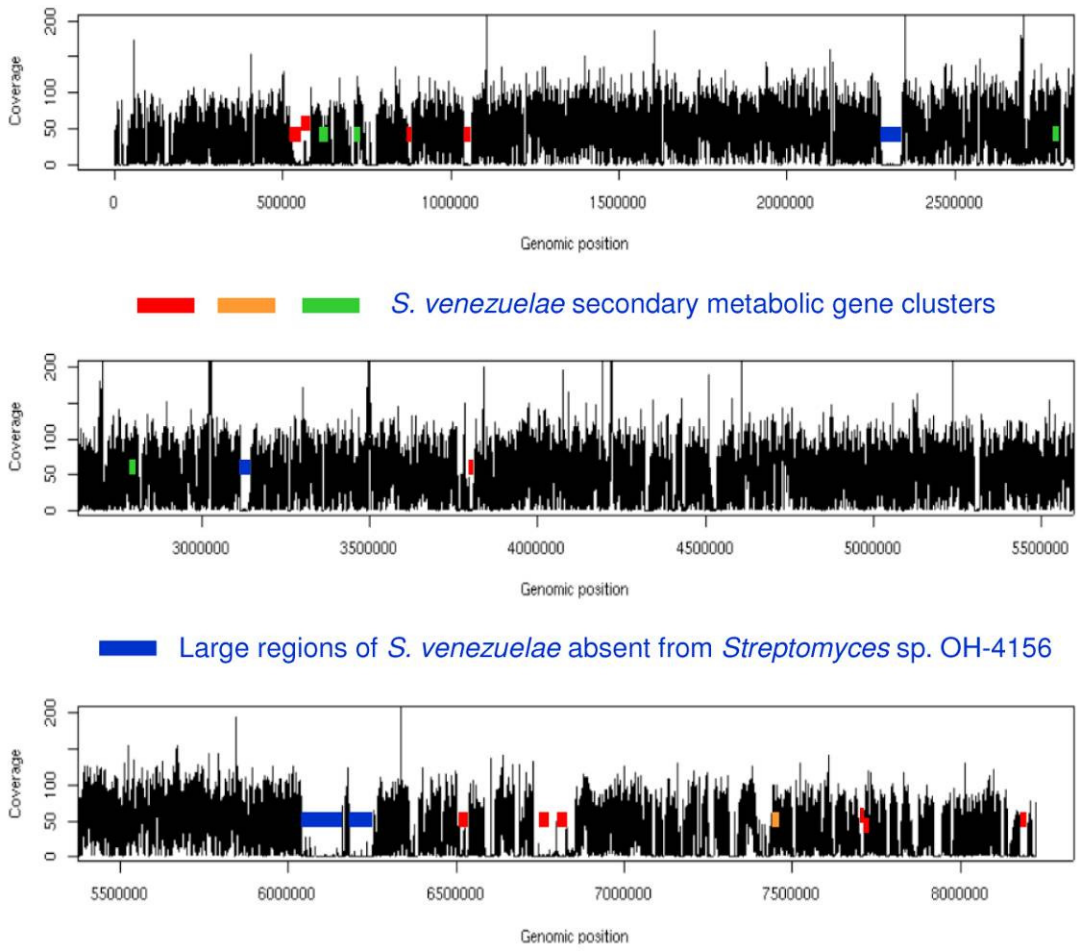


Figure III.7. Comparison between *S. sp.* OH-4156 genome scanning data and the *S. venezuelae* genome. Orthologous *S. sp.* OH-4156 contigs were mapped onto the genome sequence of *S. venezuelae* (>95 % identity for contigs of 100 bp or longer). Areas that are not covered by the *S. sp.* OH-4156 contigs are colour coded: Orthologous predicted or known secondary metabolism gene clusters present in both species (green), partially present in *S. sp.* OH-4156 (orange) and absent from *S.sp.* OH-4156 (red) are shown. Regions of the *S. venezuelae* genome not represented in *S. sp OH-4156* and lacking predicted secondary metabolite gene clusters are shown in blue.

An interesting observation was made when the gaps corresponding to the *S. venezuelae* secondary metabolite gene clusters were investigated in more detail. The predicted extent of the *S. venezuelae* gene clusters generally correlated well with the boundaries for the orthologous *S. sp.* OH-4156 sequence data. This is illustrated for the chloramphenicol gene cluster (Figure III.8.). He *et al.* (2001) reported the involvement of SMD01679 to SMD01694 in chloramphenicol biosynthesis. The region onto which orthologous *S. sp.* OH-4156 contigs map in the *S. venezuelae* genome starts eight genes to the left and ends four genes to the right of the chloramphenicol gene cluster, suggesting that these additional flanking genes may also be involved in chloramphenicol biosynthesis. Similar examples of corresponding gaps are found for seven additional *S. venezuelae* gene clusters (data not shown).

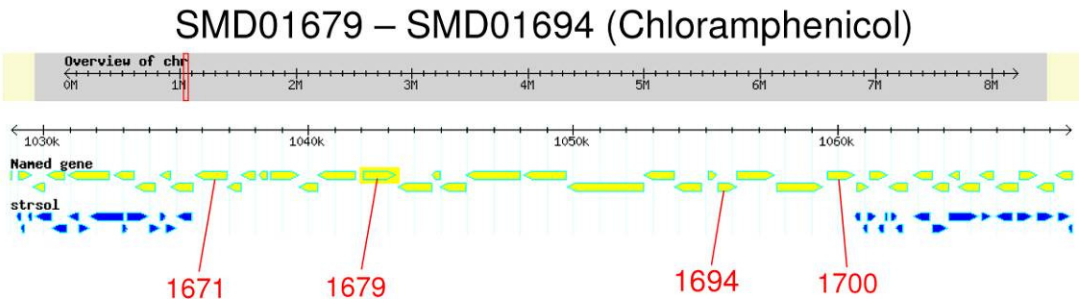


Figure III.8. Analysis of the *S. venezuelae* chromosomal region harbouring the chloramphenicol biosynthetic gene cluster. *S. venezuelae* genes are represented by yellow arrows, orthologous *S. sp.* OH-4156 contigs by blue arrows. The published chloramphenicol gene cluster spans from SMD01679 to SMD01694. Note: the *S. venezuelae* gene annotations do not have successive numbers in this region of the genome.

III.5. Identification of the cosmids bearing the cypemycin biosynthetic cluster

III.5.1. Generation of specific *cypA* probe and library screen

The Solexa data was used to generate a specific 480 bp probe by PCR using primers *cypA* T1 and T2 that encompassed *cypA* and flanking sequences. The probe was labelled with α-^{32}P-dCTP and hybridized to a nylon membrane spotted with the *S. sp.* OH-4156 genomic cosmid library (Chapter II.3.13.). Fourteen hybridizing cosmids were identified (Figure III.9.). Very low levels of background hybridisation were observed in this experiment compared to the previous hybridisation with the Cyp probe 2 (Figure III.4., Chapter III.3.2.). The two cosmids identified with the Cyp probe 2 were not picked up with the specific 480 bp probe, suggesting earlier spurious hybridisation.

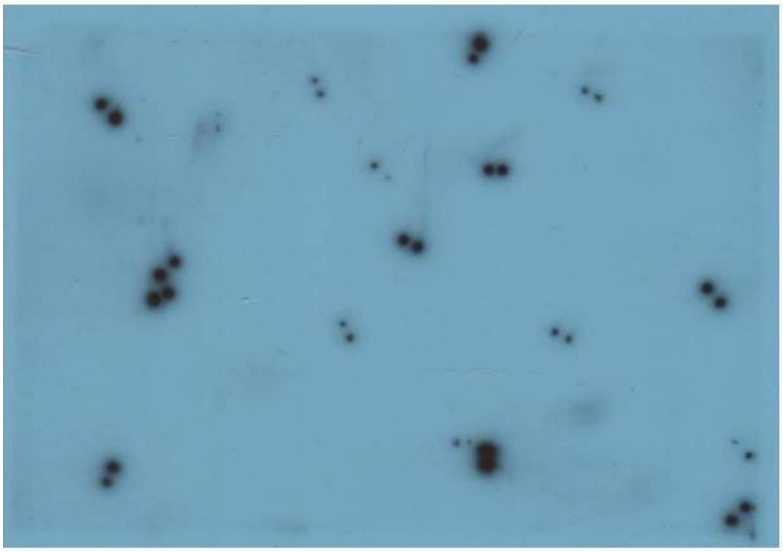


Figure III.9. Hybridisation of a nylon membrane spotted with 3072 *S. sp.* OH-4156 genomic cosmid library clones with a 480 bp ^{32}P-labelled *cypA* probe. Fourteen putative *cypA*-containing cosmids were identified from their double dot patterns.

III.5.2. Heterologous expression of the positive cosmids in *S. venezuelae*

To determine which of the cosmids contained all of the genes required for cypemycin production, the vector backbones of nine of the hybridizing cosmids were PCR-targeted with a 5.2 kb *Ssp*I fragment from pIJ10702 containing *oriT*, enabling the cosmid to be conjugated into a *Streptomyces* host, and a ϕC31 integrase gene and phage attachment site for integration of the cosmid into the genome of the heterologous host. Stable integration at the ϕC31 *attB* site obviates the need for antibiotic selection, greatly facilitating subsequent bioactivity assays. *Streptomyces venezuelae* was chosen as an initial heterologous host because of the relatively high level of nucleotide sequence identity between its genome and the Solexa data from *S. sp.* OH-4156. The nine exconjugants and a control strain with the integrated cosmid backbone were assessed for cypemycin production in a *M. luteus* bio-assay and by MALDI-ToF MS. Six of the nine strains produced cypemycin with MALDI-ToF peaks of: $[M+H]^+$ = 2096 Da, $[M+Na]^+$ = 2118 Da and $[M+K]^+$ = 2134 Da, indicating that the integrated cosmids contained all of the genes required for cypemycin production (Figure III.10.).

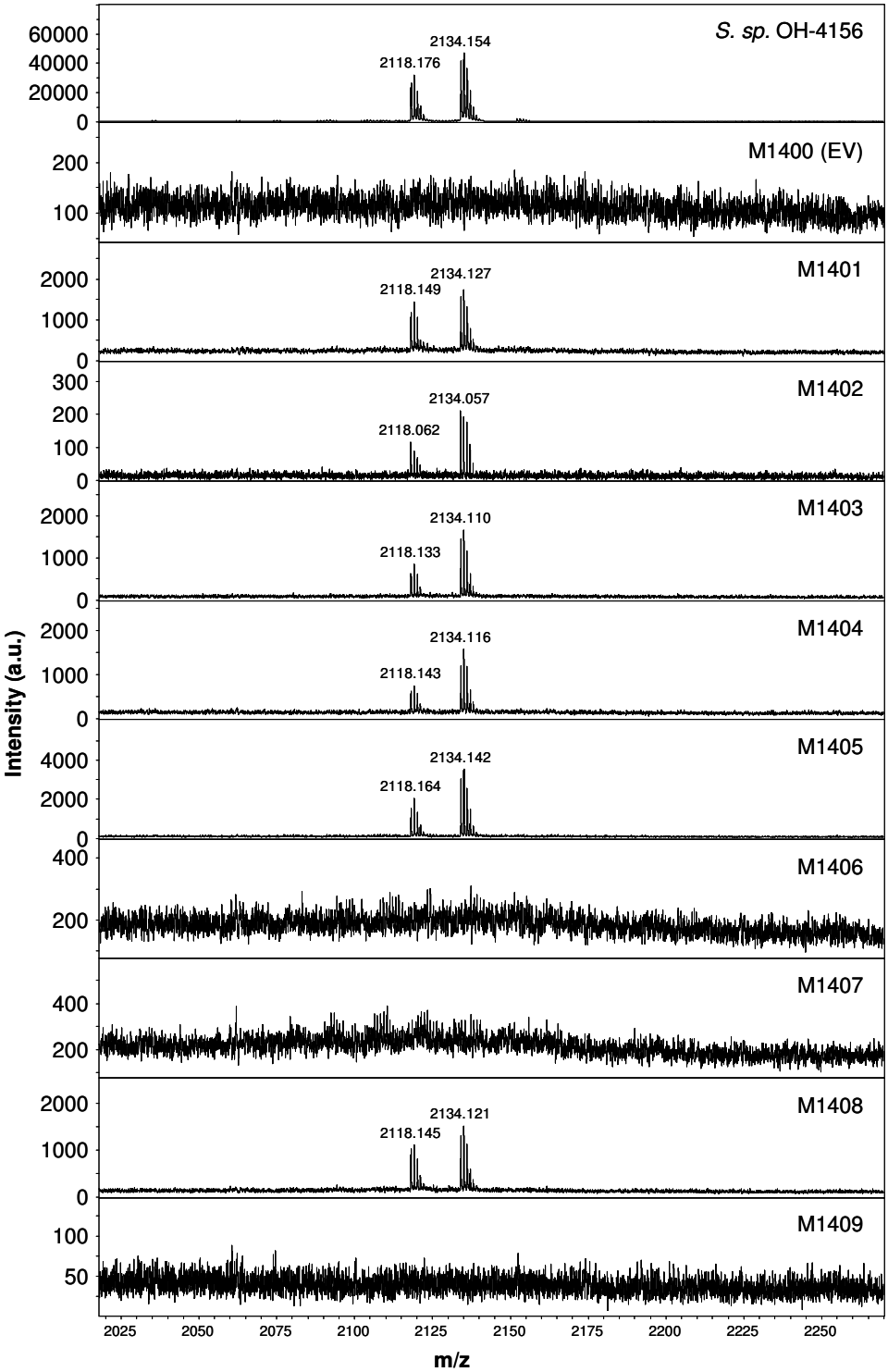


Figure III.10. Heterologous expression of the cypemycin biosynthetic gene cluster in *S. venezuelae.* MALDI-ToF spectra for *S. sp.* OH-4156, *S. venezuelae* with the empty vector pIJ10702 (M1400), and *S. venezuelae* derivatives with cosmids containing all of the genes required for cypemycin production (M1401, M1402, M1403, M1404, M1405 and M1408) and those with a cosmid that does not (M1406, M1407 and M1409).

Levels of cypemycin production in the *S. venezuelae* exconjugants, judged by the sizes of inhibition zones (Figure III.11.), were much lower than those of the natural producer. Since it gave the largest zone of inhibition upon heterologous expression, pIJ12404 was chosen for sequencing.

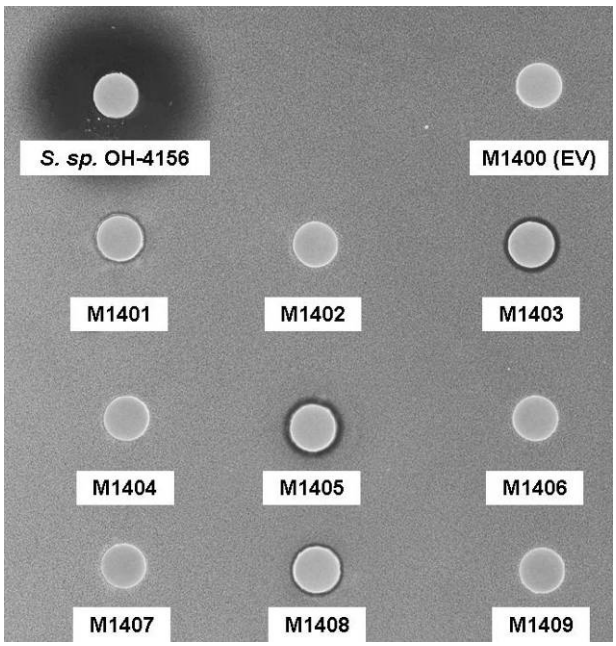


Figure III.11. Heterologous expression of the cypemycin biosynthetic gene cluster in *S. venezuelae. M. luteus* bio-assay of the nine *S. venezuelae* derivatives with a *cypA*-containing cosmid. *S. sp.* OH-4156 positive control, M1400 (EV, empty vector) negative control, and M1401-M1409 are the *S. venezuelae* clones containing hybridizing library cosmids. The halos observed for M1401, M1403, M1405 and M1408 are much smaller than that of the natural producer; while clones M1402 and M1404 failed to produce a halo, cypemycin production was confirmed by MALDI-ToF analysis.

III.6. The cypemycin biosynthetic cluster

Analysis of open reading frames (ORFs) in the pIJ12404 nucleotide sequence identified a putative biosynthetic gene cluster of nine genes (Figure III.12.A.). Upstream of *cypA* and divergently transcribed from it is *orf1*, which encodes a putative transcriptional regulator. Seven genes with predicted biosynthetic and transport functions lie directly downstream of *cypA*, the first six of which are likely to be co-transcribed since their coding sequences overlap by several nucleotides (i.e. they appear to be translationally coupled). The first gene downstream of *cypA* is *cypH*, which encodes a product with no significant homology to functionally characterized proteins. CypH is likely a two-domain protein with partial homology to a conserved horizontally transferred transmembrane helix domain at its N-terminus (Schultz 2004) and a possible α/β hydrolase fold located towards its C-terminus. Following *cypH* is *cypL*, again with no functionally identified homologs. The next gene, *cypD*, encodes a decarboxylase of the HFCD family (Kupke *et al.* 2000), and could be responsible for the introduction of AviCys, analogous to LanD enzymes in lantibiotic biosynthesis. *cypM* is the fourth gene downstream of *cypA* and encodes an *S*-adenosyl methionine (SAM)-dependent methyltransferase presumably required for methylation of the N-terminal alanine of cypemycin. These biosynthetic genes are followed by an adenosine triphosphate (ATP)-binding subunit of an ATP-binding cassette (ABC) transporter, designated *cypT*. The last gene in which the start codon overlaps with the upstream ORF is *cypP*, encoding a 516 amino acid protein with 12 predicted transmembrane helices and with no functionally annotated homologues in the public databases (Figure III.12.B.). Because of its location downstream of *cypT*, it was suggested that this protein forms a pore in the cytoplasmic membrane to allow cypemycin export from cell. No leader peptidase domains were found in CypT or CypP. *cypI* is the last gene that is transcribed in the same direction, but its start codon does not overlap with the coding sequence of *cypP*. CypI is a member of the large DUF255 family of conserved proteins with a thioredoxin domain. Analysis of the remaining ORFs in the pIJ12404 cosmid sequence did not identify any lantibiotic dehydratase homologues indicating, contrary to previous classification, that cypemycin is not a member of the lantibiotic family of modified peptides (Chatterjee *et al.* 2005).

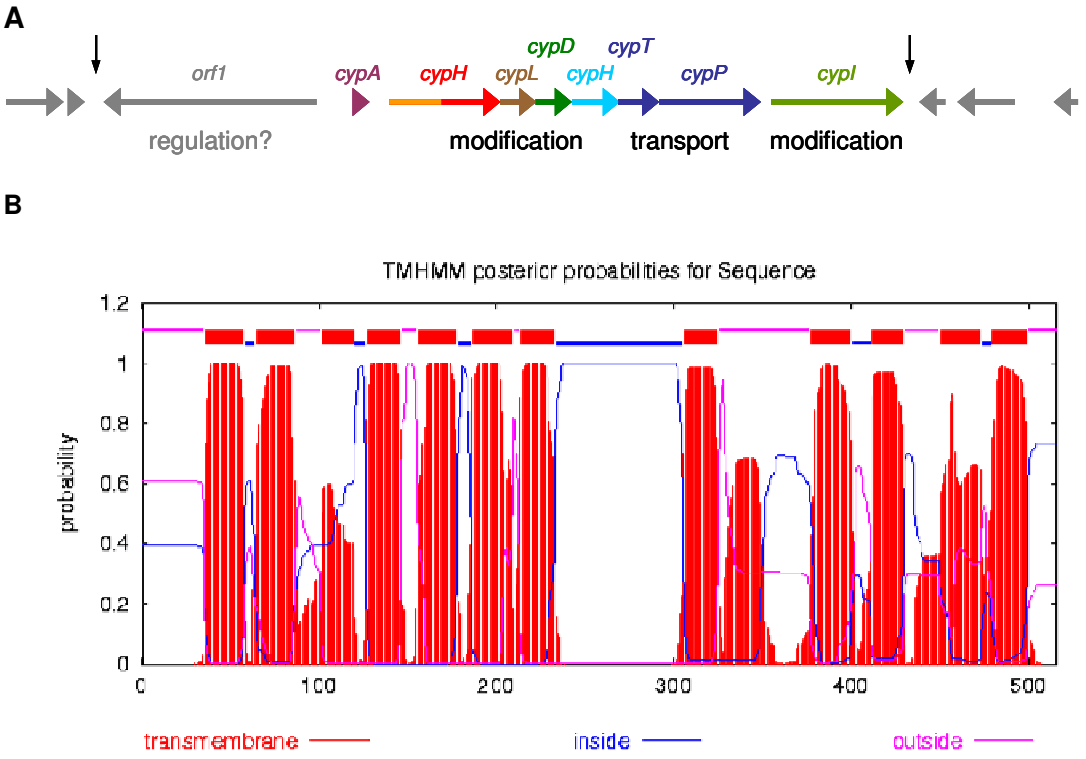


Figure III.12. A) Schematic representation of the cypemycin biosynthetic gene cluster. Flanking genes not expected to be involved in cypemycin biosynthesis are in grey. Vertical arrows delineate the putative *cyp* gene cluster used to generate a minimal gene set. B) CypP is a putative pore protein. TMHMM analysis identified 12 putative transmembrane domains (red). Predicted cytoplasmic domains are indicated in blue and exterior domains in pink.

III.7. Discussion

Various attempts were made to identify the gene cluster responsible for the biosynthesis of the unusual post-translationally modified peptide antibiotic cypemycin. This section is devoted to a discussion of these approaches. Since the next chapter describes the genetic characterisation of the cypemycin gene cluster, the reader is referred to Chapter IV.6. for a detailed discussion of the functions of individual genes.

Several experiments described in this chapter relied on the use of degenerate primers or probes to obtain sequence information for *cypA*. The main factors that likely contributed to the failure of these experiments were the short length and degree of degeneracy of the primers/probes, and the reduced sensitivity of DIG-labelling compared to radioactive labelling. The predicted CypA propeptide sequence turned out to be correct apart from the occurrence of a Cys at position 19 where Ser was expected (compare Figure III.3. and III.6.B.). Although this probably had only a minor effect on Southern hybridisations with Cyp probe 2, the PCR reaction using Cyp deg R would have been markedly compromised.

Alternative approaches that could have been used to identify the cypemycin gene cluster could have included a genetic screen for likely biosynthetic enzymes and an activity based screen. However, when employed for cypemycin, both would have been problematic. Starting from the assumption that cypemycin was a lantibiotic, a genetic screen could have been directed at the identification of lantibiotic dehydratases, for example by using primers complementary to conserved regions found in the corresponding genes (Dodd *et al.* 2006). Since it later emerged that cypemycin belongs to a different peptide family, this would have been a futile exercise. An activity based screen could have been performed by introducing the *S. sp.* OH-4156 cosmid library in a heterologous host. Lantibiotic and other post-translationally modified peptide gene clusters are relatively small and there is a reasonable chance that at least a couple cosmids from a genome library will contain all of the necessary genes. However, it is difficult to predict whether the cluster will be expressed, and if so, at what level. In the case of cypemycin, heterologous production was observed in *S. venezuelae*, *S. lividans* and *S.*

coelicolor, but production levels varied markedly between these strains and the expression levels were generally too low to allow for a screen with single colonies against *M. luteus*.

In summary, with the benefit of hindsight, it can be concluded that the genome scanning approach was particularly useful for the identification of the cypemycin gene cluster. The unusual precursor peptide and the novelty of the biosynthetic enzymes would have made the cluster difficult to find using other means.

III.8. Bullet point summary

- A genomic cosmid library was generated for the cypemycin producer *S. sp.* OH-4156. Approaches based on degenerate nucleotide sequences derived from mature cypemycin failed to identify cosmids harbouring the biosynthetic gene cluster.

- A genome scanning approach yielded a contig containing *cypA*. The CypA propeptide contains a Cys at position 19, where a Ser would be expected based on previously documented peptides with the AviCys modification.

- A specific probe for the structural gene region identified cosmids from the library that contained *cypA*. Subsequent heterologous expression identified cosmids that contain all of the genes required for cypemycin biosynthesis.

- Bio-informatic analysis of the cypemycin gene cluster revealed novel enzymes and indicated that cypemycin is not a lantibiotic.

Chapter IV – Mutational analysis of cypemycin biosynthesis

IV.1. Introduction

The cypemycin biosynthetic gene cluster was identified in Chapter III and heterologous expression in *S. venezuelae* demonstrated. This chapter describes the genetic characterisation of the gene cluster. First, the generation of a minimal gene set will be described, identifying the genes required for cypemycin production. Next, the function of each individual gene from this minimal set will be analysed by constructing in-frame deletion mutants in the heterologously expressed cluster and marked mutants in the natural producer. A bioinformatic search for related gene clusters revealed a novel family of post-translationally modified peptides. Finally, an attempt to generate cypemycin variants will be mentioned.

IV.2. Generation of a minimal gene set

Bioinformatic analysis suggested that the cypemycin biosynthetic gene cluster extended from *cypA* (or possibly *orf1*) to *cypI* (Figure III.12.A.). The region to the left of and including *orf1* is syntenous with SCO4966 to SCO4969 in *S. coelicolor* (the latter being the *orf1* ortholog) (Bentley *et al.* 2002). Genes to the left of *orf1* are predicted to be involved in mycothiol detoxification and no function in cypemycin biosynthesis is envisaged. Genes to the right of *cypI* encode rodlins and a chaplin (homologues of SCO2716 to SCO2719 (Bentley *et al.* 2002)) that have been implicated in morphological development in *S. coelicolor* (Claessen *et al.* 2004) and therefore are also unlikely to be involved in cypemycin biosynthesis.

Because no convenient restriction sites were available to excise the putative minimal cypemycin biosynthetic gene cluster and subsequently confirm its identity, a PCR targeting strategy was devised to introduce unique restriction sites flanking the putative gene cluster (Figure IV.1.) Briefly, pIJ12404 was PCR-targeted to the left of *orf1* to introduce a unique *Xba*I restriction site. The antibiotic resistance cassette was removed by FLP-mediated recombination and the resulting cosmid

targeted a second time downstream of *cypI*, introducing a unique *Ssp*I restriction site. The 12.2 kb fragment containing the putative minimal gene cluster was subsequently excised by digestion with *Xba*I and *Ssp*I and ligated into *Xba*I/*Eco*RV-digested pSET152 to give pIJ12421. *S. venezuelae* was abandoned as a heterologous host because of low levels of cypemycin production. Instead, *S. coelicolor* M1146, from which four antibiotic gene clusters had been deleted and which lacks antibiotic activity, was used. Upon conjugation, pIJ12421 integrated as a single copy into the ϕC31 phage attachment site of M1146, yielding a stable strain (M1412) that did not require antibiotic selection to maintain the construct. Cypemycin production was confirmed by both an inhibition assay against *M. luteus* and MALDI-ToF analysis. The halo produced in the bio-assay was comparable in size to that produced by M1411 (M1146 harboring pIJ12413), indicating that the putatively assigned minimal gene set was indeed sufficient for cypemycin production in a heterologous host (Figure IV.2.A.).

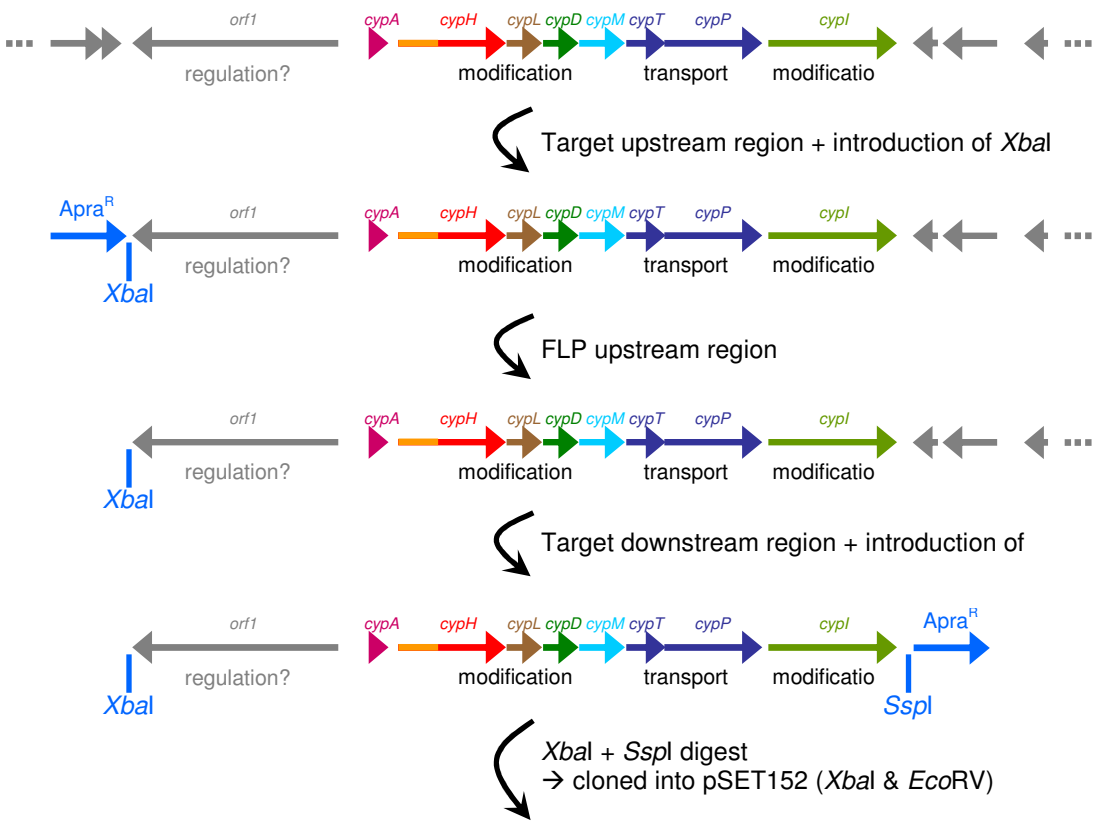


Figure IV.1. Strategy for the construction of a minimal *cyp* gene set. Starting with cosmid pIJ12404, PCR-targeting was used to introduce an apramycin resistance cassette and unique *Xba*I restriction site upstream of *orf1*. The resistance gene was removed by FLP recombination and the resulting construct subjected to a second targeting event downstream of *cypI* with a different apramycin resistance cassette, this time introducing a unique *Ssp*I site. The 12.2 kb gene cluster was then excised by digestion with *Xba*I and *Ssp*I, and cloned into pSET152 digested with *Xba*I and *Eco*RV.

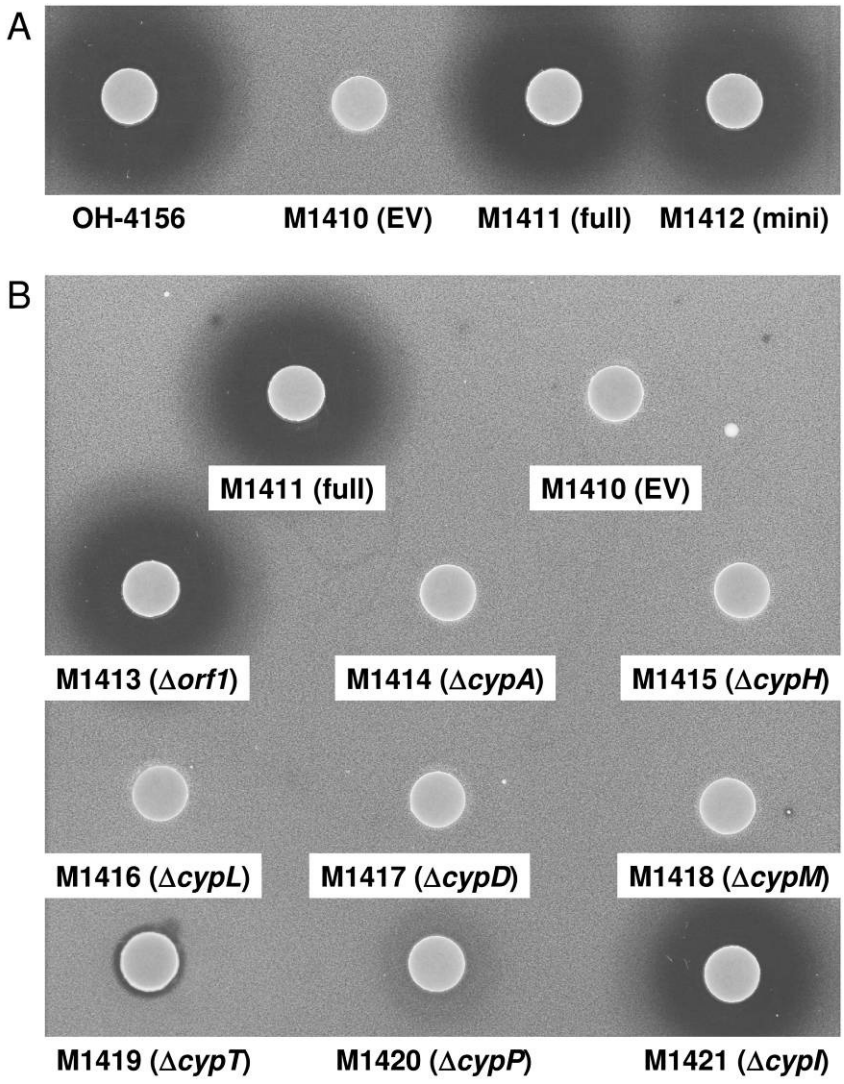


Figure IV.2. *M. luteus* bio-assays for A) The minimal gene set construct. B) the different scar mutants in the heterologous host M1146.

IV.3. Mutational analysis of the cypemycin cluster

To investigate the function of each individual gene within the minimal gene set, individual in-frame 'scarred' deletion mutants were generated by PCR-targeting of pIJ12404. The backbones of the mutagenized cosmids were subsequently targeted with the 5.2 kb *Ssp*I fragment from pIJ10702 to allow integration into the ϕC31 attachment site of M1146. Data obtained from heterologous expression in M1146 (Figure IV.2.B. & IV.3.) was corroborated by generating apramycin-marked deletions of all genes in the minimal gene set in *S. sp.* OH-4156 (Figure IV.4.). Unless otherwise stated, in the following discussion the mutant described is always the heterologously expressed in-frame deletion mutant. The Δ*cypA* strain did not produce cypemycin (confirmed by MALDI-ToF) and no halo was produced in the bio-assay with *M. luteus* (Figure IV.2.B.).

Deletion of the putative regulatory gene *orf1* had no effect on cypemycin biosynthesis (confirmed by bioassay and MALDI-ToF analysis) consistent with production of cypemycin by the Δ*orf1* mutant of *S. sp.* OH-4156. Individual deletion of *cypH* and *cypL* abolished cypemycin production. The Δ*cypD* knock-out strain produced a compound with a mass corresponding to non-decarboxylated cypemycin ($[M+H]^+$ = 2142 Da, $[M+Na]^+$ = 2164 Da and $[M+K]^+$ = 2180 Da) that did not inhibit growth of *M. luteus*. It was hypothesized that Cys19 in this compound would be modified to Dha and that Cys22 would still be intact because of the lack of decarboxylation. To confirm this, alkylation with iodoacetamide was carried out to identify the free thiol groups in the peptide. A mass increased of 57 Da indicated the presence of only one free thiol group (Figure IV.5.A.). To confirm that this thiol group was derived from the C-terminal Cys, the presence of the Dha residue at the position of the first Cys was verified by Q-ToF analysis (Figure IV.5.B.).

M1146 (Δ*cypM*) also produced a truncated version of cypemycin detectable by MALDI-ToF ($[M+H]^+$ = 2068 Da, $[M+Na]^+$ = 2090 Da and $[M+K]^+$ = 2106 Da), but failed to produce a zone of inhibition against *M. luteus*. The mass corresponds to a non-methylated version of cypemycin, which was confirmed by Q-ToF analysis (Figure IV.5.C.).

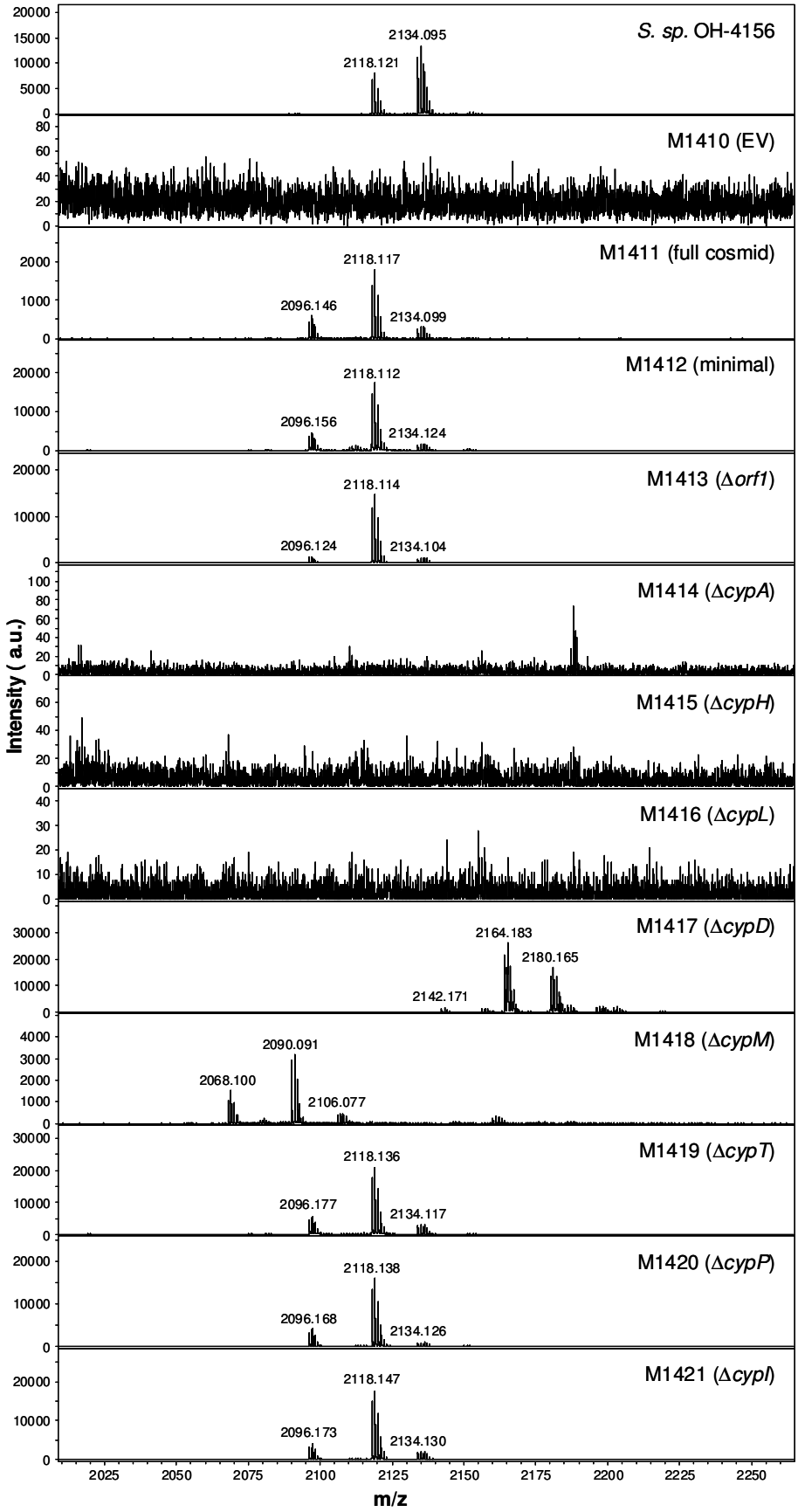

S. sp. OH-4156
2134.095
2118.121
M1410 (EV)
M1411 (full cosmid)
2118.117
2096.146
2134.099
M1412 (minimal)
2118.112
2096.156
2134.124
M1413 (Δorf1)
2118.114
2096.124
2134.104
M1414 (ΔcypA)
M1415 (ΔcypH)
M1416 (ΔcypL)
M1417 (ΔcypD)
2164.183
2180.165
2142.171
M1418 (ΔcypM)
2090.091
2068.100
2106.077
M1419 (ΔcypT)
2118.136
2096.177
2134.117
M1420 (ΔcypP)
2118.138
2096.168
2134.126
M1421 (ΔcypI)
2118.147
2096.173
2134.130
Intensity (a.u.)
m/z

Figure IV.3. MALDI-TOF spectra for scar mutants in M1146. Spectra for *S. sp.* OH-4156 positive control, M1146 with the inserted empty vector pIJ10702 (M1410), full cosmid (M1411), minimal gene set (M1412) and individual mutants in *orf1* to *cypI* (M1413-M1421). Cypemycin is detected as $[M+H]^+$ = 2096 Da, $[M+Na]^+$ = 2118 Da and $[M+K]^+$ = 2134 Da. The non-decarboxylated version of cypemycin (in M1417) appears as $[M+H]^+$ = 2142 Da, $[M+Na]^+$ = 2164 Da and $[M+K]^+$ = 2180 Da and the non-methylated intermediate (in M1418) as $[M+H]^+$ = 2068 Da, $[M+Na]^+$ = 2090 Da and $[M+K]^+$ = 2106 Da.

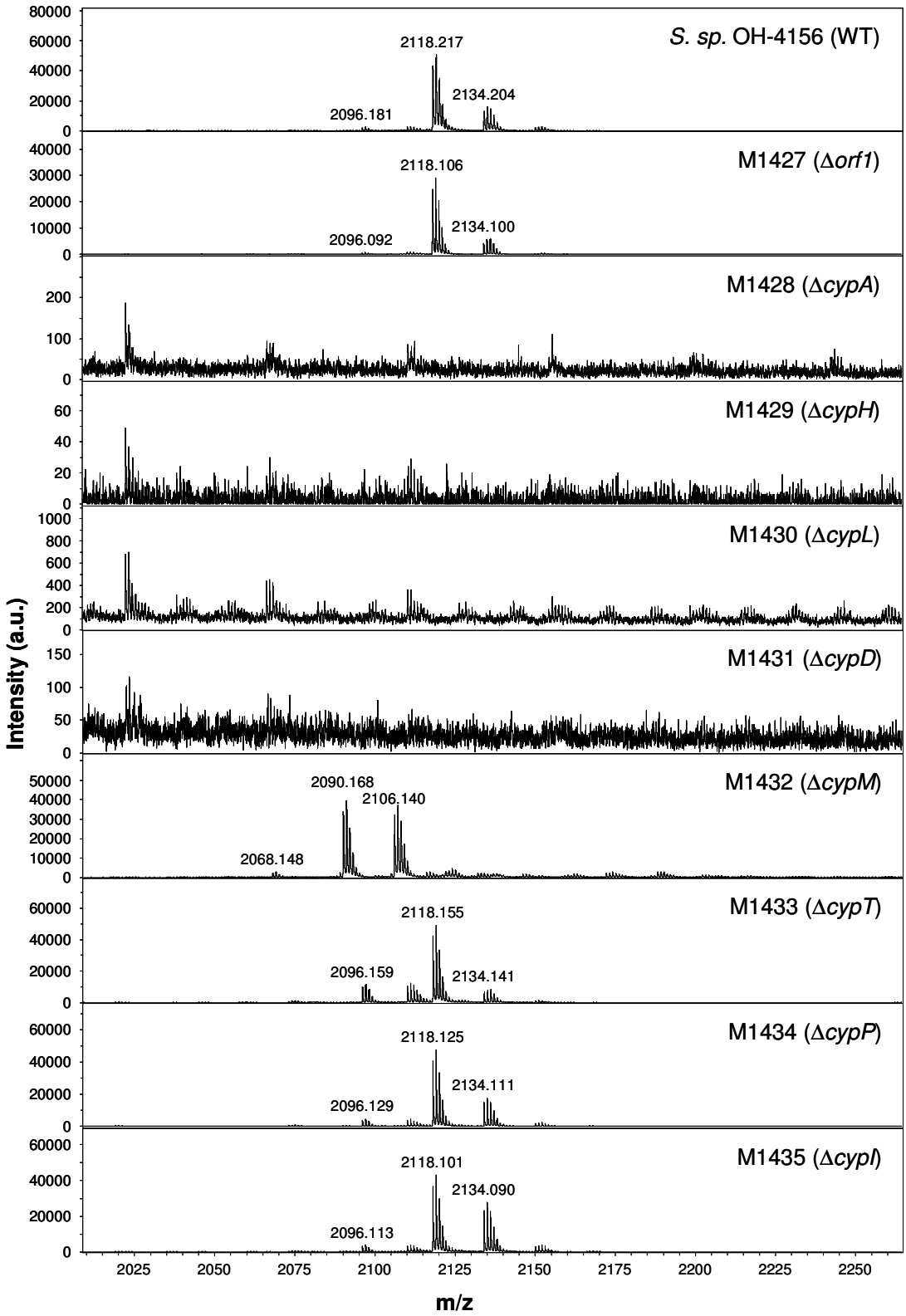
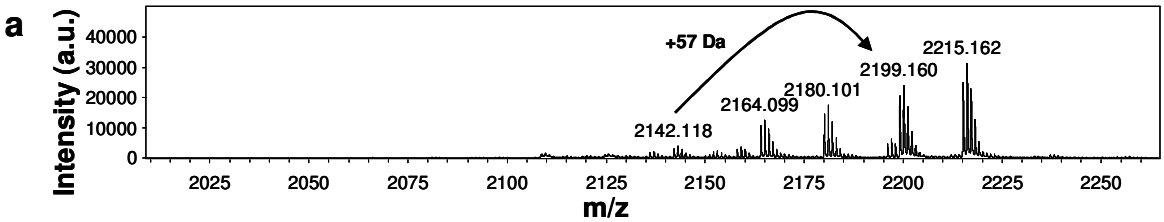

S. sp. OH-4156 (WT)
2118.217
2134.204
2096.181
M1427 (Δorf1)
2118.106
2134.100
2096.092
M1428 (ΔcypA)
M1429 (ΔcypH)
M1430 (ΔcypL)
M1431 (ΔcypD)
M1432 (ΔcypM)
2090.168
2106.140
2068.148
M1433 (ΔcypT)
2118.155
2096.159
2134.141
M1434 (ΔcypP)
2118.125
2134.111
2096.129
M1435 (ΔcypI)
2118.101
2134.090
2096.113
Intensity (a.u.)
m/z

Figure IV.4. MALDI-ToF spectra of the apramycin-marked mutants in *S. sp.* OH-4156. Spectra for M1427 (Δ*orf1*) to M1435 (Δ*cypI*) show the same phenotypes as the heterologously expressed mutants (See Fig. S5). No cypemycin intermediate was detected in the culture of M1431 (Δ*cypD*), whereas its heterologous counterpart (M1417) produced non-decarboxylated cypemycin. A possible explanation for this is that the replacement of *cypD* in this mutant with the pIJ773-derived cassette is causing a polar effect on the downstream *cypM* gene, making M1431 a Δ*cypDM* double mutant. The product of this mutant (cypemycin lacking the AviCys and the Me_2-Ala) is thought to be unstable, but has been isolated from a heterologously expressing Δ*cypDM* double mutant (results not shown). Differences in protease activity between M1146 and *S. sp.* OH-4156 could account for this observation.



b Non-decarboxylated

#	Residue	b-ions Calculated	b-ions Observed	y-ions Calculated	y-ions Observed
17	Dhb	1626.864	1626.844	599.323	599.317
18	*a*-Ile	1739.948	1739.972	516.286	516.280
19	Dha	1808.970	-	403.202	403.197
20	Leu	1922.054	-	334.180	334.188
21	Val	2021.122	-	221.096	221.089
22	Cys	2141.134	-	122.027	-

c Non-methylated

#	Residue	b-ions Calculated	b-ions Observed	y-ions Calculated	y-ions Observed
1	Ala	72.045	-	2069.113	-
2	Dhb	155.082	155.088	1997.068	-
3	Pro	252.135	252.138	1914.031	-
4	Ala	323.172	323.179	1816.978	-
5	Dhb	406.209	406.217	1745.941	1745.860
6	Pro	503.262	503.275	1662.904	1662.832

Figure IV.5. Structural verification of the truncated cypemycin derivatives. A) After alkylation of the free thiol groups in the product of the Δ*cypD* mutant, a single mass shift of +57 Da was observed indicating the presence of only one free thiol group. Non-decarboxylated cypemycin appears as $[M+H]^+$ = 2142 Da, $[M+Na]^+$ = 2164 Da and $[M+K]^+$ = 2180 Da. The alkylated version is observed as $[M+H]^+$ = 2199 Da and $[M+O]^+$ = 2215 Da. B) *Q*-ToF analysis of the compound produced by the Δ*cypD* mutant shows that Cys19 has been dehydrated to dehydroalanine. C) Q-ToF analysis confirming that the mass difference of the compound produced by the Δ*cypM* mutant is localized to the first two amino acids and thus is very likely to reflect the lack of the two N-terminal methyl groups.

Deletion of *cypT* or *cypP* resulted in reduced halo sizes, but MALDI-ToF analysis revealed some level of cypemycin production. The *S. sp.* OH-4156 Δ*cypT* and Δ*cypP* mutants displayed similar phenotypes when analysed by bio-assay and MALDI-ToF. It is possible that an alternative transporter could export cypemycin out of the cell or, given the hydrophobic nature of the peptide, that some level of diffusion through the membrane can occur.

Deletion of *cypI* had no effect on the production or activity of cypemycin. By a process of elimination, CypI is a possible candidate for isomerization of the Ile side chains to form L-*allo*-Ile, a modification that cannot be studied by MALDI-ToF since no mass change accompanies the isomerization. Analysis of the *S. sp.* OH-4156 Δ*cypI* strain showed an identical phenotype. If *cypI* is required for introduction of the *allo*-Ile residues in cypemycin, they are not required for antibacterial activity.

The five M1146 derivatives that contained a mutagenized cosmid with a biosynthetic phenotype (M1414 (Δ*cypA*), M1415 (Δ*cypH*), M1416 (Δ*cypL*), M1417 (Δ*cypD*) and M1418 (Δ*cypM*)) were complemented with wild-type copies of the respective genes to confirm that the mutant phenotypes were indeed due to the targeted mutations. The coding sequence of each gene was cloned downstream of the constitutive *ermE** promoter and EF-Tu ribosome binding site of pIJ10257. Upon conjugation, this vector integrates into the ϕBT1 phage attachment site of the *S. coelicolor* M1146 genome; again, antibiotic selection is not required to maintain the construct. Complementation of the Δ*cypA* through Δ*cypM* mutants was verified by MALDI-ToF analysis. Interestingly, in the complemented *cypL* mutant, non-decarboxylated cypemycin was also produced (Figure IV.6.).

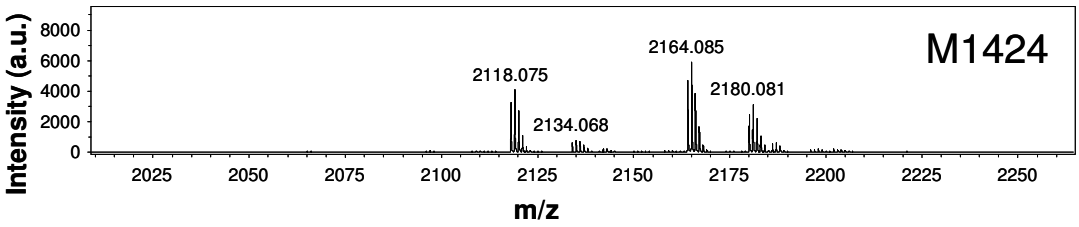


Figure IV.6. Analysis of the complementation of the Δ*cypL* mutant M1424 confirms production of both cypemycin and the non-decarboxylated form.

IV.4. Cypemycin is a member of a larger family of post-translationally modified peptides, the linaridins

A search of the NCBI sequence database identified ten *cypL* homologues. Their occurrence always coincided with the presence, in close proximity, of genes with homology to *cypH* (sometimes just to the 5' half of *cypH*) (Figure IV.7.). Interestingly, in *Streptomyces viridochromogenes*, the *cypL* homologue is present in a cluster with two genes, one of which is homologous to the 5' half of *cypH* while the other to the 3' half. All of these clusters, with the exception of the one in *Solibacter usitatus,* contain a short ORF that could encode a prepropeptide (Table IV.1.). Both *cypL* and *cypH* were essential for cypemycin biosynthesis and were required for introduction of dehydrobutyrine (dehydrated threonine) residues into the mature peptide. Given the rarity of the *cypH* and *cypL* homologues, and their invariant co-occurrence, I propose that the gene clusters in Figure IV.7. represent a new family of post-translationally modified peptides. Given that these peptides are predicted to be linear (lin) and non-cyclised, and to contain dehydrated (arid) amino acids, I propose the name linaridins. Although most of the linaridin gene clusters identified thus far occur in Actinomycetes (Gram-positive bacteria known for their ability to produce a wide range of secondary metabolites), they also occur in other bacterial phyla and even in Archaea (Figure IV.7.).

So far, the occurrence of a decarboxylase (the *cypD* homologue) is restricted to cypemycin and the orthologous cluster from *S. griseus*. SAM-dependent methyltransferases are more prevalent in linaridin gene clusters, indicating that a subset of compounds could be modified by methylation of their N-terminal amino acid. Examples of genes other than those in the cypemycin cluster that could play a potential role in post-translational modification are oxidoreductase-dehydrogenases (SCAB84211 and SSQG_04385), a metallophosphoesterase (FrEUN1DRAFT_0187) and a multi-copper oxidase (Acid_3062). Interestingly, homologous genes encoding a protein of unknown function are present in or immediately adjacent to linaridin gene clusters in *Mycobacterium vanbaalenii* (Mvan_2779), *Solibacter usitatus* (Acid_3059) and *Haloterrigena turkmenica* (two copies: Htur_3023 and Htur_3024; black in Figure IV.7.). The presence of this

homologue only in the linaridin gene clusters suggests a role that is dedicated to the production of modified peptides.

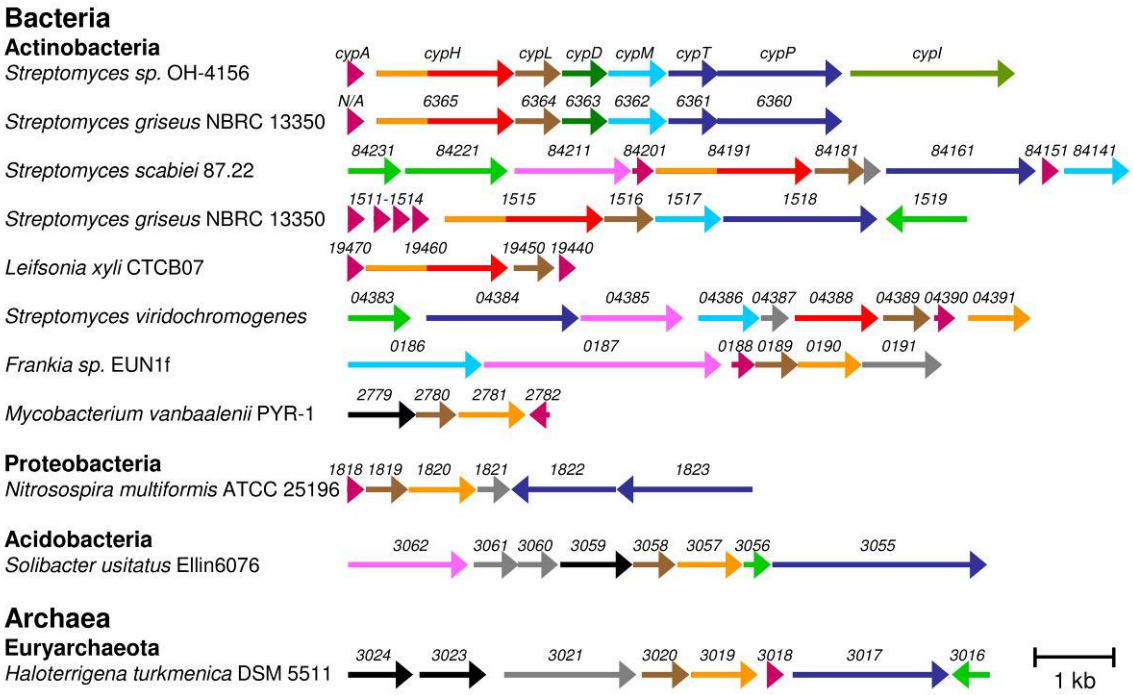


Figure IV.7. The biosynthetic gene clusters of the linaridin family of peptides. Genes are color coded according to the predicted function of their products: structural genes (purple), *cypH* homologues (orange for the N-terminus, red for the C-terminus), *cypL* homologues (brown), decarboxylases (dark green), SAM-dependent methyltransferases (light blue), transporters (dark blue), regulation (light green), catalytic (pink), unknown (grey), unknown but specific to the filaridin clusters (black).

CypA	VR**S**EM**T**L**T**S**T**N**S**AEALAAQDFAN**T**VL**S**AAAPGFHAD<u>C</u>E**T**PAMA**T**PA**T**P**T**VAQFVIQG**ST**I<u>C</u>LV<u>C</u>
SGR_N/A	MRLD**S**IA**T**QE**T**A**T**ALPE**S**MA**T**QDFAN**S**VLAGAVPGFH**S**DAE**T**PAMA**T**PAVAQFVIQG**ST**I<u>C</u>LV<u>C</u>
SCAB_84201	M**SS**IENALN**S**VEIPVEGVVYVAARP**T**LG**T**PRIARIGRIAQAAEGIGAIAAAA**T**AGVGVAQAAEANNLAAEANAQNAAALAAVGGAPA**S**
SCAB_84151	MNDFLLIP**S**VVALGIVGFLIA**T**RAV**ST**PAVIGVALVVLWGA<u>C</u>**S**QAR**ST**VHPRH**TSS**KKRRQHP
SGR_1511	MPEFRQPGW**T**RGVAPLDE**S**AGGQVFGGA**S**PVAA**T**PAVVA**T**AGAVVVAFAAGVAARHLANGGNVELPM
SGR_1512	MA**S**PHRPP**TS**EEK**S**LN**T**A**S**IPAVL**S**E**T**GAL**S**E**S**DHGRALLD**T**VPVA**S**V**T**F**T**M**T**A<u>C</u>VEV**S**V<u>C</u>L**T**G**S**VIQLPQ
SGR_1513	MNK**SS**APAVL**T**A**T**GAL**T**E**S**DAGV**T**LP**T**LAPVAA**T**PVAIAA**T**MGVAFVAGYAAGRAA**T**GNVELPM
SGR_1514	M**S**M**S**P**T**PAALRGAGGL**S**E**S**DPGRAL**SS**LAPV**T**A**T**PGVVAGVALGVALVNAFAAGYNH<u>C</u>GGNVELPM
Lxx19470	V**S**LQR**T**AQAL**S**DLVA**T**ELN**T**DV**T**AG**SS**VPLHEV**T**<u>C</u>LGVIAVVAA**T**AMDIAAYDVA**S**GAAALAV**TT**L**S**V
Lxx19440	M**S**LQK**T**AQ**T**LNDRV**TT**EIDPDA**T**A**STS**VLHREN**T**<u>C</u>IGVV**T**IIVGAAADVAAFDVA**T**GA**S**ALAA**T**ALAI
SvirD4_22614	VPRE**S**AP**S**RR**T**AGELALRI**T**DRKR**TT**M**S**VVADFAN**T**ELADV**T**PGRIGNDA**T**P**T**ML**T**PLAALA**T**PEGVAV**T**AA**T**AYALNEV**T**NDLAG
FrEUN1fDRAFT_0188	M**S**MR**S**EPG**S**LRL**S**QLARIDALI**T**EAQ**S**RGFGL**S**DRFRIHI**T**EEQAAA**T**PDAHHPLFDL**S**EHDREILNQIIEL**T**GQLEH**TTS**IGELVEMRAQVVQG
Mvan_2782	LRVIVIIV**TT**AERFRAAAAHPAA**T**V**S**IAGVRWP**T**YKVV**S**LLVGLGVFGVVAVA**TT**AAAPAVL**S**GAGVA**T**LVWLGLGLYR**TS**RR
Nmul_A1818	MM**T**KLAEAELAGLDAVIEARRM**TS**D**TS**DKYEQIVIPIGNAG**T**AVAGDNAGL**S**GGEGL**S**LGVLMKLRENAL**S**
Htur_3018	M**SS**Q**TT**FGW**S**LF**TS**GIV**T**LVLKALPGD**S**LWWGLMLLAVGLVLLYYR

Table IV.1. Putative structural peptides. Ser and Thr are represented in bold face. Cys residues are underlined. SGR_N/A = gene not annotated.

IV.5. Generation of cypemycin variants

In this section, the importance of particular modified residues for cypemycin activity was investigated. This analysis could also have indicated whether altered substrates were still recognised by the modification enzymes. I opted for *cypA* expression from a high copy number plasmid since only low levels of cypemycin production were obtained in strain M1422 (M1146 containing the integrated Δ*cypA* cypemycin cosmid complemented with *cypA in trans*) (Chapter IV.3.). pIJ86 ($Apra^R$) can be used to express genes from the constitutive *ermE** promoter, but it could be introduced into M1422 since no Apra selection would be possible due to the $Apra^R$ marker present within the integrated construct. Consequently, the *neo* gene from SuperCosI that confers kanamycin resistance was PCR-amplified with primers *neo Nhe*I F and R that contained an extended tail with a *Nhe*I restriction site. This PCR fragment was cloned in the *Nhe*I site of pIJ86, resulting in pIJ12477, an expression plasmid that can be used in a host that is already $Apra^R$.

Three altered *cypA* genes were generated for the production of cypemycin variants (Figure IV.8.). In the first, a synthetic *cypA* gene (Biomatik, Cambridge Ontario), all four Thr codons (Thr2, Thr5, Thr7 and Thr17) were replaced by Ser codons. The construct was cloned in pIJ12477 yielding pIJ12479 (*cypA* 4S). Successful production of this variant could indicate whether CypH/CypL can dehydrate Ser residues and what the effect of this would be on bio-activity. The other two constructs were generated by PCR from *cypA* WT template with a mismatch reverse primer. In the second pIJ12477-based construct, Ser16 codon was replaced by a Thr codon, yielding pIJ12480 (*cypA* S-T). This cypemycin variant could provide information on why Ser16 escapes dehydration. The third *cypA* variant gene contained a Thr codon in place of the Cys19 codon. The construct resulting from the cloning of this gene in pIJ12477 was called pIJ12481 (*cypA* C-T). Should the Thr19 in this variant be dehydrated, this could result in the formation of a C-terminal MeAviCys.

The constructs described above, as well as the vector pIJ12477 and a construct containing the WT *cypA* (pIJ12478), were introduced into M1414 (M1146 containing the integrated Δ*cypA* cypemycin cosmid) by conjugation. The resulting

strains M1459 (pIJ12477 (EV)), M1460 (pIJ12478 (*cypA* WT)), M1461 (pIJ12479 (*cypA* 4S)), M1414 (pIJ12480 (*cypA* S-T)) and M1414 (pIJ12481 (*cypA* C-T)) were grown under cypemycin production conditions with Kan selection to maintain the plasmid. Production of cypemycin variants was analysed by MALDI-ToF MS after $CHCl_3$ extraction of the supernatants. As expected, no cypemycin peaks were observed in the mass spectrum of the empty vector control strain (M1459) and cypemycin was detected for the positive control strain containing the introduced *cypA* WT gene (M1460). No mass peaks were observed in the spectra of M1461 (*cypA* 4S) or M1463 (*cypA* C-T) that could be attributed to a cypemycin variant. This suggests that these variant CypA precursors may be degraded rapidly *in vivo* either because modification of important residues cannot occur (for example if they are no longer accepted as a substrate by the modification enzymes) or because an altered modified residue prevents the correct introduction of later modifications (for example in the formation of a MeAviCys).

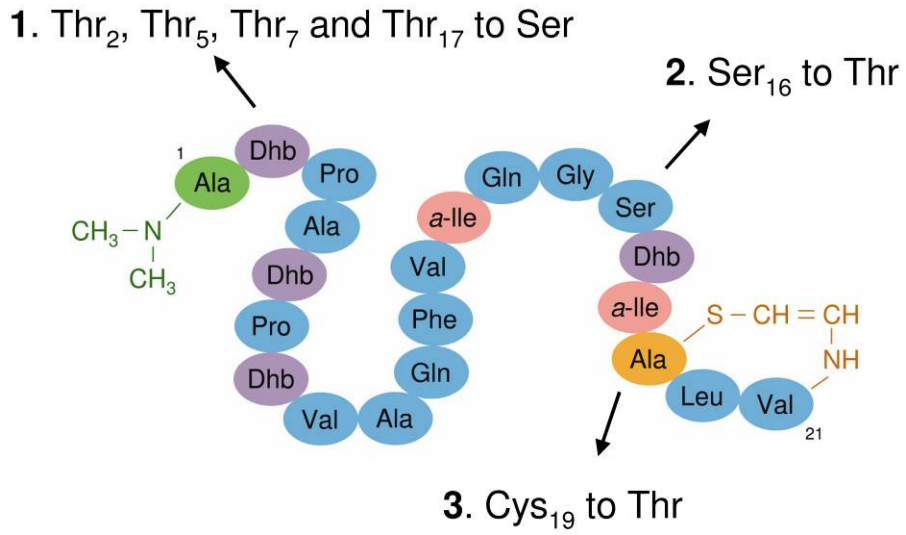
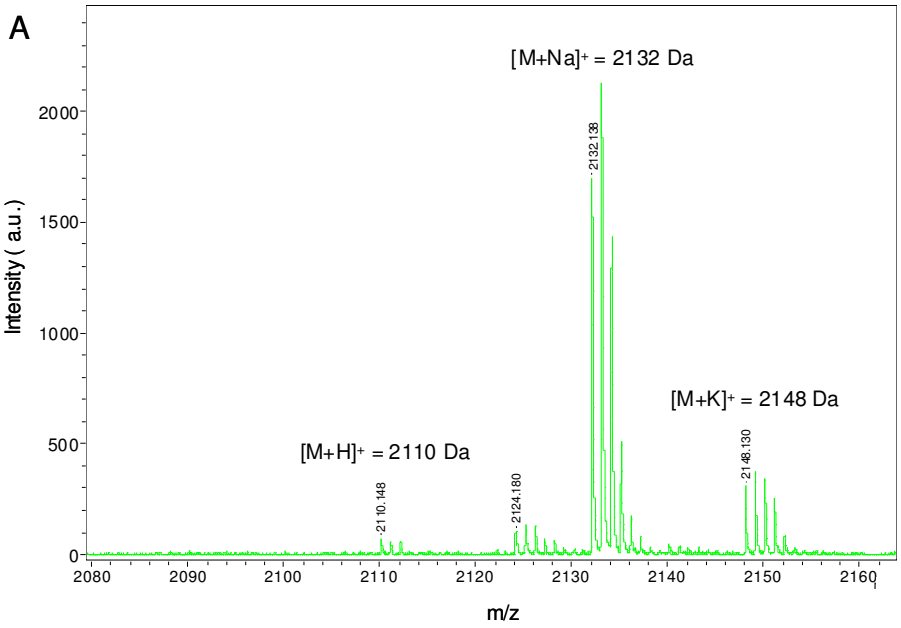


Figure IV.8. Schematic representation of the three cypemycin variants.

MALDI-ToF MS analysis of the M1462 (*cypA* S-T) strain identified a compound with $[M+H]^+$ = 2110 Da, $[M+Na]^+$ = 2132 Da and $[M+K]^+$ = 2148 Da (Figure IV.9.A.). This corresponds to the calculated mass for the S-T cypemycin variant with one of the Thr residues left unmodified. Q-ToF analysis revealed that the Thr16 residue had escaped dehydration and that a Dhb was present at position 17 (Figure IV.9.B.). This indicates that the context of the amino acid residue at position 16 in cypemycin is important for dehydration, regardless of whether it is Ser or Thr. A similar observation has been made with some lantibiotics (Rink *et al.* 2005). Insufficient S-T cypemycin was produced to determine whether the amino acid change had an effect on bio-activity against *M. luteus*.

A



B

#	Residue	b-ions Calculated	b-ions Observed	y-ions Calculated	y-ions Observed
13	*a*-Ile	1271.715	1271.711	952.513	952.526
14	Gln	1399.774	1399.747	839.445	839.445
15	Gly	1456.795	-	711.387	-
16	Thr	1557.843	1557.821	654.365	-
17	Dhb	1640.880	-	553.318	553.295
18	*a*-Ile	1753.964	1753.934	470.280	-

Figure IV.9. Mass spectrometric analysis of M1462 (*cypA* S-T). A) The MALDI-ToF spectrum shows that one Thr remained unmodified in the cypemycin variant. B) Q-ToF analysis shows that Thr16 escaped dehydration and that Thr17 is modified to Dhb.

IV.6. Discussion

I have identified, by genome scanning, the gene cluster for cypemycin biosynthesis and showed, contrary to previous classification, that cypemycin is not a member of the lantibiotic family (Chatterjee *et al.* 2005), but instead represents the first characterized member of a novel family of post-translationally modified peptides. A putative minimal gene cluster was constructed and nine genes contained in a 8.3 kb region were shown to be required for cypemycin production in a heterologous host. Individual in-frame deletions of these genes allowed identification of their functions in post-translational modification and transport.

Cypemycin contains two modifications that are unique to ribosomally synthesized peptides, an N-terminal Me_2-Ala and two L-*allo*-Ileis (Minami *et al.* 1994). Recently, D-*allo*-Ile residues were also identified in the aerucyclamides, a family of cyclic cyanobacterial peptides (Portmann *et al.* 2008). The gene responsible for the isomerization of the side chains of Ile13 and Ile18 to form the L-*allo*-Ile residues present in cypemycin could not be unambiguously identified. CypI is a candidate isomerase and belongs to the DUF255 family of proteins that contain a thioredoxin domain and which show similarity to the N-acetyl-D-glucosamine epimerase superfamily. As yet, I have been unable to obtain enough pure peptide from the Δ*cypI* mutants to verify whether the L-*allo*-Ile modifications are absent.

N,*N*-dimethylation of the N-terminal Ala of cypemycin is unprecedented in ribosomally-produced peptides (McIntosh *et al.* 2009). I unambiguously identified CypM, a SAM-dependent methyltransferase, as the enzyme that introduces these two methyl groups. Non-methylated cypemycin was not active in a *M. luteus* bio-assay.

In addition to these two unique modifications, cypemycin contains structural motifs, Dhb and AviCys, found in many lantibiotics. However, much to our surprise, these are introduced through novel enzyme activities and/or via a pathway distinct from their lantibiotic counterparts. No conventional dehydratase was identified in the cluster, but both *cypH* and *cypL* were required for cypemycin biosynthesis. Our mutational analysis dictates, by a process of elimination, that CypH and/or CypL

must be responsible for dehydration of the Thr residues of cypemycin to Dhb. It is possible that this modification is an early step in biosynthesis and required for the subsequent modifications, hence the absence of other modified forms of the propeptide. Alternatively, the lack of the Dhb residues may lead to rapid protease-mediated degradation. The mature cypemycin molecule contains an unmodified Ser; interestingly, lantibiotic Ser residues are also more likely to escape dehydration than Thr (Rink *et al.* 2005).

The occurrence of Cys at position 19 in the propeptide, rather than the expected Ser (Blaesse *et al.* 2000; Chatterjee *et al.* 2005), initially led us to think that formation of AviCys in cypemycin occurred via a novel mechanism. Analysis of the Δ*cypD* mutant revealed that CypD is solely involved in decarboxylation of the C-terminal Cys22. Alkylation of the peptide produced by the *cypD* mutant with iodoacetamide and subsequent Q-ToF analysis confirmed that Cys19 had been converted to Dha. To our knowledge, such a dethiolation reaction is unprecedented. Since the chemistry behind this reaction is similar to the dehydration of Thr to Dhb, I propose that Cys dethiolation is also catalyzed by CypH and/or CypL (Figure IV.10.).

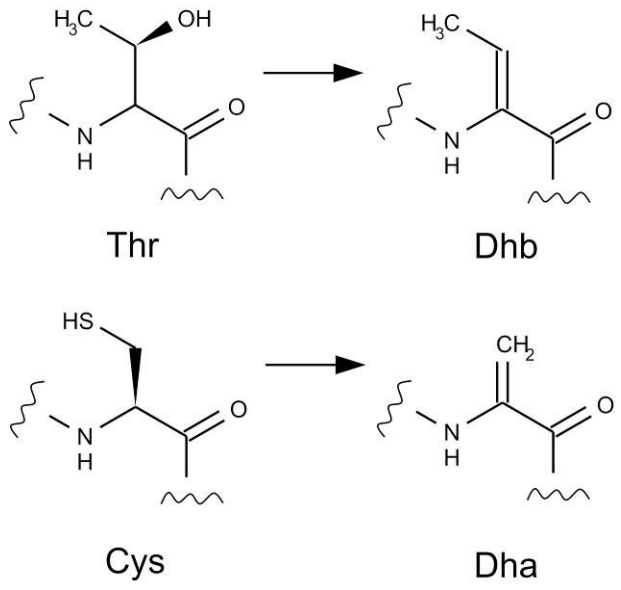


Figure IV.10. CypH and/or CypL are involved in the dehydration of Thr residues to form Dhb and catalyse a similar dethiolation reaction that converts Cys to Dha.

I have shown that CypD decarboxylates the C-terminal Cys in a manner similar to the analogous reaction in lantibiotic biosynthesis, despite its relatively low identity to LanD enzymes. The product of the Δ*cypD* mutant is the first example of the isolation of a natural product lacking its (Me)AviCys residue. This might reflect a general role for AviCys in peptide stability, protecting the modified peptide from degradation by (carboxy)peptidases. AviCys, like the Me_2-Ala residue, is essential for activity against *M. luteus*. A Δ*cypDM* double mutant produced a truncated peptide lacking both modifications. As expected, this compound did not display bio-activity.

The cypemycin leader sequence is removed by an as yet unidentified protease. No putative peptidase domains were identified in either CypT or CypP, and their respective mutants produced lower levels of cypemycin, indicating that they are not essential for modification or cleavage of the peptide, and presumably serve solely to export cypemycin.

Deletion of *orf1* had no effect on cypemycin production, which was perhaps not surprising since SCO4969, its orthologue in *S. coelicolor*, is syntenous with the other genes found upstream of the cypemycin cluster. Thus the cypemycin gene cluster (*cypA* – *cypI*) appears to be contained in a 8.3kb chromosomal fragment (Figure IV.7.). Interestingly, *cypH*, the first gene in the likely biosynthetic operon, contains a rare TTA codon suggesting that cypemycin production might be developmentally regulated and controlled by the tRNA-encoding *bldA* (Leskiw *et al.* 1991). The cypemycin operon appears to be transcribed from a promoter upstream of *cypA*. A predicted transcriptional attenuator (a stem-loop structure with a calculated Δ*G* of -31.2 kCal) lies between *cypA* and the other biosynthetic genes, and may ensure appropriately different levels of production of the modification enzymes and their peptide substrate.

Bioinformatic analysis identified ten additional gene clusters that contain a *cypL* homologue together with a gene with full or partial homology to *cypH* and, with one exception, a gene encoding a putative prepropeptide. These clusters are expected to be capable of producing linear (non-cyclised) dehydrated peptides I call linaridins. The gene clusters are phylogenetically widespread, occurring in different

phyla of bacteria and Archaea, suggesting that they play an important adaptive role in microbial physiology. While cypemycin has antibiotic activity against *M. luteus*, it is inactive against many other bacteria and fungi. This suggests that the primary role of cypemycin, and potentially the linaridin family of peptides, may be to function as extracellular signaling molecules.

Interestingly, some of the putative linaridins from gene clusters with a full length *cypH* homologue contain Cys residues while peptides from clusters with genes homologous to just the 5' end of *cypH* are devoid of Cys (Figure IV.7. & Table IV.1.). It is tempting to speculate that the C-terminal domain of CypH is involved in the dethiolation of Cys. The occurrence of other potential biosynthetic genes (depicted in pink in Figure IV.7.) suggests that other members of the linaridin family contain novel post-translational modifications.

In vitro analysis of the cypemycin modification enzymes as well as the identification and characterization of other members of the linaridin family are ongoing. I anticipate that these peptides will reveal novel post-translational amino acid modifications that will add significantly to the four new biosynthetic enzymes described here, and that they will contribute to an expanding toolbox for post-translational peptide tailoring. The phylogenetically widespread distribution of linaridins and the identification of ten gene clusters in ~2200 sequenced prokaryote genomes, suggests that many more remain to be discovered.

IV.7. Bullet Point Summary

- A reduced gene cluster was constructed and nine genes contained in a 8.3 kb region were shown to be capable of producing cypemycin in a heterologous host.

- Individual in-frame deletions of the genes in the minimal gene set confirmed roles in post-translational modification and transport.

- Mass spectrometric analysis of a non-decarboxylated cypemycin intermediate provided proof for a novel enzymatic conversion: dethiolation of Cys to Dha.

- A bioinformatic search for related gene clusters revealed a novel family of post-translationally modified peptides, the linaridins.

- A cypemycin variant in which the non-dehydrated Ser16 was replaced by a non-dehydrated Thr revealed the importance of sequence context for amino acid dehydration.

Chapter V – Biochemical analysis of cypemycin biosynthesis

V.1. Introduction

Investigation of the enzymatic mechanisms responsible for the post-translational modifications of peptides is a relatively young field of research. The first successful report dates back to 1994, describing the *in vitro* reconstitution of EpiD activity. EpiD is a flavin-dependent decarboxylase involved in the oxidative decarboxylation of the C-terminal cysteine in the lantibiotic epidermin, leading to the formation of the AviCys residue (Kupke *et al.* 1994). Interestingly, the biochemical characterisation of EpiD led to novel insights into the catalytic mechanism of PPC decarboxylases, related enzymes that are involved in coenzyme A biosynthesis (Kupke *et al.* 2000).

The first report of the *in vitro* biosynthesis of a modified peptide followed in 1996 with the bacterial DNA gyrase inhibitor microcin B17 (Li *et al.*). The microcin B17 synthase was purified and consists of three different proteins. This complex introduces four oxazole and four thiazole rings into the microcin B17 prosequence, thereby generating the mature antibiotic.

The *in vitro* biosynthesis of lantibiotics was delayed for a long time because the dehydratase activity could not be reconstituted *in vitro* (van der Donk 2006). Consequently, the *in vitro* formation of Lan and MeLan bridges could not be studied due to the lack of a substrate for the cyclase reaction. The turning point was the analysis of the bifunctional dehydratase/cyclase LctM (Xie *et al.* 2004). This breakthrough paved the path for the biochemical analysis of the LanC enzyme involved in nisin biosynthesis (Li *et al.* 2006).

Lantibiotic biosynthesis has been well characterised and has led to the identification of compounds based on genome sequence information from the producing organisms. Examples include the identification of the two-component lantibiotic haloduracin (McClerren *et al.* 2006) and more recently venezuelin, an

unidentified *Streptomyces* lantipeptide that is generated by a novel type of bifunctional enzyme, LanL (Goto *et al.* 2010). This last example will be more thoroughly discussed in chapter VI of this thesis. Thanks to many efforts, mainly by the van der Donk group, the lantibiotic family is the most extensively biochemically studied family of post-translationally modified peptides.

In this chapter, the *in vitro* analysis of the enzymes involved in cypemycin biosynthesis will be discussed. The generation of suitable substrates will be described, followed by the *in vitro* reconstitution of the CypD and CypM activities. The *in vitro* dehydration of the four Thr residues and the dethiolation of the internal Cys in pro-CypA was also attempted.

V.2. Substrate peptides for the enzymatic reactions

Two different peptide substrates were required to analyse the cypemycin biosynthetic enzymes *in vitro*. The first one, prepro-CypA, was needed because the leader sequence is predicted to be involved in recognition by CypH and/or CypL, the enzymes that introduce the Dhb residues. In lantibiotic biosynthesis, the leader peptide was shown to be important for recognition of the prepropeptide by the dehydration and cyclisation enzymes (Chen *et al.* 2001; Kluskens *et al.* 2005; Chatterjee *et al.* 2006; Li *et al.* 2006; Rink *et al.* 2007). The prepro-CypA substrate can also be used in a CypD assay to investigate the decarboxylation of the C-terminal Cys. EpiD, a related LanD enzyme that was characterised *in vitro,* could use the prepro-sequence of EpiA as a substrate (Kupke *et al.* 1994), although it does not require the leader peptide for activity and oligopeptide substrates were also modified by EpiD (Kupke *et al.* 1995).

A second substrate that lacks the leader sequence (pro-CypA), was required for the *in vitro* methyltransferase activity, since this modification is predicted to occur after leader cleavage *in vivo*. A TEV protease cleavage site cannot be located N-terminal to the pro-CypA sequence in a fusion protein because a Gly or Ser residue would then be left in front of the first Ala upon cleavage. This would interfere with CypM activity, as a free N-terminal Ala is likely required for methylation. Instead, a factor Xa site could be introduced, since cleavage with this

enzyme results in a free N-terminus. Unfortunately, factor Xa is not an efficient protease. When a MBP-EpiD fusion protein was made with an internal factor Xa site, only 3 % of the fusion protein was cleaved after 20 hours of incubation (Kupke *et al.* 1992).

Several methods have been employed to obtain the peptide substrates for the *in vitro* reconstitution of enzyme activities that introduce post-translational modifications (Kupke *et al.* 1993; Majer *et al.* 2002; Xie *et al.* 2002; Xie *et al.* 2004; McClerren *et al.* 2006; Li *et al.* 2009). For the production of substrates for the cypemycin modification enzymes, a Gateway-based vector system was used to yield proteins with different 6His-tagged fusion partners (Busso *et al.* 2005). This technology allows rapid screening of different conditions to obtain a product with good expression levels and solubility. This method was employed not only to generate substrates (pro-CypA and prepro-CypA) for the enzymatic conversions, but also to purify the biosynthetic enzymes themselves (CypH, CypL, CypD, CypM and CypI). All enzyme constructs were generated with a TEV protease site directly N-terminal of the desired enzyme, allowing it to be liberated by enzymatic cleavage after purification. Expression of 6His-MBP fusions resulted in soluble protein for all fusion constructs (Table II.1.). Purification of the fusion proteins was achieved by Ni^{2+}-affinity chromatography (Chapter II.7.1.). SDS-PAGE analysis revealed few contaminating proteins after this one step purification and the yield was reasonable to very good for all expressed substrates and enzymes. The SDS-PAGE analysis for the purification of the 6His-MBP-proCypA fusion protein is shown in Figure V.1.

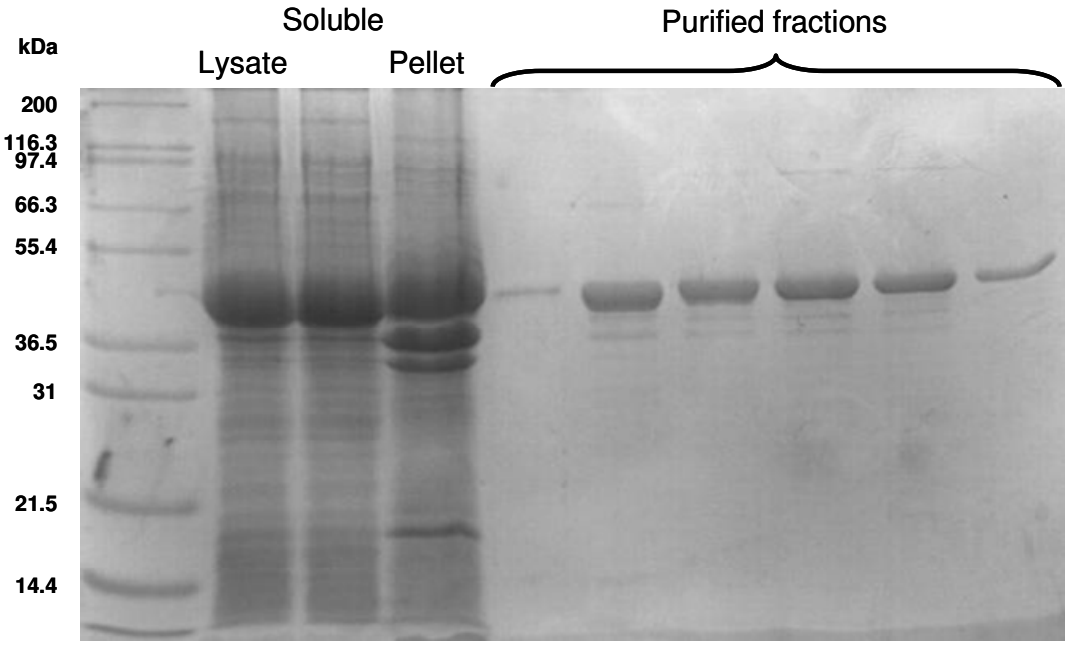


Figure V.1. SDS-PAGE analysis of the 6His-MBP-proCypA fusion protein purification. Purified fractions eluted from the column at imidazole concentrations between 20 – 35 %.

V.3. Formation of the AviCys residue

Genetic analysis of cypemycin biosynthesis unambiguously identified CypD as the decarboxylase involved in the oxidative decarboxylation of the C-terminal Cys (Chapter IV.3.). CypD belongs to the HFCD family of Cys decarboxylases that consists of PPC decarboxylases and LanD enzymes (Kupke *et al.* 2000). When CypD is used as a query in a BLAST search, surprisingly no LanD enzymes are identified. Instead, all hits are annotated as potential PPC decarboxylases, which is surprising, given that the substrate of CypD (a C-terminal Cys residue in a small peptide) resembles the substrates of the LanD enzymes more than that of the PPC decarboxylases ((*R*)-4'-phospho-N-pantothenoylcysteine). This could indicate that CypD has evolved from a common HFCD ancestor in a separate event from the LanD enzymes to perform a similar function.

The structures of three HFCD members - two LanD enzymes (EpiD and MrsD) and one PPC decarboxylase (AtHAL3a) - have been determined (Blaesse *et al.* 2000; Blaesse *et al.* 2003; Steinbacher *et al.* 2003). These studies revealed the reaction mechanism of the oxidative decarboxylation and identified the residues that are important for catalysis and cofactor and substrate binding.

Two striking differences become obvious from an amino acid sequence alignment of CypD (and the closely related GrmD from the grisemycin cluster of *S. griseus*, Chapter VI.2.1.) with all functional LanD enzymes (EpiD, GdmD, BsaD, MutD, MrsD and MicD) and the two biochemically characterised PPC decarboxylases (Dfp and AtHAL3a) (Figure V.2.). Firstly, the essential and strictly conserved active site His is not present in CypD and GrmD, but instead a Glu is found. Mutation of this His residue to Asn in EpiD completely inactivated the enzyme (Kupke *et al.* 2000). A search of the NCBI database for other, functionally uncharacterised HFCD members did not identify any proteins with a different residue at the location of the conserved His. Since Glu has a comparable pKa (4.07 compared to 6.10 for His), the catalytic mechanism could still be similar to the other HFCD enzymes. A second sequence difference is the absence of the PXMNXXMW motif that has been shown to assist cofactor binding and fixes the backbone amide group of the substrate's C-terminal Cys (Blaesse *et al.* 2000).

```
EpiD          -----------MYG------KLLICATASINVININHYIVELKQH-FDEVNILFSPS--S 40
GdmD          -----------MHG------KLLICATASINVVNINHYIVELKQY-FEEVNILFSPS--S 40
BsaD          ------MEAYIMGE------NVLICLCGSVNSINISHYIIELKSK-FDEVNVIASTN--G 45
MutD          ------MEEQNIEK------KILLCLTGSGPLLGIAEYITFLTVR-FKHVRVIISDN--A 45
MrsD          -----MSISILKDK------KLLIGICGSISSVGISSYLLYFKSF-FKEIRVVMTKT--A 46
CypD          -----VNVEKFEGA------ELHVHVTGSISAALVPWWIHWLREF-QPELVVNVSVTPAA 48
GrmD          -----MNVEQFEGA------ELHLHVTGSISAALVPWWIHWLRQL-NPDVVVNVSVSRSA 48
MibD          -----MTAHSDAGGDPRPPERLLLGVSGSVAALNLPAYIYAFRAAGVARLAVVLTPA--A 53
Dfp           --------MSLAGK------KIVLGVSGGIAAYKTPELVRRLRDR-GADVRVAMTEA--A 43
AtHAL3a       MENGKRDRQDMEVNTTPRKPRVLLAASGSVAAIKFGNLCHCFTEWAEVRAVVTKSSLHFL 60
                                   .: :   ..            :         :  :

EpiD          KNFINTDVLKLFCDNLYD-----EIKDPLLNHINIVEN--HEYILVLPASANTINKIANG 93
GdmD          KKFINTDVLNLFCDNLYD-----ETQDPLLNHINIVEN--HEYILVLPASADTINKIASG 93
BsaD          RKFINGEILKQFCDNYYD-----EFEDPFLNHVDIANK--HDKIIILPATSNTINKIANG 98
MutD          AKMLPVAAITQLCEKVYTD--EVSFTDKQKSHIALTRW--ADITVVLPATANIIGKVANG 101
MrsD          EDLIPAHTVSYFCDHVYS---EHGENGKRHSHVEIGRW--ADIYCIIPATANILGQTANG 101
CypD          SRFLAVRALRHLANGKVWV--DSWDDPDVPPEVNSGKSGASECFLVFPATLDTVMRLAQG 106
GrmD          TQFVTVKALRHLANGQVWT--DAWDDPSLPAEVNSGQSGAAECFIVFPATLDTLMRLAQG 106
MibD          EGFLPAGALRPIVDAVHT------EHDQGKGHVALSRW--AQHLLVLPATANLLGCAASG 105
Dfp           KAFITPLSLQAVSGYPVSDSLLDPAAEAAMGHIELGKW--ADLVILAPATADLIARVAAG 101
AtHAL3a       DKLSLPQEVTLYTDEDEWS--SWNKIGDPVLHIELRRW--ADVLVIAPLSANTLGKIAGG 116
                :     :                      .:   .    :   : * : : :   * *

EpiD          ICDNLLTTVCLTGY--QKLFIFPNMNIRMWGNPFLQKNIDLLKNNDVKVYSPDMN--KSF 149
GdmD          ICDNLLTTVCLTGY--KSLYIFPNMNIKMWENPFLQKNIDLLKNNSVKVYPPDVN--KSF 149
BsaD          ICDNLLLTICHTAF--EKLSIFPNMNLRMWENPVTQNNIRLLKDYGVSIYPANIS--ESY 154
MutD          IADNFMTTTLLSSS--KPVLIYPCMNNIMWGNPVVQKNVEVLSGTQYKVIVGQES--ESF 157
MrsD          VAMNLVATTVLAHP--HNTIFFPNMNDLMWNKTVVSRNIEQLRKDGHIVIEPVEI--MAF 157
CypD          RADSPALMMLQLTD--APLVIADTFPG---SNEIVENNVQTLKLRPNVEFAPRVN--GVR 159
GrmD          RADSPALLMLQVTD--RPIVIADTLPG---SNEIVESNLKTLRLRPNIAFAPRVT--GVR 159
MibD          LAPNFLATVLLAAD--CPITFVPAMNPVMWRKPAVRRNVATLRADGHHVVDPLPG--AVY 161
Dfp           MANDLVSTICLATP--APVAVLPAMNQQMYRAAATQHNLEVLASRGLLIWGPDSGSQACG 159
AtHAL3a       LCDNLLTCIIRAWDYTKPLFVAPAMNTLMWNNPFTERHLLSLDELGITLIPPIKKRLACG 176
               . .                .   :            ::  *

EpiD          EISSGRYKNNITMPNI----------ENVLNFVLNNEKRPLD------------------ 181
GdmD          EISSGLYKNSITMPNI----------ENVLSFILNSEKRPLD------------------ 181
BsaD          ELASKTFKKNVVAPEP----------YKVLEFI---------------------------- 177
MutD          ELASGKMKKNIAIPSL----------DELRRVVLENLQEER------------------- 188
MrsD          EIATGTRKPNRGLITP----------DKALLAIEKGFKERTKHPSLT------------- 194
CypD          ASNRQTAEVGFNLPGA----------LAAANRMRKEGRSGE------------------- 190
GrmD          ASNRAAAEVGFNLPGA----------IAVANEMVKKGSVHE------------------- 190
MibD          EAASRSIVEGLAMPRP----------EALVRLLGGGDDGSPAGPAGPVGRAEHVGAVEAV 211
Dfp           DIGPGRMLDPLTIVDMAVAHFSPVNDLKHLNIMITAGPTREPLDPVRYISNHSSGKMGFA 219
AtHAL3a       DYGNGAMAEPSLIYST-------------VRLFWESQAHQQTGGTS-------------- 209
                                                .
```

Figure V.2. Alignment of CypD with other HFCD-family decarboxylases. For the HFCD proteins, the catalytic His and conserved PASANT and PXMNXXMW motifs are boxed and depicted in red. Additional residues involved in cofactor binding are shown in yellow and pututative substrate binding clamps in blue. Sequences shown belong to the following organisms: EpiD from *Staphylococcus epidermis* (P30197.1), GdmD from *Staphylococcus gallinarum* (ABC94905.1), BsaD from *Staphylococcus aureus subsp. aureus* TCH70 (ZP_04829285.1), MutD from *Streptococcus mutans* (AAG48568.1), MrsD from *Bacillus sp.* HIL-Y85/54728 (Q9RC23.1), CypD from *S. sp.* OH-4156 (this study), GrmD from *S. griseus subsp. griseus* NBRC 13350 (YP_001827875.1), MibD from *Microbispora corallina* NRRL 30420 (ADK32557), DfP from *E. coli str.* K-12 *substr.* W3110 (AP_0014152.1) and AtHAL3a from *Arabidopsis thaliana* (NP_188430.1).

The PASANT motif characteristic of HFCD proteins, which supports the flavin cofactor, is conserved as well as other scattered residues that are involved in cofactor binding (indicated in yellow in Figure V.2.) (Blaesse *et al.* 2000). It could not be determined from the alignment whether CypD binds FMN (like EpiD) or FAD (like MrsD) as a cofactor since the affinity for either cofactor could not be attributed to a few distinct residues. Instead, a larger number of small contributions is thought to determine cofactor affinity (Blaesse *et al.* 2003).

All characterised members of the HFCD family contain either FAD or FMN as a redox cofactor (Majer *et al.* 2002). Purified MBP-CypD has a bright yellow color (Figure V.3.A), indicating that the protein is likely to be folded with an associated flavin cofactor. The UV-Vis absorption spectrum of the purified MBP-CypD fusion protein was determined (Figure V.3.B) and the maxima at 377.97 nm and 455.70 nm are characteristic of a flavin-containing enzyme, comparable to MBP-MrsD and MBP-EpiD (Majer *et al.* 2002). The cofactor is not covalently attached to CypD; it could be removed from the enzyme by boiling at 100 °C for 10 minutes and remained in solution after subsequent centrifugation. MALDI-ToF mass spectrometry was employed to identify the cofactor (Figure V.3.C.). The $[M+H]^+$ of the cofactor was 786.171 Da, in good agreement with a calculated $[M+H]^+$ of 786.164 Da for FAD (the $[M+H]^+$ of FMN would be 457.113 Da). This makes CypD the second HFCD enzyme with a FAD cofactor (together with MrsD). All other HFCD enzymes characterised to date contain an FMN cofactor (Majer *et al.* 2002).

The activity of the LanD enzymes EpiD and MrsD and the PPC decarboxylases Dfp and AtHAL3a has been reconstituted *in vitro* (Kupke *et al.* 1994; Kupke *et al.* 2000; Kupke *et al.* 2001; Majer *et al.* 2002). As is the case for the LanD enzymes, the natural substrate of CypD is a peptidyl-Cys. The substrate specificity of EpiD was investigated using oligopeptide libraries and mass spectrometric analysis (Kupke *et al.* 1995). Peptides as short as four amino acid residues were decarboxylated by EpiD and the enzyme exhibited broad substrate specificity. Most peptides with the C-terminal sequence [V/I/L/(M)/F/Y/W]-[A/S/V/T/C/(I/L)]-C could be used as a substrate. Note that the C-terminus of preproCypA (LVC) obeys this rule and would thus likely be modified by EpiD. Interestingly, a Val at the penultimate position (like in preproCypA) was found to give a catalytic reaction rate

that was close to the maximum observed with the natural substrate sequence (Kupke *et al.* 1995).

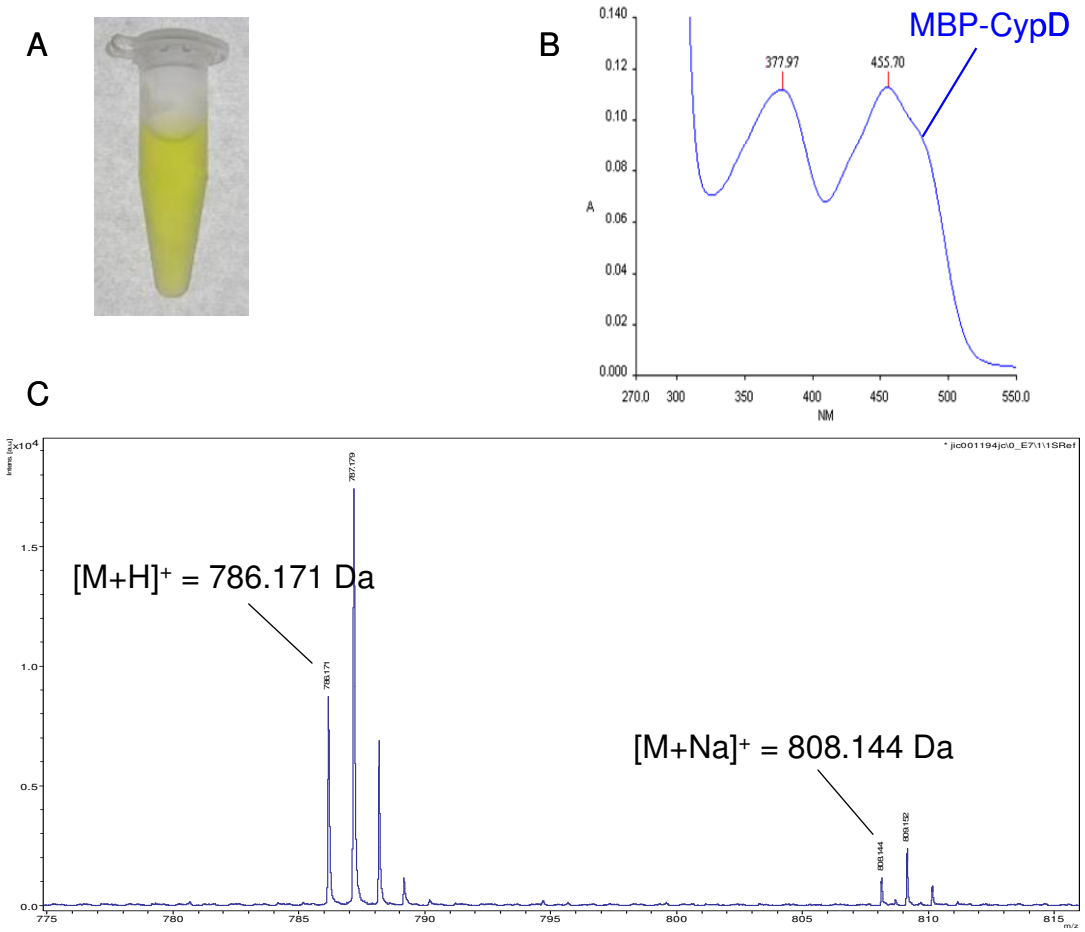


Figure V.3. Characterisation of the CypD cofactor. A) Purified 6His-MBP-CypD has a yellow colour. B) The UV-Vis absorption spectrum for 6His-MBP-CypD has maxima that are characteristic for a flavin-containing enzyme. C) MALDI-ToF spectrum identifies the purified cofactor as FAD.

PreproCypA was used as substrate for the CypD *in vitro* assay, and was obtained by purification of the fusion protein 6His-MBP-preproCypA (Table II.1.) and subsequent cleavage with TEV protease. A maximum efficiency of 60 % cleavage was obtained and the resulting mixture was used as a substrate in the CypD assay without further purification from the MBP and full length fusion protein (Figure V.4.).

6His-MBP-CypD (~ 25 µM) was incubated with the preproCypA substrate (~ 120 µM) at 30 °C in a 20 µl volume at pH 7.9 (50 mM Tris HCl) under reducing conditions (2 mM TCEP). After six hours of incubation, the reaction mixture was analysed by MALDI-ToF MS. The substrate peak with a mass of 6087 Da (as observed in the control reaction without the enzyme) was partially converted into a peak of 6041 Da (Figure V.5.). The 46 Da mass difference is in good agreement with the removal of two hydrogen atoms and one molecule of CO_2. When the pH of the reaction was changed to 7.2, less substrate was decarboxylated (data not shown).

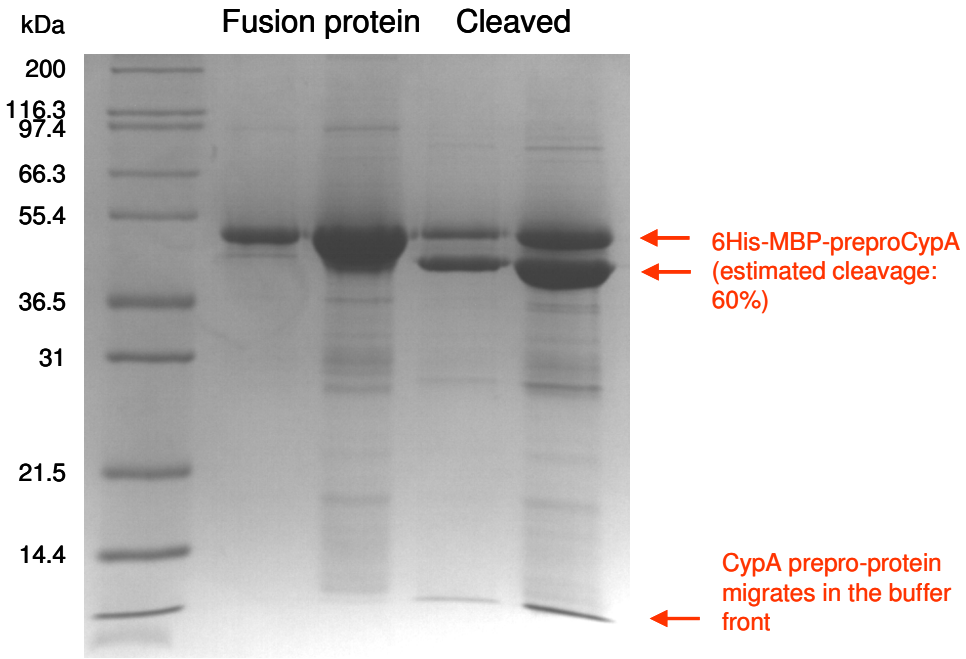


Figure V.4. Substrate preparation for the CypD decarboxylase assay. Cleavage of 6His-MBP-preproCypA with TEV protease occurs at ~ 60 % efficiency. Liberated preproCypA migrates with the buffer front.

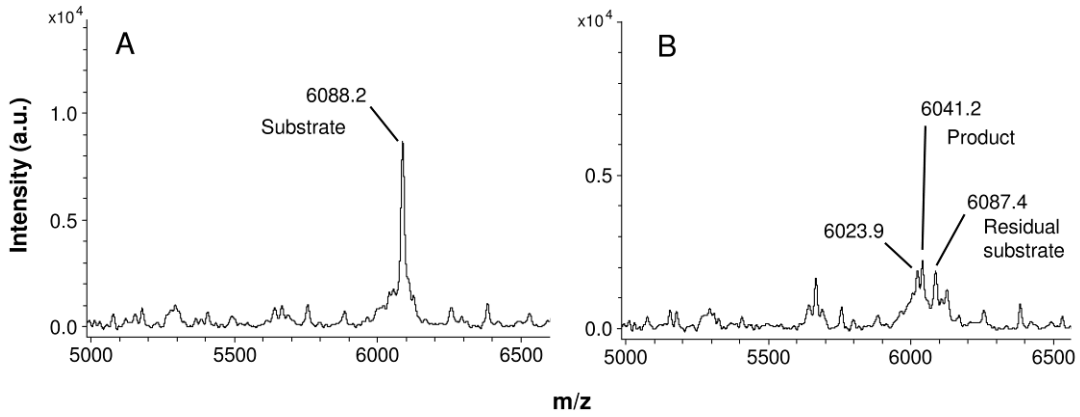


Figure V.5. CypD catalyses the *in vitro* decarboxylation of preproCypA. Upon incubation with the CypD enzyme, the mass peak corresponding to the 6087 Da substrate is partially converted to the 6041 Da decarboxylated product.

V.4. Methylation of the N-terminal alanine residue

cypM was shown to be required for the *N,N*-dimethylation of the N-terminal Ala residue in cypemycin, leading to the formation of Me_2-Ala (Chapter IV.3.). CypM is the first SAM-dependent methyltransferase that has been implicated in the post-translational modification of a ribosomally synthesised peptide. Many examples of peptide bond amide *N*-methylation are known in NRPs and for some of these, the modification confers favourable pharmacokinetic properties on the resulting compounds (Lawen and Zocher 1990; Subtelny *et al.* 2008). A better biochemical understanding of the CypM methyltransferase could thus provide a valuable contribution to an enzyme toolbox for rational peptide design.

CypM is predicted to catalyse the last modification reaction during cypemycin maturation. Removal of the leader peptide liberates the N-terminal Ala that can then be dimethylated on the free amine. Attempts to cleave the purified 6His-MBP-proCypA fusion protein with factor Xa protease resulted in very low proCypA yields because of inefficient protease activity and a small peptide (2.7 kDa) to fusion protein (45.8 kDa) ratio. When proCypA was fused to a 6His-tag, no expression was obtained in *E. coli*. A fusion of proCypA to a smaller partner (thioredoxin, 14.6 kDa) was made in *E. coli*, but nearly all of the fusion protein was found to be insoluble.

The *in vivo* product of M1418 (Δ*cypM*) was shown to share all post-translational modifications except for Me_2-Ala (Chapter IV.3.). Since it is very likely the natural substrate of CypM, this non-methylated version of cypemycin was used as the substrate in the *in vitro* dimethylation assay.

6His-MBP-CypM (~7.5 μM) was incubated with non-methylated cypemycin (unknown concentration) at 30 °C in a 20 μl volume at pH 7.2 (50 mM Tris HCl) under reducing conditions (1 mM TCEP) in the presence of added SAM (2.5 mM). After three hours of incubation, the reaction mixture was analysed by MALDI-ToF MS. The peaks corresponding to the non-methylated substrate ($[M+H]^+$ = 2068 Da, $[M+Na]^+$ = 2090 Da, $[M+K]^+$ = 2106 Da) were mostly converted into mature cypemycin ($[M+Na]^+$ = 2118 Da, $[M+K]^+$ = 2134 Da) by the addition of CypM

(Figure V.6.). A peak was also observed at 2110 Da, corresponding to the $[M+H]^+$ of monomethylated cypemycin. A similar substrate conversion was observed when the reaction was incubated at pH 6.5 or 7.9. While some SAM-dependent methyltransferases require a divalent metal ion for activity (Kozbial and Mushegian 2005), CypM does not, as shown by the reaction conditions above. The addition of 1-10 mM $MgCl_2$ or 1-10 mM $ZnSO_4$ had no apparent effect on the *in vitro* reaction.

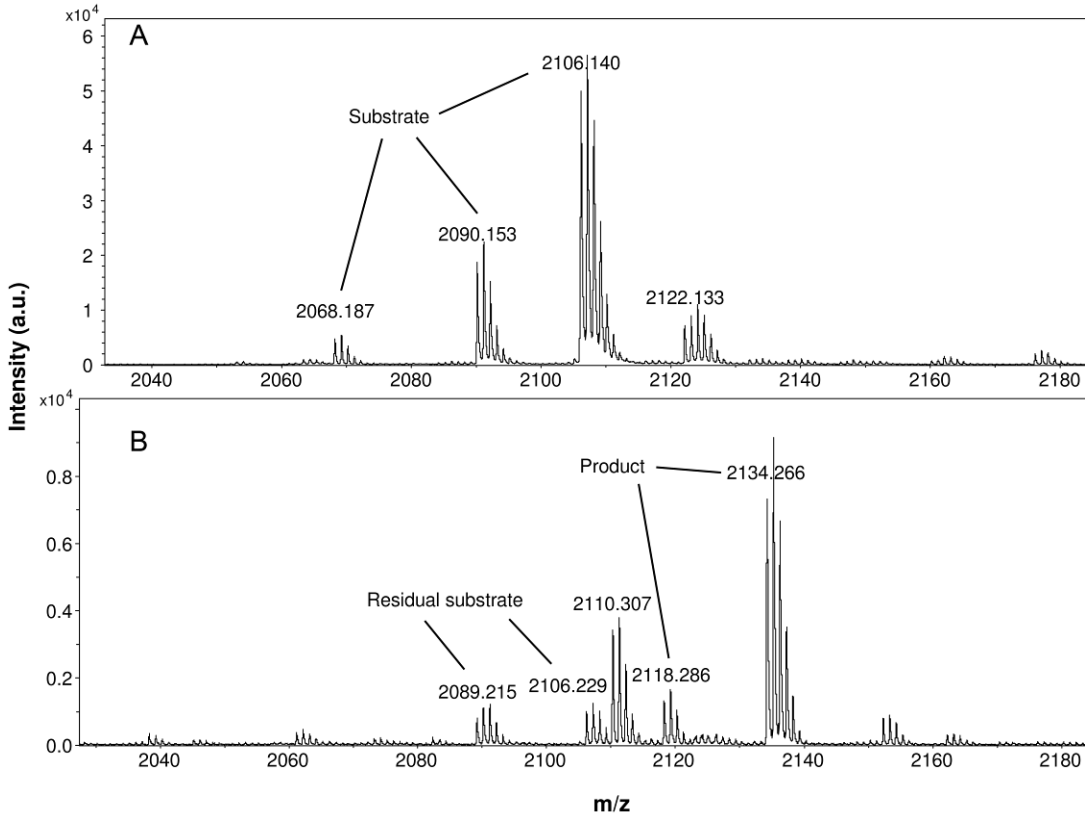


Figure V.6. CypM catalyses the *in vitro* dimethylation of the N-terminal Ala in non-methylated cypemycin. The non-methylated substrate ($[M+H]^+$ = 2068 Da, $[M+Na]^+$ = 2090 Da and $[M+K]^+$ = 2106 Da) was converted to mature cypemycin ($[M+Na]^+$ = 2118 Da and $[M+K]^+$ = 2134 Da) in the CypM methylation assay. Monomethylated cypemycin ($[M+H]^+$ = 2110 Da) was also formed.

V.5. Formation of the Dhb residues

Genetic analysis of the cypemycin biosynthetic gene cluster did not identify a cypemycin intermediate that lacked either the Thr dehydration or the Cys dethiolation modifications (Chapter IV.3.). A possible explanation for this observation is that the intermediates lacking these modifications are unstable and rapidly degraded *in vivo*. Both the Δ*cypH* (M1415) and the Δ*cypL* (M1416) mutants resulted in loss of cypemycin production, making CypH and/or CypL candidates for the dehydratase/dethiolase activity. A BLAST search of the public databases with each as a query failed to identify homologues with an experimentally verified function (Chapter III.6.)

The modification that most resembles the dehydration/dethiolation is the LanB/LanM dehydratase activity that results in the generation of Dha and Dhb residues from Ser and Thr. Unfortunately, the *in vitro* reaction conditions for LanB dehydratases remain to be elucidated. *In vitro* dehydratase activity has been shown for two other types of lantipeptide modification enzyme families, namely LanM (Xie *et al.* 2004) and LanL (Goto *et al.* 2010). The domain that catalyses the dehydration is located in the N-terminal half of the protein in both types. These domains do not share homology with each other or with the LanB enzymes.

The assay conditions used in an attempt to reconstitute CypH and CypL activity *in vitro* were based on methods that have been applied successfully for the LanM and LanL dehydratase activities (Xie *et al.* 2004; Chatterjee *et al.* 2005; Chatterjee *et al.* 2006; McClerren *et al.* 2006; Miller *et al.* 2006; Goto *et al.* 2010). PreproCypA was used as the substrate for the CypH and CypL reactions, since the leader peptide is predicted to be required for molecular recognition by these modification enzymes. The invariant co-occurrence of *cypH* and *cypL* homologs in all linaridin clusters suggests that both could be required together to introduce the modifications. Purified CypH and CypL were thus not only added individually, but also combined in the reaction mixes. Unfortunately, none of the conditions tested resulted in the reconstitution of the dehydration/dethiolation activity.

V.6. Discussion

The *in vitro* characterisation of enzymes that introduce post-translational modifications into ribosomally synthesised peptides may well allow their subsequent use as tools in the design of novel compounds. Post-translational modifications contribute to the rigidity of the peptides and influence their activities (Rink *et al.* 2007; Walsh and Nolan 2008). They also have an effect on stability by increasing the peptide's resistance to proteases. For example, dehydroamino acids have been shown to reduce proteolytic degradation and to enhance biological activity of linear peptides (Lombardi *et al.* 1998; Li *et al.* 2006). A recent study by Rink *et al*. (2010) showed that modification of the N- and C-terminus of pharmaceutical peptides protects them against amino- and carboxypeptidase action. Cypemycin is modified by an N-terminal Me_2-Ala and a C-terminal AviCys (Minami *et al.* 1994). A better biochemical understanding of CypM and CypD, which are responsible for the introduction of these post-translational modifications, could lead to their use in protecting other peptides against peptidases.

The cofactor of CypD was determined and the decarboxylase activity has been reconstituted *in vitro*. CypD is the fifth HFCD decarboxylase for which the *in vitro* activity has been demonstrated and the second one with an associated FAD (Majer *et al.* 2002; Blaesse *et al.* 2003). A BLAST search with CypD as a query only identified PPC decarboxylases (which decarboxylate (*R*)-4'-phospho-N-pantothenoylcysteine) and no LanD enzymes. This suggests a separate evolutionary path for CypD from a common HFCD ancestor to perform a function similar to the LanDs, namely the decarboxylation of a peptidyl-Cys substrate. Unfortunately, the preproCypA substrate used the *in vitro* CypD decarboxylation assay is not easily quantifiable. In order to determine the kinetic parameters for the enzymatic reaction, it might be possible to use shorter peptides with the same amino acid sequence as the C-terminus of CypA as a substrate. For EpiD, peptides as short as four amino acids can be decarboxylated (Kupke *et al.* 1995).

CypM is the first methyltransferase involved in the posttranslational modification of a ribosomally synthesised peptide for which the activity has been reconstituted *in vitro*. The non-methylated cypemycin substrate used in this reaction is not readily

quantifiable. It would be interesting to investigate the substrate specificity of CypM for short peptides that resemble the N-terminus of cypemycin and to determine the kinetic parameters for the dimethylation reaction. SAM-dependent methyl-transferases are quite prevalent in the linaridin clusters (Chapter IV.4.), indicating that cypemycin might not be the only peptide that is *N*-methylated. If a structure could be obtained for CypM, this would shed light on the catalytic mechanism and the substrate binding pocket. Information on the substrate binding pocket would be useful for possible applications where CypM is used to methylate other peptide substrates.

Biochemical reconstitution of the remaining two types of post-translational modification present in cypemycin remain to be established. Determining assay conditions for the dehydration/dethiolation of Thr and the internal Cys in cypemycin is hampered by the lack of CypH and CypL homology to characterised enzymes. Moreover, both enzymes might have to be added together in the reaction mix to obtain activity, which adds an extra level of complexity compared to a one enzyme reaction. Historically, the *in vitro* dehydration reaction in lantibiotic biosynthesis has proven to be notoriously difficult to reconstitute *in vitro* (van der Donk 2006). The dehydratase activity of LanM and LanL enzymes has been characterised (Xie *et al.* 2004; Goto *et al.* 2010), but remains an enigma for the LanB enzymes. The problem in studying the activity of the remaining modification is of a different nature. To study the likely CypI-mediated Ile isomerisation, a suitable assay for the detection of this modification will have to be first devised. This is difficult since no mass difference is associated with the Ile isomerisation. Once a suitable assay has been established, the *in vivo* product of the Δ*cypI* mutant (M1421) could be a candidate substrate for the *in vitro* reaction.

V.7. Bullet point summary

- Peptides were produced to serve as substrates for *in vitro* modification reactions with cypemycin biosynthetic enzymes by producing both as fusion proteins in *E. coli*.

- CypD was shown to decarboxylate preproCypA in an *in vitro* assay.

- *In vitro* activity was shown for the dimethyltransferase CypM, using non-methylated cypemycin as a substrate.

- The Thr dehydratase/Cys dethiolase activity of CypH and/or CypL was not reconstituted *in vitro*.

Chapter VI – Characterisation of grisemycin, a linaridin produced by *S. griseus*

VI.1. Introduction

Characterisation of the cypemycin biosynthetic gene cluster led to the bio-informatic identification of several related clusters, defining the linaridin family of linear post-translationally modified peptides (Chapter IV.4.). A linaridin gene cluster that closely resembles the cypemycin cluster was found in the genome sequence of *S. griseus* IFO 13350 (Ohnishi *et al.* 2008). This 6.4 kb cluster contains seven genes (SGR_6365.2 (not annotated) through SGR_6360) that are presumed orthologs (functionally equivalent homologue) of *cypA* through *cypP*. A *cypI* homologue (SGR_2558) was identified in *S. griseus*, but its genomic location is not adjacent to the cluster. The *S. griseus* cluster is predicted to produce a linaridin, named here as grisemycin. The grisemycin biosynthetic genes (SGR_6365.2 – SGR_6360) will hereafter be referred to as *grmA* through *grmP*, analogous to their counterparts in the cypemycin cluster (Figure VI.1.A.). An apparently orthologous cluster was later found in the draft genome sequence of *S. sp.* ACT-1. Its constituent ORFs share an average nucleotide sequence identity of 99 % with *grmA* through *grmP*.

In this chapter, the identification and structural verification of grisemycin will be discussed. Genes required for grisemycin production were identified by heterologous expression of a minimal gene set in *S. coelicolor* M1146 and by the analysis of a Δ*grmA* mutant of *S. griseus*.

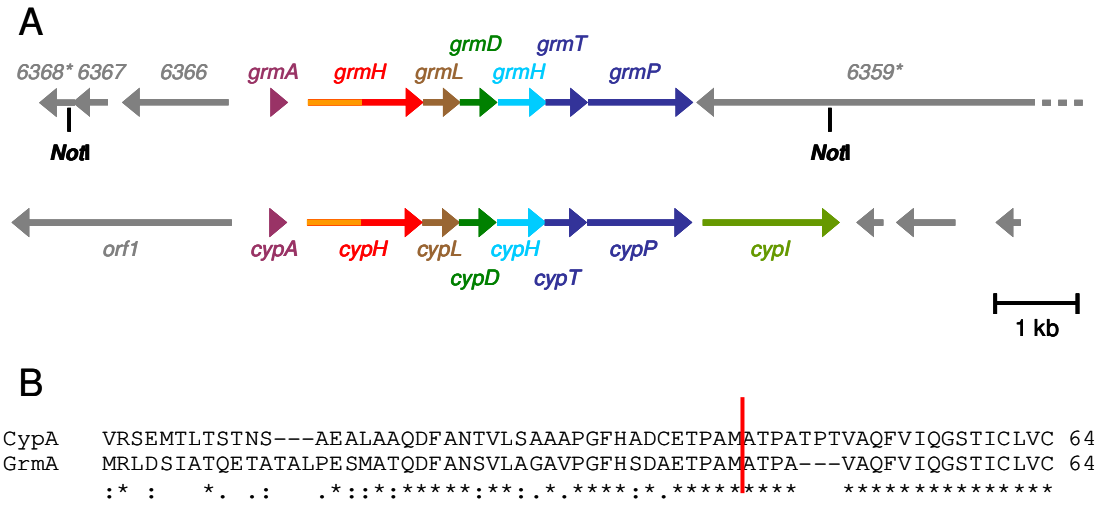


Figure VI.1. A) Comparison of the biosynthetic gene clusters for cypemycin and grisemycin production. The location of the *Not*I sites that flank the grisemycin gene cluster and that were used to construct pIJ12474 are indicated. Genes for which a partial ORF is present in the sub-cloned fragment are indicated with *. Only part of SGR_6359 is shown on the figure. B) Alignment of the precursor peptides GrmA and CypA. The leader peptidase cleavage site for both peptides is indicated by a red line.

VI.2. Identification of grisemycin

VI.2.1. Bio-informatic analysis of the grisemycin gene cluster

The flanking regions on either side of the grisemycin biosynthetic cluster (*grmA* - *grmP*) are not syntenous with the regions adjacent to the cypemycin gene cluster (Ohnishi *et al.* 2008). The grisemycin cluster is flanked by a putative Type I PKS cluster upstream of *grmA* and a cluster of genes of unknown function downstream of *grmP*. Upstream of the cypemycin cluster, genes predicted to be involved in mycothiol detoxification are found and downstream, rodlins and a chaplin are present (Chapter IV.2.). Comparison of these two linaridin gene clusters and their flanking sequences aided in defining the likely boundaries for the cypemycin minimal gene set (Chapter IV.2.).

Directly upstream of *grmA* and divergently transcribed is a gene encoding a putative *Streptomyces* antibiotic regulatory protein (SARP) (Bibb 2005). This putative regulator does not show homology to the product of *orf1* that lies adjacent to the cypemycin gene cluster. Deletion of *orf1* showed that this putative regulator was not involved in the regulation of cypemycin biosynthesis (Chapter IV.3.). Extrapolation leads to the prediction that the putative SARP might not be required for grisemycin production but could instead be involved in regulating expression of the adjacent Type I PKS gene cluster.

Amino acid sequence alignment of GrmA and CypA revealed that the propeptide sequence of GrmA is shorter by three amino acids, Thr-Pro-Thr, corresponding to positions 5-7 of cypemycin (Figure VI.1.B.). All other residues in both propeptides were identical. Comparison of the GrmA and CypA leader sequences also showed a high degree of conservation. Leader sequences of post-translationally modified peptides are proposed to act as a recognition signal for the biosynthetic enzymes (Oman and van der Donk 2010).

Based on the modification enzymes that are encoded by the grisemycin gene cluster, GrmA is predicted to be modified in a similar way to CypA in cypemycin biosynthesis. The two Thr residues in the GrmA propeptide are predicted to be dehydrated to form Dhb, the N-terminal Ala would be di-methylated and an AviCys is predicted to be formed at the C-terminus after dethiolation of the internal Cys and decarboxylation of the C-terminal Cys. Since the *S. griseus cypI* orthologue is not present near the grisemycin cluster, it is not certain whether the Ile residues at position 10 and 15 of the GrmA propeptide are modified to form L-*allo*-Ile.

VI.2.2. Grisemycin identification and structural verification

S. griseus was grown under similar conditions to *S. sp.* OH-4156 for cypemycin production (a three day preculture in SOC medium followed by three days in MarM production medium). Culture supernatant was separated from the mycelium by centrifugation and extracted with $CHCl_3$. The solvent was evaporated and the residual pellet dissolved in 5 % formic acid for MALDI-ToF analysis or MeOH for use in a *M. luteus* paper disc bio-assay. No activity against *M. luteus* was observed

for the *S. griseus* supernatant extracts (data not shown). However, MALDI-ToF analysis showed a set of three peaks with masses $[M+H]^+$ = 1833 Da, $[M+Na]^+$ = 1855 Da and $[M+K]^+$ = 1871 Da (Figure VI.2). If grisemycin is modified in a similar way to cypemycin, this would mean that the compound is shorter by the three N-terminal residues Dhb-Pro-Dhb. The difference in mass between the two compounds would thus be 263 Da, which is in perfect agreement with the mass peaks observed for cypemycin ($[M+H]^+$ = 2096 Da, $[M+Na]^+$ = 2118 Da and $[M+K]^+$ = 2134 Da, Chapter III.2.2.).

The identification of grisemycin by MALDI-ToF in the same *S. griseus* culture supernatant that was used for the *M. luteus* bio-assay suggests that the compound does not possess potent antibacterial activity. Even when the residue from the $CHCl_3$ extract was concentrated or dissolved in different solvents, it failed to produce a zone of inhibition in the *M. luteus* bio-assay.

To verify that the compound observed in the MALDI-ToF spectrum was indeed grisemycin, the amino acid sequence of the mass peak $[M+K]^+$ = 1871 Da was determined by Q-ToF mass spectrometry (Figure VI.3.). The analysis showed that both Thr residues had been dehydrated to Dhb. N-terminal di-methylation of grisemycin was inferred from the 183.12 Da b-ion, which corresponds to Me_2-Ala-Dhb (in good agreement with a calculated mass of 183.11 Da) and the 1688.85 Da y-ion (calculated mass 1688.86 Da). The presence of the C-terminal AviCys residue was inferred from the 395.16 Da y-ion (calculated mass 395.15 Da). No mass peak was observed corresponding to the calculated mass for a b-ion from the grisemycin fragment lacking its AviCys. This is probably due to a progressively decreasing signal intensity of the b-ion peaks at higher masses in the grisemycin Q-ToF spectrum. At present, it is uncertain whether grisemycin contains L-*allo*-Ile residues, since this modification can not be identified by mass-spectrometric analysis.

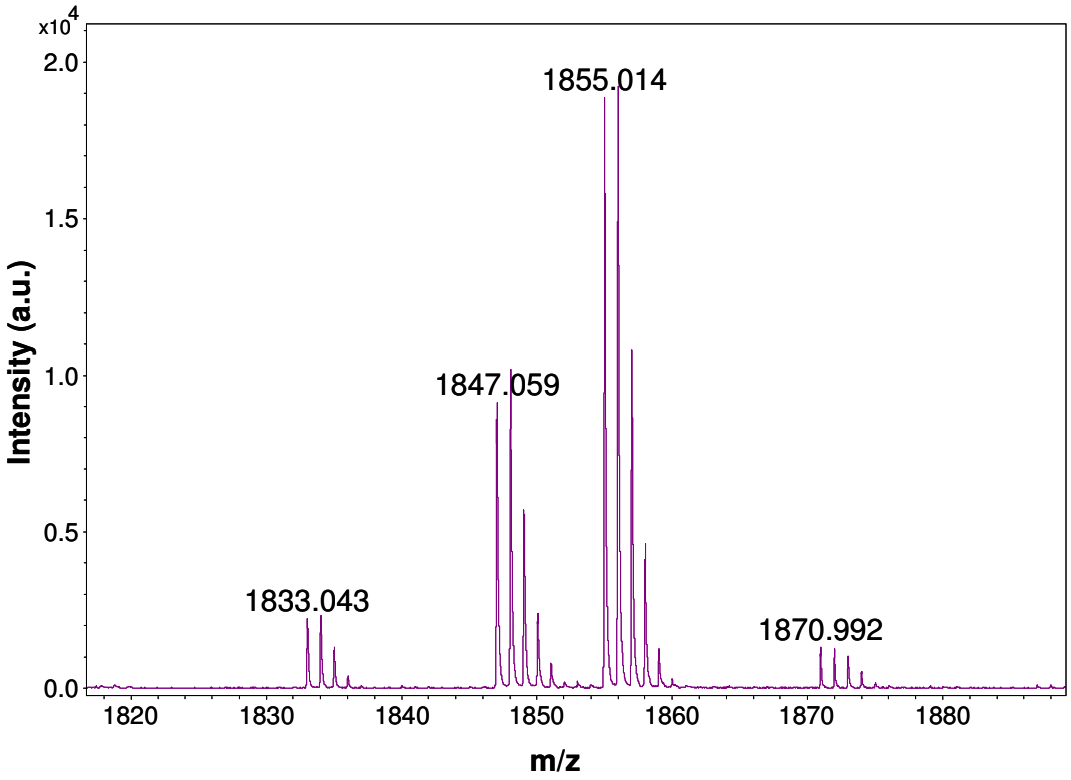


Figure VI.2. MALDI-ToF MS spectrum of grisemycin ($[M+H]^+$ = 1833 Da, $[M+Na]^+$ = 1855 Da and $[M+K]^+$ = 1871 Da) produced by *S. griseus*.

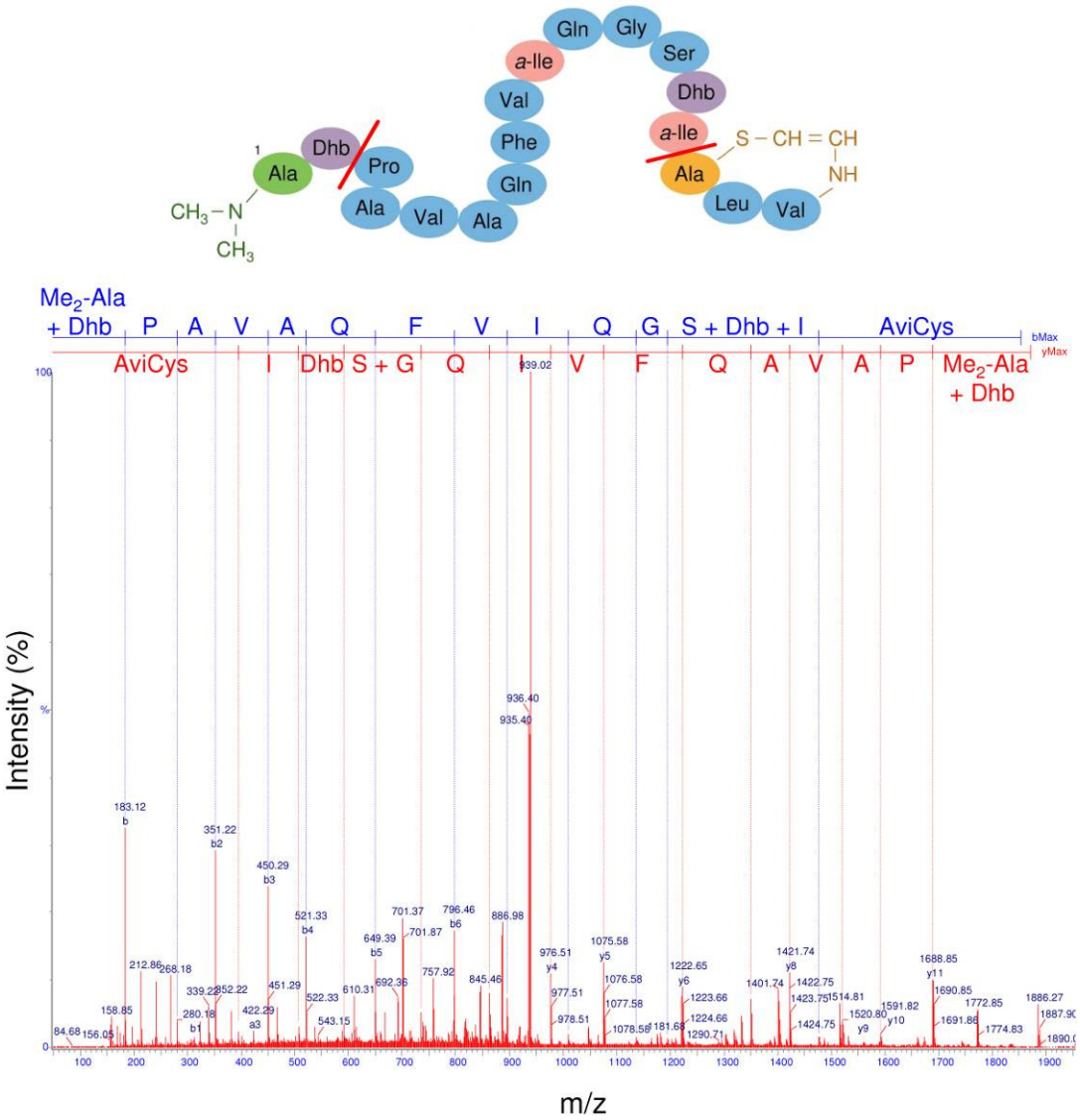


Figure VI.3. Q-ToF spectrum of grisemycin. Fragmentation of the modified linear peptide allows determination of the amino acid sequence, and the nature and locations of modified residues. The amino acid sequence between Pro3 and the AviCys residue (which does not fragment into easily interpretable masses) can be readily discerned and is shown above the spectrum. Note: It has not been determined whether grisemycin contains the L-*allo*-Ile residues at positions 10 and 15.

VI.3. Cloning and heterologous expression of the cluster and generation of a Δ*grmA* mutant

VI.3.1. Cloning of the gene cluster and heterologous expression in M1146

The likely boundaries of the minimal gene set for grisemycin biosynthesis were determined by comparison with the cypemycin gene cluster (Figure VI.1.A. and Chapter III.12.A.). The putative SARP-family regulator situated upstream of *grmA* was included in this construct to account for the possibility that it might be required for regulation of grisemycin biosynthesis. Sequence analysis identified a *Not*I restriction site on either side of the putative grisemycin gene cluster. Excision of the 10.6 kb *Not*I fragment would result in truncation of ORFs SGR_6359 and SGR_6368, leaving only one full length ORF encoding a putative oxidoreductase (SGR_6367) in the minimal gene set construct that was highly likely to be superfluous for grisemycin production (Figure VI.1.A.). Since no *S. griseus* cosmid library was available, genomic DNA was isolated and digested with *Not*I. The DNA was separated on a 0.8 % agarose gel and fragments excised between approximately 10.1 kb and 11.2 kb. The fragments were purified, ligated with *Not*I-digested pSET152 and introduced in *E. coli* DH5α by transformation. 56 DH5α transformants were analysed by colony PCR using primers *grmA* T1 and *grmA* T2, which amplify a 960 bp region containing the *grmA* structural gene. Two DH5α transformants were identified that appeared to contain the grisemycin minimal gene set cloned in pSET152 (hereafter referred to as pIJ12474). The constructs were isolated and verified by restriction digest analysis and by two different diagnostic PCRs. pIJ12474 was introduced into the *S. coelicolor* M1146 by conjugation, yielding strain M1457. Heterologous production of grisemycin was confirmed by MALDI-ToF analysis of $CHCl_3$ extracted supernatant from a MarM production culture of M1457 (Figure VI.4.A.). This indicates that the minimal gene set construct pIJ12474 (Figure VI.1.A.) contains all of the genes required for grisemycin production.

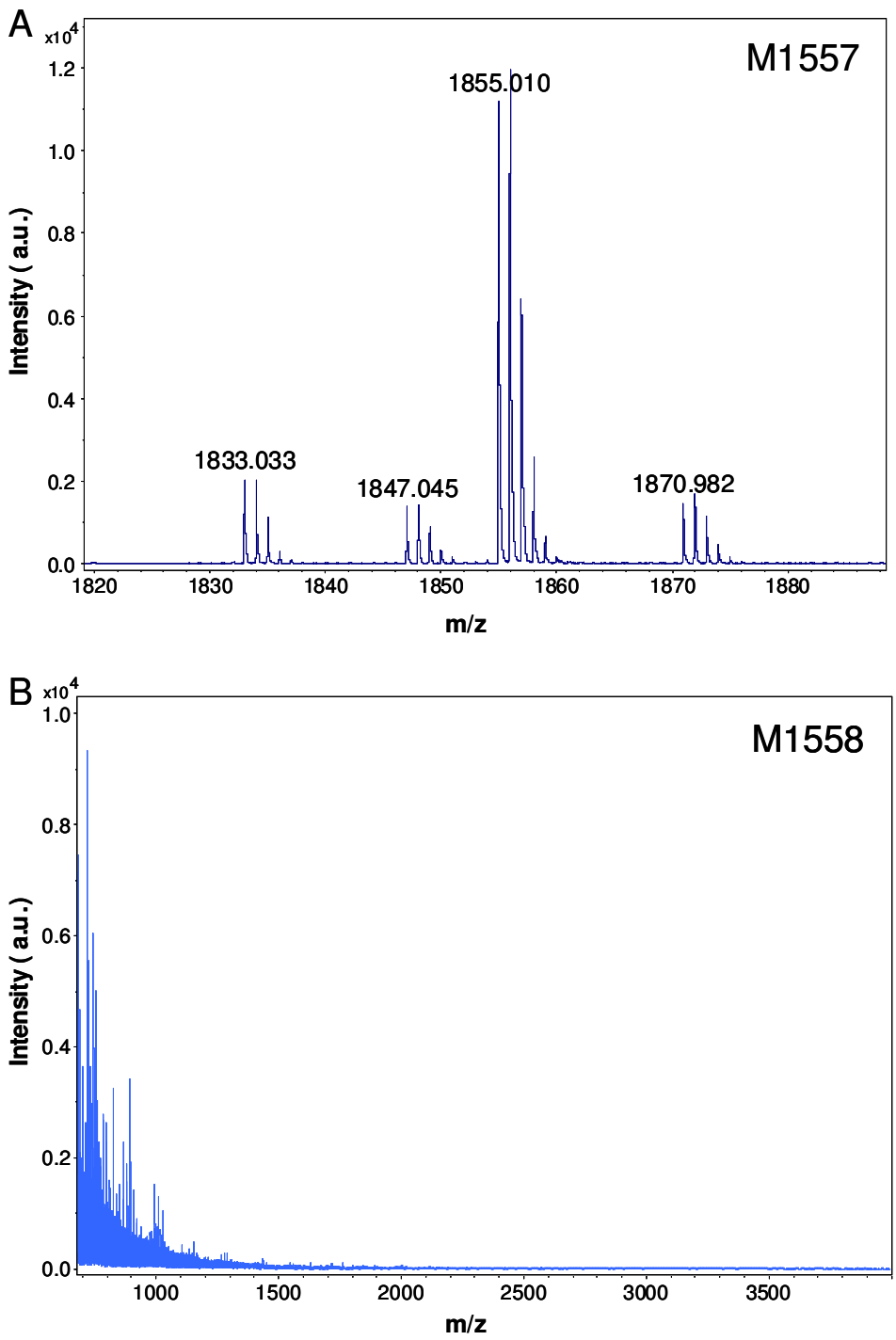


Figure VI.4. MALDI-ToF MS spectra for A) M1457, the heterologously expressed grisemycin minimal gene cluster in *S. coelicolor* M1146 and B) M1458, the *S. griseus* Δ*grmA* deletion strain.

VI.3.2. Generation of a *grmA* deletion mutant

pIJ12474 could not be used to generate a Δ*grmA* deletion construct since its pSET152 backbone would cause it to integrate into the ϕC31 attachment site upon conjugation into *S. coelicolor*. pIJ12474 was digested with *Not*I to liberate the 10.6 kb insert, which was subsequently cloned into *Not*I-digested SuperCosI, yielding pIJ12475. *grmA* was replaced in pIJ12475 with an Apra-resistance cassette by PCR targeting (Chapter II.6.). The resulting construct, pIJ12476, was introduced into *S. griseus* by conjugation. Double homologous recombinants were identified by their $Apra^R$-Kan^S phenotype and verified by PCR analysis. The resulting Δ*grmA* strain (M1458) was grown in grisemycin production medium and the culture supernatant extracted with $CHCl_3$. MALDI-ToF analysis confirmed that the $[M+H]^+$ = 1833 Da, $[M+Na]^+$ = 1855 Da or $[M+K]^+$ = 1871 Da mass peaks were not present in the spectrum, indicating that grisemycin was no longer produced (Figures VI.2 and VI.4.B.). This experiment confirmed that *grmA* is the structural gene for grisemycin biosynthesis in *S. griseus*.

VI.4. Discussion

Several linaridin gene clusters were identified by bioinformatic analysis (Chapter IV.4.). The linaridin gene cluster most closely related to the cypemycin cluster was found in *S. griseus*. The pro-sequence of the structural peptide only differs from cypemycin by the absence of three amino acids, Thr-Pro-Thr, corresponding to positions 5 to 7 in cypemycin. The compound produced by the *S. griseus* cluster was identified by MALDI-ToF and Q-ToF analyses and was called grisemycin.

It remains to be determined whether grisemycin contains L-*allo*-Ile residues. Apart from L-*allo*-Ile, all modifications present in cypemycin were also identified in grisemycin by mass spectrometric analysis. A reduced gene set construct was generated and used to produce grisemycin heterologously in *S. coeilicolor* M1146. To complement this data, a *S. griseus* Δ*grmA* mutant failed to produce grisemycin. Taken together, these data identify the gene cluster that is required for grisemycin biosynthesis in *S. griseus*.

No antibacterial activity could be shown for grisemycin in a bio-assay with *M. luteus*. Although there is no quantitative data on levels of production from the two gene clusters, this might indicate that the antibacterial activity of cypemycin against *M. luteus* is dependent on peptide length, since this is the most obvious structural difference between the two compounds. The mechanism of action of cypemycin is not known, but it could exert its activity by forming pores in the bacterial cytoplasmic membrane. A shorter peptide like grisemycin might not be long enough to cross the phospholipid bilayer. A study on the *in vivo* mode of action of pore-forming lantibiotics showed that the pore-forming ability of gallidermin depended on membrane thickness (Bonelli *et al.* 2006). It was speculated that the greater susceptibility to pore formation of micrococcal strains compared to lactobacilli could reflect in part their relatively thinner phospholipid bilayer. Bonelli *et. al* (2006) pointed out that the membranes of lactobacilli contain considerable amounts of phospholipids with an average chain length of 17 carbon atoms (In 't Veld *et al.* 1991), whereas in micrococcal membranes the most predominant phospholipid has a C_{15} acyl chain (Welby and Tocanne 1982).

The lack of antimicrobial activity observed for grisemycin and the very narrow spectrum reported for cypemycin (Komiyama *et al.* 1993) suggests that these two compounds might serve a different role in the biology of their producing streptomycetes, for example as signalling molecules.

The grisemycin cluster was shown to be highly activated by the bacterial γ-butyrolactone hormone A-factor (Hara *et al.* 2009). This led us to investigate the DNA sequence upstream of *grmA* (and *cypA*) for AdpA binding sites. AdpA is a transcriptional regulator of the AraC/XylS family and *adpA* is the sole target of the A-factor receptor protein ArpA (Kato *et al.* 2004). AdpA thus acts as the master regulator downstream of A-factor signaling and transcriptionally activates many genes involved in secondary metabolism and morphological differentiation (Ohnishi *et al.* 2005).

Putative AdpA binding sites were identified in the DNA sequence upstream of *grmA* (5'-TTGCGGGATT-3') and *cypA* (5'-TGGCCGGATG-3') that are in good agreement with the consensus sequence 5'-TGGCSNGWWY-3' determined by

Yamazaki *et al.* (2004). It would be interesting to experimentally confirm the regulatory role of AdpA in grisemycin and cypemycin biosynthesis.

Interestingly, the rare leucine codon UUA is found in the *cypH* transcript (no corresponding UUA codon is present in the mRNA of the *S. griseus* orthologue *grmH*, Figure VI.5.). This suggests that cypemycin biosynthesis could be under the regulatory control of *bldA*, which encodes the tRNA for the rare leucine codon UUA (Leskiw *et al.* 1991). Interestingly, in *S. coelicolor*, a species that does not produce A-factor, one of the *bldA* targets is its *adpA* orthologue (Takano *et al.* 2003). This suggests that if *S. sp.* OH-4156 does not have a γ-butyrolactone signaling system, cypemycin production could still be developmentally regulated not only through an effect from its *bldA* on its *adpA* orthologues, but also directly because of the UUA codon present in the *cypH* transcript.

```
                Q   G   T   L   R   H   R   I   G   R   V   L   D   H
cypH   181   CAA GGC ACA TTA CGA CAC CGG ATA GGC CGC GTA CTC GAC CAT   222
grmH   181   CCG GGA AGT CTG CGA CAC CGG CTC GGC CGC ACG CTC GAC CAT   222
             *   **  *    *  *** *** ***  *  *** ***     *** *** ***
                P   G   S   L   R   H   R   L   G   R   T   L   D   H
```

Figure VI.5. Alignment of the partial nucleotide sequences of *cypH* and *grmH. cypH* contains a rare UUA leucine codon (depicted in red), while the *S. griseus* orthologue *grmH* does not. The Leu residue is conserved in both proteins (bold in the partial acid sequences for CypH and GrmH above and below the alignment respectively).

VI.5. Bullet point summary

- A second member of the linaridin family was identified in *S. griseus* and called grisemycin.

- The propetide sequence of grisemycin closely resembles that of cypemycin, but is shorter by three amino acids.

- Grisemycin was isolated from a *S. griseus* culture and its predicted structure verified by MALDI-ToF and Q-ToF mass spectrometry.

- The functionality of the grisemycin biosynthetic gene cluster was shown by cloning and heterologous expression in *S. coeilicolor* M1146 and confirmed by generating a Δ*grmA* mutant in *S. griseus*.

Chapter VII - Cryptic lantipeptide gene clusters from streptomycetes

VII.1. Introduction

Among several different strains of *S. venezuelae* described in the literature, the two most well characterised ones are the original chloramphenicol producer ATCC10712 (or ISP5230) (Ehrlich *et al.* 1948) and strain ATCC15439, which is studied for its production of the macrolide antibiotics methymycin, neomethymycin, narbomycin and pikromycin (Xue *et al.* 1998). From here on, only the former strain will be discussed.

S. venezuelae ATCC10712 was first isolated over 60 years ago from a soil sample taken near Caracas, Venezuela (Ehrlich *et al.* 1948). When grown on solid agar medium, *S. venezuelae* develops relatively quickly compared to other streptomycetes such as *S. coelicolor*. Its mycelium has fewer branches and is generally straight or slightly curved but no spiralling occurs (Ehrlich *et al.* 1948). Sporulating *S. venezuelae* colonies have a dark green colour and liquid grown cultures also form spores in some media, an ability that *S. coelicolor* lacks.

The production of “chloromycetin”, the old name for chloramphenicol, has long been the main reason of interest in *S. venezuelae*. The antibiotic was introduced into clinical practice in 1948 and would later become the first antibiotic to be produced synthetically at large scale (McGhee and Anastas 1996). Its use as a broad spectrum antibiotic is currently in decline because of resistance problems, but it is still used as topical treatment for eye infections (Wareham and Wilson 2002).

The biosynthesis of chloramphenicol is linked to the shikimate pathway that assembles aromatic metabolites. It branches off this primary metabolic pathway at chorismic acid with the generation of *p*-aminophenylalanine (PAPA) (Teng *et al.* 1985). The chloramphenicol biosynthetic gene cluster has been identified and cloned, and gene disruption studies have identified key enzymes (He *et al.* 2001).

S. venezuelae also produces a group of polyketide-derived antibiotics, the jadomycins (Doull *et al.* 1993; Doull *et al.* 1994). These angucyclines, of which jadomycin B is the main component, are generally produced only under certain environmental stress conditions, such as phage infection, and heat or ethanol shock (Doull *et al.* 1993; Doull *et al.* 1994). This observation fuelled an interest in the regulation of jadomycin biosynthesis. *jadR2* was identified as a repressor gene that negatively influences jadomycin production. Disruption of *jadR2* uncoupled jadomycin production from environmental stress (Yang *et al.* 1995). A more thorough analysis of the regulatory mechanism showed that $jadR_2$ encodes a γ-butyrolactone receptor homologue that has been hypothesized to perceive a signal produced by $jadW_1$, an *afsA* homologue (Wang and Vining 2003). Inactivation of $jadW_1$ not only affected the growth rate and ability to sporulate of *S. venezuelae*, its disruption also resulted in the loss of both jadomycin and chloramphenicol production (Wang and Vining 2003). In a recent study, *S. venezuelae* was engineered to produce jadomycin without the need for stress induction by replacing the pathway specific regulators and the promoter upstream of *jadJ* – the first gene in the biosynthetic operon of the jadomycin gene cluster – with the constitutive promoter *ermE** (Zheng *et al.* 2007). This resulted in a two-fold increase in production compared to a stress-induced wild-type strain.

The genome sequence of *S. venezuelae* ATCC10712 was determined by Diversa Corp. (now Verenium Corp.) as part of an effort to create a generic strain for industrial antibiotic production (Mervyn Bibb and Diversa Corp., unpublished data). The genome sequence contains several additional gene clusters that are predicted to be involved in the production of secondary metabolites, including lantipeptides and a thiopeptide.

This chapter starts with a general bioinformatic analysis of all cryptic lantipeptide gene clusters that have been identified in sequenced streptomycetes. After this, attention will return to *S. venezuelae*. The expression of secondary metabolite gene clusters in *S. venezuelae* will be discussed in the context of bio-assays for putative lantipeptides made by the strain. Various approaches will be described that were taken with the aim of identifying the products of cryptic lantipeptide gene clusters in *S. venezuelae*. Finally, the *in vitro* approach of collaborators that led to

the identification of venezuelin, a cryptic *S. venezuelae* lantipeptide, will be discussed.

VII.2. Bioinformatic analysis of cryptic lantipeptide gene clusters in sequenced streptomycetes

At the time this PhD project started, the full genome sequence of four streptomycetes had been determined. The genome of the model strain *S. coelicolor* was sequenced in 2002 (Bentley *et al.*), followed shortly by the avermectin producer *S. avermitilis* (Ikeda *et al.* 2003). The genome of potato pathogen *S. scabies* was in a late stage of annotation and the *S. venezuelae* genome had been sequenced by a private company (Diversa Corp.). During the course of the PhD project, the genome sequence of *S. griseus* was completed (Ohnishi *et al.* 2008) and the Broad Institute (Cambridge, MA) provided sequence data from 17 additional *Streptomyces* species.

Several cryptic lantipeptide gene clusters were identified in a bioinformatic analysis of the available *Streptomyces* genome sequences. The data from this analysis are summarised in Table VII.1. It is immediately apparent that lantipeptide gene clusters are widespread and abundant in streptomycetes. All biosynthetic enzymes were examined for the presence of essential catalytic residues and clusters that were likely to be inactive because of a defunct enzyme or the lack of LanA were omitted unless otherwise stated. The clusters from both *S. roseosporus* strains contained identical LanA peptides and were only considered once. The same is true for the orthologous LanBC clusters SCO6930 from *S. coelicolor* and SSPG_07328 from *S. lividans*. Interestingly, no orthologous SCO0269 cluster could be found in *S. lividans*, possibly because of its location near the end of the chromosome. Clusters encoding predicted SapB-like morphogens that are involved in the erection of aerial hyphae during development (Kodani *et al.* 2004; Kodani *et al.* 2005) are present in all strains and will not be discussed further in this analysis. In five strains (*S. avermitilis*, *S. sviceus*, *S. pristinaspiralis*, *S. sp.* AA4 and *S. sp.* SPB78), no additional lantipeptide clusters were identified apart from the SapB/RamC orthologous cluster (Chapter I.4.3.c.) and these strains have been omitted from Table VII.1.

Table VII.1. Lantipeptide synthetases of the LanB, LanC, LanM and LanL types identified in streptomycete genomes. Enzymes that are predicted to be non-functional are indicated by *. Clusters for which no *lanA* could be identified are indicated by (-A). A gene cluster containing a functional *lanC* and partial *lanM*, but no *lanA,* is indicated by (MC).

	LanB	**LanC**	**LanM**	**LanL**
S. coelicolor A3(2)	SCO0269	SCO0270		
	SCO6930	SCO6929		
S. venezuelae ISP-5230			SMD01140*	VenL
S. scabies 87.22	SCAB32031	SCAB32041		
S. griseus IFO 13350	SGR_3847	SGR_3846	SGR_4809*	SGR_152 (-A)
		SGR_6574 (MC)	SGR_6571 (MC)	
S. albus J1074			SSHG_00049	
S. roseosporus NRRL 15998			SSGG_02013*	
			SSGG_04021	
S. clavuligerus ATCC 27064	SSCG_01497	SSCG_01498		SSCG_02809
	SSCG_03317	SSCG_03318		
S. hygroscopicus ATCC 53653			SSOG_05509	
S. sp. C			SSNG_01013*	
S. viridochromogenes DSM 40736	SSQG_04548	SSQG_04549	SSQG_04194	
			SSQG_04196	
S. sp. SPB74	SSBG_01042	SSBG_01043		
S. sp. Mg1	SSAG_01528	SSAG_01529		
	SSAG_03540 (-A)	SSAG_03541 (-A)		
	SSAG_05771*	SSAG_07472*		
S. lividans TK24	SSPG_07328	SSPG_07327		
S. griseoflavus Tu4000	SSRG_00125 (-A)	SSRG_00124 (-A)		
	SSRG_05139 (-A)	SSRG_05140 (-A)		
	SSRG_06258	SSRG_06257		
S. ghanaensis ATCC 14672			SSFG_06280*	
S. sp. E14			SSTG_04682*	SSTG_05266

Ten potentially functional LanBC clusters, four LanM clusters and three LanL clusters were identified in these streptomycete genomes. Four additional LanM enzymes were identified that do not contain the required catalytic and cofactor binding residues in their C-terminal cyclase domain, but they could be part of an active cluster as discussed below. Most genomes only seem to contain one or more clusters either of the LanBC or of the LanM type, whereas in *S. griseus* and *S. viridochromogenes*, both types are present. In *S. griseus*, a gene cluster was found with a potentially functional LanC cyclase (SGR_6574) and a gene with homology with the N-terminus of a LanM (SGR_6571). It was not possible to predict whether this truncated LanM would be capable of catalysing a dehydration reaction and no putative LanA could be identified near the cluster. Two potentially active orphan LanBs (SGR_5947 and SSFG_06284), two orphan LanCs (SCO4224/SSPG_03470 and SSMG_08135) and a LanL with no associated LanA (SGR_152) were also identified. Although not encoding a lantipeptide, a novel cryptic thiopeptide gene cluster was identified in *S. venezuelae* (SMD07779-SMD07795) as well as a previously identified thiopeptide cluster in *S. griseus* (SGR_4410-SGR_4418, Wieland-Brown *et al.* 2009).

For each enzyme type, all potentially active clusters could be subdivided into different subfamilies. This classification is based on their gene arrangement (Figure VII.1.) and the sequence of their predicted LanA (Figure VII.2.), with both criteria in good agreement.

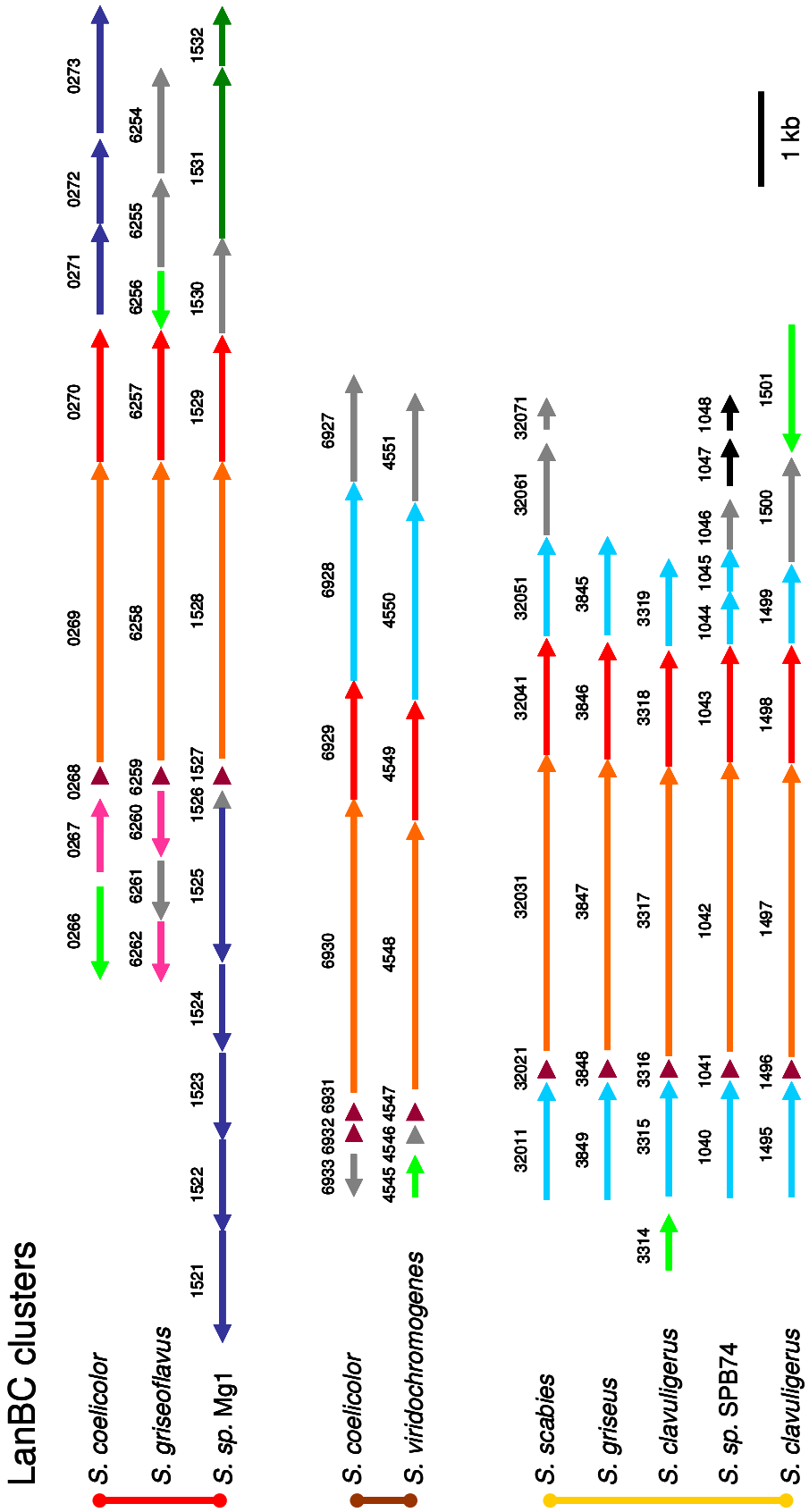


Figure VII.1. Schematic representation of gene clusters predicted to contain functional biosynthetic enzymes. Genes are colour-coded according to their predicted function: *lanA* (dark purple), *lanB* (orange), *lanC* (red), *lanM* (brown), *lanL* (olive green), *ramC*-like (yellow-green), transport (dark blue), regulation (light green), immunity (dark green), L-isoaspartyl methyltransferase homologue (light blue), other biosynthetic (pink), transposon-related (black) and unknown (grey). The different gene cluster families correspond to the LanA families shown in figure VII.2.

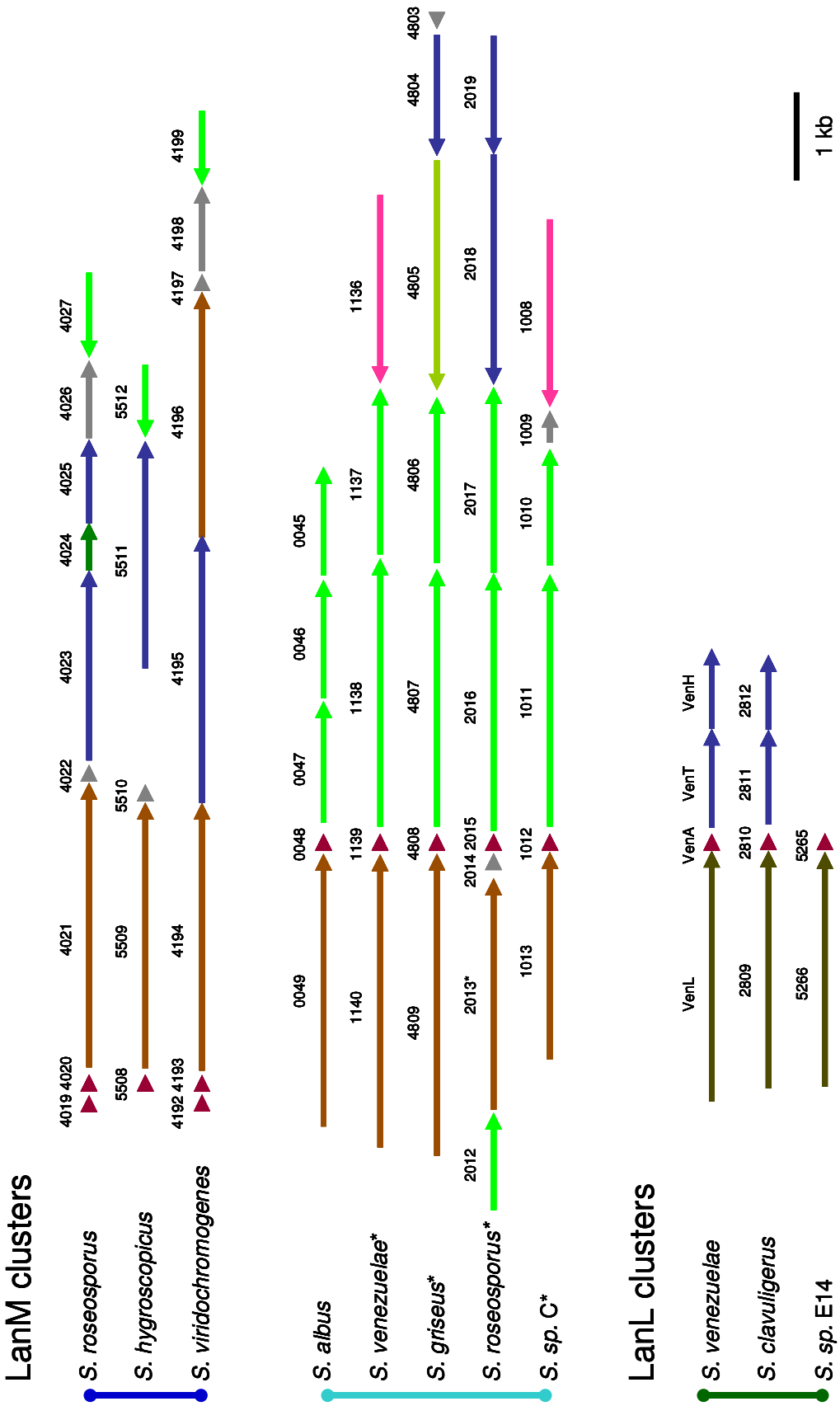


Figure VII.1. Continued.

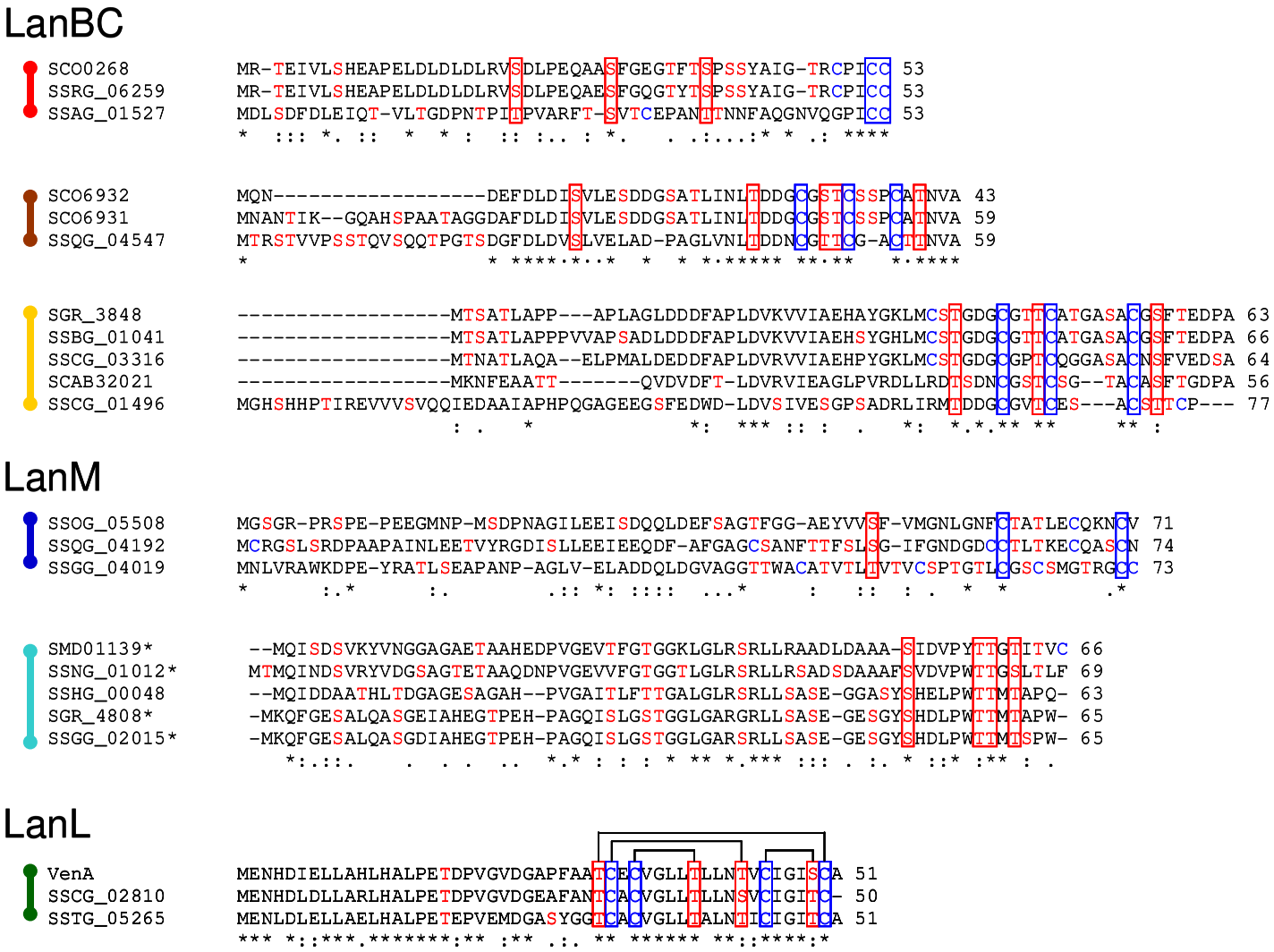


Figure VII.2. Multiple sequence alignment for related LanA peptides. Ser and Thr residues are depicted in red, Cys in blue. Ser/Thr or Cys that are conserved in all family members are boxed as they might indicate a shared ring topology. The ring topology has been indicated for the LanL family, based on the venezuelin structure (Goto *et al.* 2010). The different LanA families correspond to the gene cluster families shown in figure VII.1.

Three distinct families of LanAs were identified for the LanBC gene clusters. Conserved positions for Ser/Thr and Cys were identified for each separate family, suggesting a common ring topology for compounds within each family. The LanA peptides in the first family all contain a C-terminal Cys but no LanD enzyme was found in their biosynthetic clusters. However, several lantibiotics possess a C-terminal (Me)Lan, including salivaricin A, mutacin II and planosporicin (Ross *et al.* 1993; Krull *et al.* 2000; Maffioli *et al.* 2009), and the same may be true for this family of compounds. These clusters are predicted to be the largest LanBC clusters and are the only family for which likely transport or immunity genes could be identified. This makes this family a good candidate for mining lantipeptides with antibiotic activity. The second and third LanBC families contain LanA peptides that are well conserved in both the leader and the propeptide sequence (in the third family, only SSCG_01496 does not align well with the other LanAs). The biosynthetic gene clusters from both of these families all contain genes with homology to L-isoaspartyl methyltransferases (light blue in Figure VII.1.). This enzyme family is involved in the repair of damaged L-isoaspartyl and D-aspartate groups in proteins (Ryttersgaard *et al.* 2002). No members of this family have been associated previously with the post-translational modification of peptide natural products. However, given their invariant co-occurrence with these two LanBC family clusters in which the LanAs contain conserved Asp residues, it is possible that these methyltransferases are involved in the introduction of a novel type of post-translational modification, i.e. methylation of Asp residues.

The LanM gene clusters were divided into two families. The first family contains three similar gene cluster organisations, but the LanA sequences are quite dissimilar despite some conserved Ser/Thr and Cys positions in the propeptides. The leader sequences contain an area of greater homology, but no consensus sequence could be identified. This first LanM family appears to show the greatest variability in likely compound structures among all of the clusters analysed here. The second LanM family only contains one LanM with an active cyclase domain (SSHG_00049). The *S. venezuelae* LanM (SMD01140) has the conserved residues for the Zn-binding site, but lacks the catalytic His. All the other LanM enzymes in this family lack the entire Zn-binding motif. Thus, all but one of these clusters may be inactive. However, the amino acid sequences of the LanA peptides

are reasonably well conserved (Figure VII.2.), implying that they may still be functional. Moreover, apart from one C-terminal Cys in the *S. venezuelae* LanA (SMD01139), all of these LanAs are completely devoid of Cys residues, thus making an active cyclase domain redundant. If these gene clusters are capable of producing compounds, the resulting products are predicted to be linear dehydrated peptides. Before the cypemycin gene cluster was identified, analysis of the cryptic gene clusters of *S. venezuelae* was ongoing in parallel. At that time cypemycin was still believed to belong to the lantibiotic family of peptides, and to be the product of a cluster with a LanB or a LanM for the Ser/Thr dehydrations in combination with a LanD for the AviCys. We (wrongly) believed that SMD01139 from *S. venezuelae* represented a cypemycin-like modified peptide. The predicted product of this LanA is also a dehydrated peptide with a C-terminal Cys (no AviCys in this case because there is no LanD in the *S. venezuelae* cluster). With hindsight, this second LanM family could be involved in the production of compounds that are similar in structure to the newly identified linaridin family (Chapter IV.4.) regardless of the differences in the sequences of the respective biosynthetic genes.

For the LanL type, only one family was identified with highly similar structural peptides which extended to the leader region. The predicted ring topology, as determined for venezuelin (Goto *et al.* 2010), is likely conserved, as indicated in Figure VII.2.

VII.3. Cryptic gene clusters in *S. venezuelae*

In this section, the focus is on three cryptic lantipeptide gene clusters identified in *S. venezuelae*. One cluster (SMD01136 - SMD01140) is of the LanM-type and a second, very small cluster just consists of a gene encoding a RamC-type enzyme (Chapter I.4.3.c.) and an associated *lanA* (SMD01224 and SMD01224.2 respectively). The third cryptic cluster (SMD01252 – SMD01254) contains a novel type of modification enzyme with an N-terminal phosphoSer/Thr lyase domain, a Ser/Thr kinase domain and a C-terminal LanC-like cyclase domain with catalytic His and Zn binding site residues. The predicted *S. venezuelae* RamC cluster (SMD09638 – SMD09642) was not analysed since its predicted structural peptide is very similar to the well characterised *S. coelicolor* SapB lantipeptide (Kodani *et al.* 2004).

VII.3.1. Deletion of the modification gene for each cluster

In a first attempt to identify the products of the cryptic gene clusters, Apra^R replacement mutants were generated for the different modification genes as described in Chapter II.6. Deletion of a modification gene should result in the inactivation of the associated cluster, since the posttranslational modifications in the predicted lantipeptides are likely to be essential for biological activity. Provided that the gene cluster is expressed under the conditions used, this analysis would identify antibacterial peptides by lack of an inhibition zone in the mutant and morphogenetic peptides by an altered colony phenotype in the mutant. To maximise the chances of finding an interesting phenotype without interference from other secondary metabolites made by *S. venezuelae*, a screen was set up using several different agar media.

Deletion of individual modification genes resulted in *S. venezuelae* mutants M1436 (ΔSMD01140::(*oriT-aac(3)IV*)), M1437 (ΔSMD01224::(*oriT-aac(3)IV*)) and M1438 (ΔSMD01252::(*oriT-aac(3)IV*)). To rule out potential cross-talk between the SMD01140 and SMD01252 gene clusters, double mutants were generated as well by replacing the remaining lantipeptide synthetase in M1436 and M1438 with Hyg^R. This second round of mutation yielded strains M1439 (ΔSMD01140::(*oriT-*

aac(3)IV); ΔSMD01252::(*oriT-hyg*)) and M1440 (ΔSMD01252::(*oriT-aac(3)IV*); ΔSMD01140::(*oriT-hyg*)).

Two separate exconjugants for all of the above mentioned *S. venezuelae* mutants (M1436 – M1440) were grown on 14 different agar media (described in Chapter II.2.2.) and analysed for antibiotic activity against *M. luteus* after three and five days of growth. During incubation, mutant colony morphology was monitored in comparison to *S. venezuelae* WT on a daily basis. No morphological abnormality was observed for any of the deletion mutants and no differences in antibacterial activity against *M. luteus* were found when the mutants were compared to the WT strain on all 14 media. This screen did identify several conditions in which *S. venezuelae* WT as well as the mutants produced a zone of inhibition against *M. luteus*, indicative of the production of a biologically active secondary metabolite, possibly chloramphenicol or jadomycin.

VII.3.2. Micro-array expression data for the different clusters

It was deemed useful to identify the metabolite(s) causing the zones of inhibition in the aforementioned bio-assays, since the observed halos could be masking a smaller halo produced by the lantipeptide under study. The gene expression profiles of known and predicted secondary metabolic gene clusters from *S. venezuelae* were established by micro-array analysis (Maureen Bibb, unpublished data). It is important to note that the *S. venezuelae* RNA used for this experiment was extracted from liquid grown cultures, so the expression profiles may well be different from those existing in the agar-grown cultures used for the bio-assays. The data were therefore used only as a guide to identify possible candidate gene clusters responsible for the production of the inhibitory activity.

The micro-array analysis indicated that several of the putative secondary metabolite gene clusters were transcriptionally active. Among these were two developmentally expressed putative siderophore gene clusters. Figure VII.3. shows the expression profiles for three clusters from which the predicted product could have antibacterial activity, namely chloramphenicol, jadomycin and a cryptic NRP cluster. The chloramphenicol cluster (SMD01679 – SMD01700) is transcribed

transiently at higher levels at 14 hours and returns to basal levels of expression shortly after. A similar spike is seen in the profile for the jadomycin cluster (SMD08944 – SMD08982) and the cryptic NRP cluster (SMD10419 – SMD10434) at ten hours. Expression levels become very low for these three clusters at later time points and the time frame is much shorter than that used for the agar-grown cultures (three and five days). However, stable metabolites could have been produced from one or more of these gene clusters that interfered later with the bio-assay.

The expression profiles of the cryptic lantipeptide gene clusters and their surrounding regions were also analysed. The LanM cluster (SMD01140) appeared to be expressed only at very low levels under the conditions tested (Figure VII.3.). The gene expressed strongly after 14 hours (SMD01145) is located upstream of the cluster and encodes a protein of unknown function. The small cluster with the RamC-like enzyme (SMD01224) was expressed at low levels (data not shown), but the *lanA* from the third cluster (SMD01252.2) was expressed at high levels, reaching its maximum at a late stage of growth (Figure VII.3.). The expression profiles of the other genes in this cluster are quite flat, suggesting they are expressed at low levels. A small increase in expression can be observed at 16 hours for the two transport genes (SMD01253 and SMD01254), which coincides with the largest increase in *lanA* expression. These data suggest that the cluster might be producing a lantipeptide under the conditions tested. The expression profile for the RamC-orthologous cluster (SMD09638 – SMD09642) was included for comparison (Figure VII.3.). The two genes that are most active are first the *ramS* ortholog and then the *ramR* ortholog. *ramC* itself appears to be expressed at low levels. Note that is unknown whether active SapB is formed, since *ramC* was only expressed at very low levels and there would be no need for the surfactant activity during liquid growth.

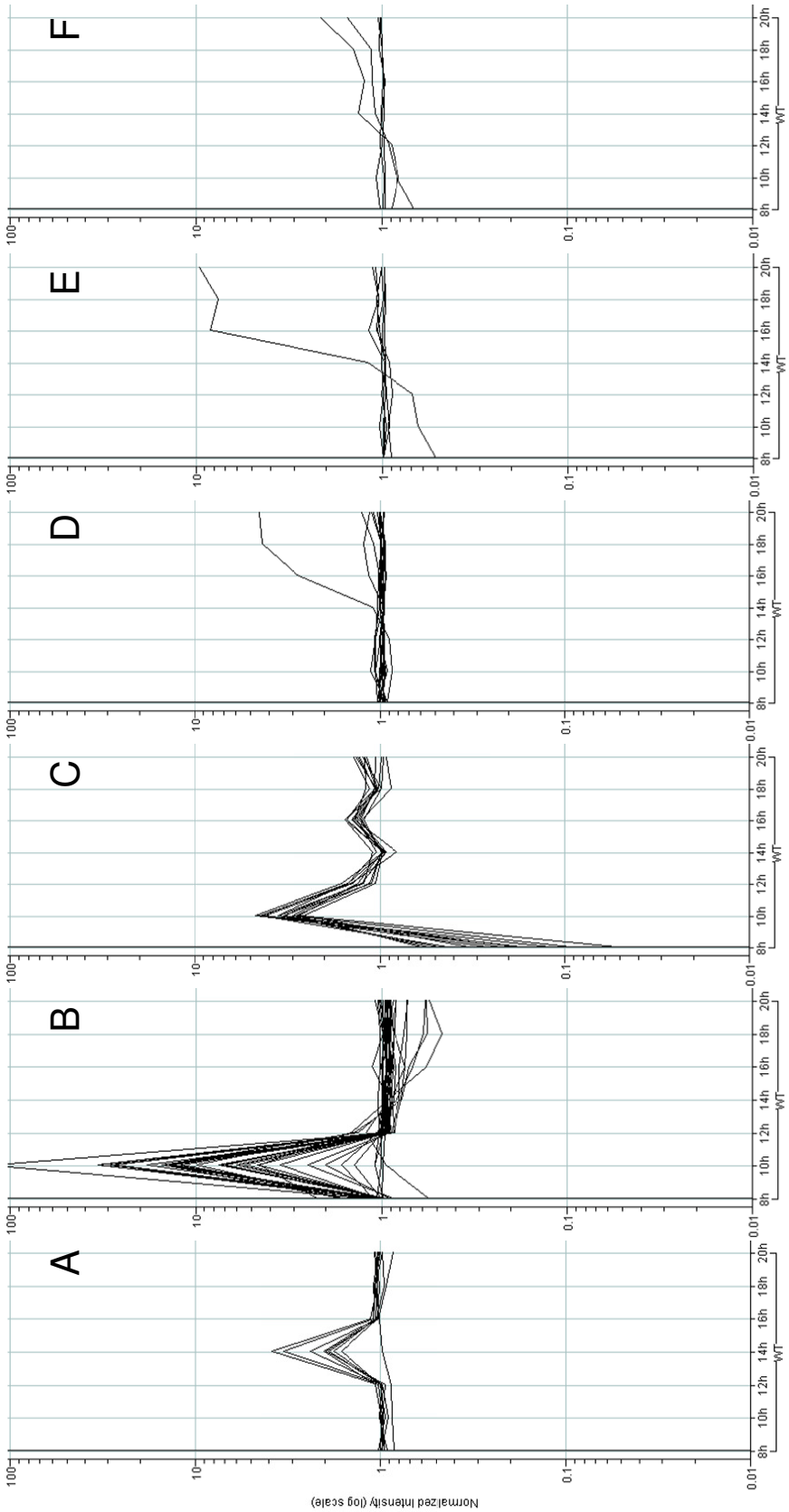


Figure VII.3. Gene expression profiles for selected secondary metabolite gene clusters from *S. venezuelae*. RNA from liquid cultures was analysed at two hour intervals between eight and 20 hours of growth post-inoculation. The profile is shown for biosynthetic gene clusters that are predicted to produce A) chloramphenicol (SMD01679 – SMD01700), B) jadomycin (SMD08944 – SMD08982), C) a cryptic NRP (SMD10419 – SMD10434), D) a cryptic lantipeptide (SMD01130 – SMD01145), E) another cryptic lantipeptide (SMD01246 – SMD01255) and F) the *S. venezuelae* SapB (SMD09638 – SMD09642).

To investigate whether the SMD01252 gene cluster produced an active compound in liquid culture, *S. venezuelae* WT and M1438 (ΔSMD01252) were grown under conditions similar to those used for the micro-array experiment. When the culture supernatant was analysed in a paper disk assay against *M. luteus*, a zone of inhibition was formed for both the WT and mutant strains, again indicating interference by another secondary metabolite in the bio-assay.

The chloramphenicol gene cluster was the most likely candidate for causing interference in the bio-assays and thus it was inactivated by generating the Apra^R replacement mutant M1441 (Δ*cmlPH*::(*oriT-aac(3)IV*)) as described in Chapter II.6. Gene disruption studies by He *et al.* (2001) have shown previously that the lack of either *cmlP* or *cmlH* results in total loss of chloramphenicol production without imbalancing the shikimate pathway.

The antibiotic activity of M1441 (Δ*cmlPH*) against *M. luteus* was assayed after growth on the same agar media that were used in the previous screens. Only one condition (MarM medium, three days of growth prior to overlay) was identified where the observed activity could be solely attributed to chloramphenicol production. On all of the other media, M1441 (Δ*cmlPH*) behaved in an identical fashion to *S. venezuelae* WT. This indicates that all other zones of inhibition were caused by one or more secondary metabolite other than chloramphenicol. No attempts were made to inactivate additional gene clusters or to identify the antibiotic activity.

VII.3.3. Heterologous expression of the clusters

In an attempt to avoid the interference of other secondary metabolites produced by *S. venezuelae* in the bio-assays, the lantipeptide gene clusters were introduced into a heterologous host. *S. lividans* was chosen since it was known not to produce antibacterial activity under a variety of growth conditions.

The predicted LanM cluster (SMD01140) lies close to the ends of two cosmid inserts in the *S. venezuelae* cosmid library (in cosmids 1-B5 and 4-G03). The region directly up- or down-stream of the predicted cluster is missing from one or

other cosmid. To account for the possible involvement of one or other of these flanking regions in lantipeptide synthesis, both cosmids were used for heterologous expression. Heterologous expression of the small RamC-like cluster (SMD01224) was not attempted. Cosmid 4H08 was used for heterologous expression of the SMD01252 cluster, which is located in the centre of the cosmid insert. The cosmid backbones were targeted with the integrative fragment of pIJ10702 (Chapter II.6.) and introduced into *S. lividans* by conjugation, resulting in strains M1447 (1-B5), M1448 (4-G03) and M1449 (4H08). A control strain (M1446) was generated by introducing just the vector pIJ10702. The strains were grown on the same media that had been used in the previously described screens. Morphology was monitored during growth and overlay assays were performed with *M. luteus* as an indicator after three and five days. No differences were observed between the strains with integrated cosmids and the WT and empty vector controls.

VII.3.4. Chemical induction of antibiotic production

Several studies have investigated the effect of growth media supplements on antibiotic production. *N*-acetylglucosamine (GlcNAc) has been identified by Rigali *et al.* (2008) as a starvation signal in *S. coelicolor* that elicits a response from the pathway-specific activators of antibiotic biosynthetic gene clusters. GlcNAc stimulated antimicrobial activity in several streptomycetes, suggesting that it may provide a valuable strategy for the activation of cryptic antibiotic gene clusters. A second small molecule that has been linked to increased antibiotic production is SAM. Overproduction of SAM synthetases in *S. coelicolor*, *S. avermitilis* and *S. peucetius* caused an increase in production of actinorhodin (Act), avermectin and doxorubicin, respectively (Okamoto *et al.* 2003; Yoon *et al.* 2006). Addition of exogenous SAM was also shown to enhance streptomycin production in *S. griseus* through increased *adpA* transcription (Shin *et al.* 2006).

The effect of GlcNAc and SAM on the expression of the cryptic lantipeptide gene clusters of *S. venezuelae* was assessed. *S. venezuelae* WT and the lantipeptide synthetase deletion strains M1436 – M1440 could not be used in this assay, since they produce a zone of inhibition against *M. luteus* after growth on MM (1.5% LabM, 0.5 % mannitol). Instead, *S. lividans* strains M1447 – M1449 with the

introduced cosmids were analysed. The addition of 25 µM SAM to the medium did not influence production of antibacterial activity in a *M. luteus* bioassay (Fig. VII.4.A.). However, zones of inhibition were observed for the heterologous expression strains M1447 – M1449 and not the WT or empty vector (M1446) control strains when 1 % GlcNAc was added to the growth medium (Fig. VII.4.B.). Unfortunately, while this initial observation was repeated three times within one month, it could not be reproduced five months later, when additional *S. lividans* strains had been generated containing cosmids with an in-frame deleted *lanA* (M1450 – M1452) or modification gene (M1453 – M1455). The reason for this lack of reproducibility is not known, but it could reflect, for example, a change in a medium component.

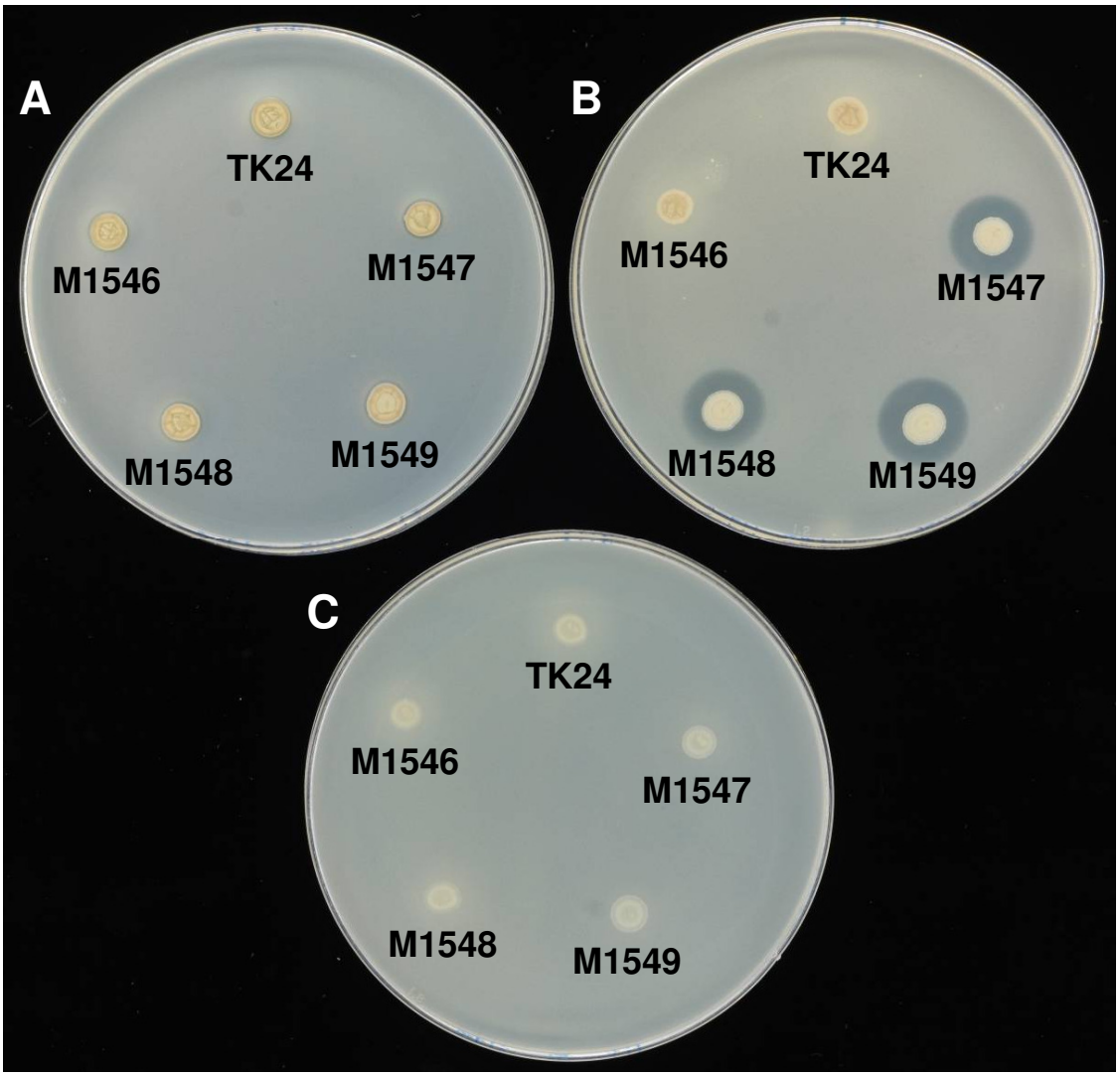


Figure VII.4. Chemical induction of the heterologously expressed *S. venezuelae* lantipeptide clusters. *S. lividans* WT, and strains containing an empty vector control (M1446), the SMD01140 cluster (M1447 (1-B5) and M1448 (4-G03)) and the SMD01252 cluster (M1449 (4H08)) were grown for three days at 30 °C on MM (1.5 % LabM, 0.5 % mannitol) containing A) 25 µM SAM, B) 1 % GlcNAc and C) no inducer. An overlay of *M. luteus* in SNA was applied and halos were scored after overnight incubation at 30 °C. The strains containing the *S. venezuelae* cosmids produced a zone of inhibition only on the medium with added GlcNAc. Unfortunately, the observation could not be repeated five months later (refer to text).

VII.3.5. Constitutive expression of the modification enzymes in each cluster

An attempt was next made to constitutively express the genes encoding the modification enzymes from each cluster. A PCR targeting construct was generated that contained an Apra^R marker, the strong constitutive promoter *ermE**p and the EF-Tu RBS from pIJ10257 (based on the design by Sean O'Rourke). This cassette (pIJ10704) can be introduced directly upstream of the start codon of a target gene, resulting in its constitutive expression.

The functionality of the *ermE** construct was tested by introducing it in front of *afsR* in *S. coelicolor*, resulting in strain M1456 (*ermE*-afsR*). Previous work had shown that introduction of multiple copies of *afsR* had a stimulatory effect on the production of Act and undecylprodigiosin (Red), the two pigmented antibiotics produced by *S. coelicolor* (Floriano and Bibb 1996). M1456 (*ermE*-afsR*) exconjugants were easily identified by the copious blue droplets of Act on their surface when grown on SFM medium (Figure VII.5.; *S. coelicolor* does not produce significant amounts of Act on SFM). This experiment demonstrated that the expression construct was functional.

The *ermE** construct was introduced upstream of the three lantipeptide synthetase genes in *S. venezuelae*, resulting in strains M1442 (*ermE**-SMD01140), M1443 (*ermE**-SMD01224) and M1444 (*ermE**-SMD01252). When these strains were tested for antimicrobial activity against *M. luteus*, no difference was observed when compared to the WT strain. The supernatant of liquid grown cultures was analysed by MALDI-ToF, but no differences were observed for the spectra of the WT, the M1442 (*ermE**-SMD01140) and M1443 (*ermE**-SMD01224) expression strains. However, in the spectrum of the M1444 (*ermE**-SMD01252) expression strain, a set of mass peaks was observed at much greater abundance than in the WT spectrum. All other peaks in both spectra were of comparable signal intensity. MALDI-ToF does not allow quantitative interpretation, but this observation led us to investigate the set of elevated peaks more thoroughly. A pattern of oxidations (+16 Da mass increments), and sodium (+22 Da) and potassium (38 Da) adducts was observed for a compound with $[M+H]^+$ = 2124.175 Da (Figure VII.6.A.). Similar

spectra have been observed for lantibiotics (Lucy Foulston and Robert Bell, personal communication) and are thought to be caused by oxidation of (Me)Lan prior to and during MALDI-ToF analysis. Up to four oxidations could be observed for the sodium adduct of the compound, indicating that if the compound with the $[M+H]^+$ of 2124 Da is a lantipeptide, it could contain four (Me)Lan bridges.

The calculated $[M+H]^+$ of a SMD01252.2-derived lantipeptide with four bridges, cleaved from the leader peptide after Ala29 is 2125 Da. This is a difference of 1 Da compared to the observed compound. MALDI-ToF analysis of the supernatant of M1438 (ΔSMD01252) revealed that all peaks apparently derived from the compound with $[M+H]^+$ = 2124 Da were still present. To confirm that this compound was definitely not a product of the SMD01252 cluster, an $Apra^R$ replacement mutant was made for *lanA* as well: M1445 (ΔSMD01252.2::(*oriT-aac(3)IV*)). None of the peaks disappeared in the mass spectrum of this mutant either.

Q-ToF analysis of the unidentified compound required a much higher collision energy compared to the energy used for cypemycin fragmentation. The spectrum did not show fragment masses from which an unambiguous amino acid sequence could be derived. No further attempts were made to identify the structure of the $[M+H]^+$ = 2124 Da compound.

It is unlikely that the observed compound is the *S. venezuelae* SapB morphogen ortholog, which could possibly be produced in liquid media (Chapter VII.3.2.). The *S. venezuelae* SapB has not been characterised, but the structural peptide is nearly identical to *S. coelicolor* SapB (Kodani *et al.* 2004). Both are very likely to be processed in a similar fashion, yielding a compound with a predicted mass of 2013 Da for *S. venezuelae* SapB (Figure VII.6.C.). Moreover, SapB only contains two Lan bridges, so this would not explain a pattern with four oxidations.

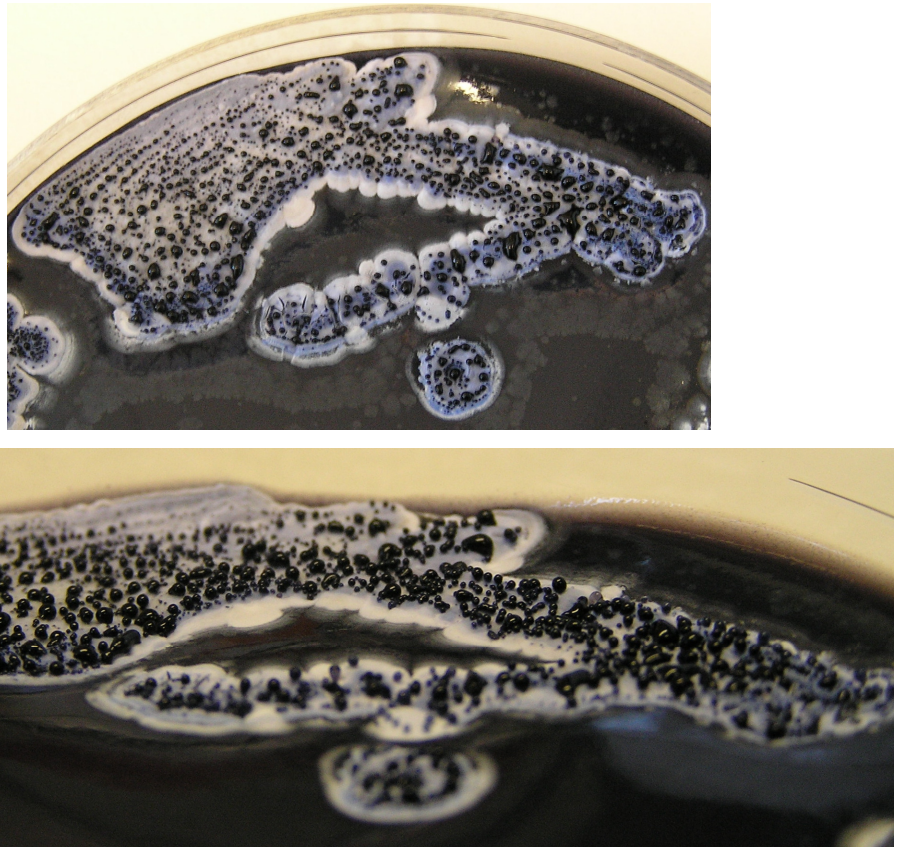


Figure VII.5. Act overproduction phenotype of *S. coelicolor* M1456 (*ermE*-afsR*) on SFM. Blue droplets of Act were observed on top of the colonies, confirming the functionality of the *ermE** construct. Plates had been incubated for five days at 30 °C when the pictures were taken.

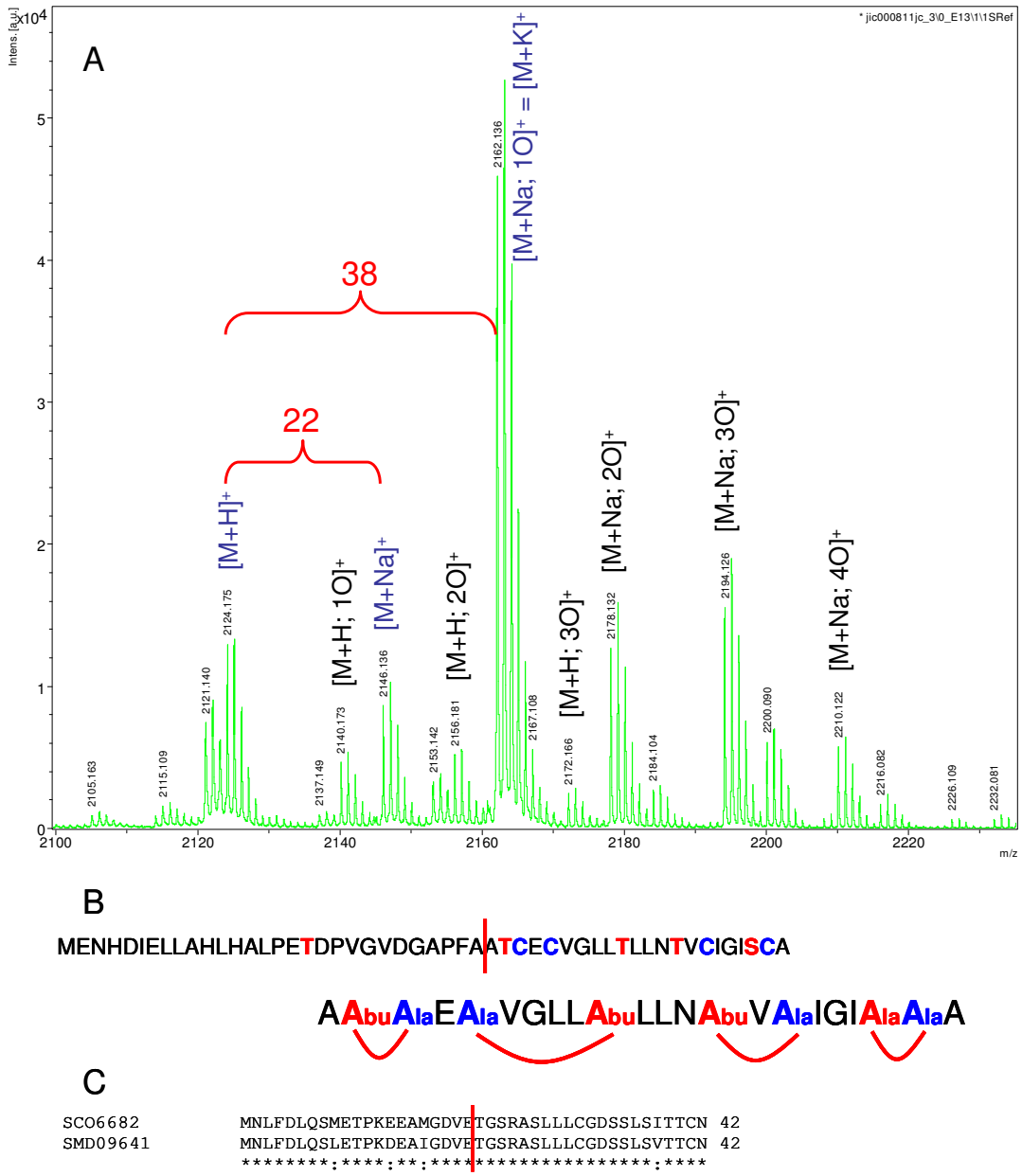


Figure VII.6. Identification of a possible product from the cryptic SMD01252 gene cluster. A) MALDI-ToF spectrum for a compound with $[M+H]^+$ = 2124 Da. Peaks representing oxidations, sodium or potassium adducts are annotated accordingly. B) Schematic representation of a possible product from the SMD01252 cluster. Ser/Thr residues in the prepropeptide sequence are depicted in red and Cys in blue. This colour coding is maintained in the model for the modified peptide which is based on a leader cleavage after Ala29 (indicated by a vertical red line) and the introduction of four dehydrations and subsequent cyclisations. The ring topology shown is completely speculative and not based on experimental data. C) Alignment of the SapB prepropeptide sequences from *S. coelicolor* (SCO6682) and *S. venezuelae* (SMD09641). The site of leader cleavage in *S. coelicolor* is indicated by a vertical red line. The only difference in the propeptide sequences is an Ile-Val substitution.

VII.4. Discussion

A bioinformatic analysis of the 21 available streptomycete genomes (counting the two *S. roseosporus* strains only once) identified several cryptic lantipeptide gene clusters. Ten LanBC, four LanM and three LanL clusters contained a *lanA* and genes for apparently functional modification enzyme(s), as judged from amino acid sequence alignments and the presence of conserved residues. Four additional clusters were identified with a LanM enzyme that lacks the catalytic and Zn-binding residues in their cyclase domain. However, the LanA peptides that are associated with these enzymes are devoid of Cys residues (with the exception of a C-terminal Cys in SMD01139). Thus these clusters could still be active and might produce linear dehydrated peptides, bringing the total for potentially active LanM-like clusters to eight.

A PCR approach with degenerate primers was previously used by Dodd *et al.* (2006) to screen about 100 actinomycete genomes for the presence of *lanA* and *lanM* genes. About 20 % of the strains tested proved positive, likely an underestimate given the restrictions imposed by the use of degenerate primers. A recent bioinformatic study by Begley *et al.* (2009) identified 89 LanM homologs in the publicly available bacterial genome sequences. The eight potentially active LanM clusters described in this chapter make up a significant proportion of this total, indicating that streptomycetes are good candidates for lantipeptide genome mining. The identification of bacteriocin gene clusters in newly sequenced genomes will be facilitated by the recently developed BAGEL2 software (de Jong *et al.* 2010). This program can be used for the annotation of lantipeptide clusters from the LanBC, LanM and even the recently characterised LanL types. All clusters could be classified into distinct families based on their gene organisations and their structural peptide sequences. This could mean that despite the relatively large number of clusters in *Streptomyces* genomes, the number of compound scaffolds with a specific (Me)Lan ring pattern could be limited. This apparent conservation may be indicative of a shared biological activity.

There are several different strategies to obtain the product of a cryptic biosynthetic gene cluster, as reviewed by Zerikly and Challis (2009). Most of the genetic

approaches to identify products from the cryptic lantipeptide clusters of *S. venezuelae* were attempted, but none of them proved to be successful.

A collaboration with the van der Donk group was initiated to employ the '*in vitro* reconstitution' strategy on the SMD01252 gene cluster containing the novel lantipeptide synthetase. This involved the purification of the structural peptide and the modification enzyme to obtain a product from their combined incubation in an *in vitro* reaction. The *in vitro* reconstitution approach had already been successfully applied to identify haloduracin, the two-component lantibiotic product of a cryptic gene cluster found in *Bacillus halodurans* C-125 (McClerren *et al.* 2006). However, results have to be interpreted with caution, as additional mutagenesis studies required a revision of the haloduracin structure (Cooper *et al.* 2008).

Upon incubation of the purified SMD01252.2 structural peptide with the purified modification enzyme SMD01252, a four-fold dehydrated and cyclised product was formed (Goto *et al.* 2010). This compound was named venezuelin and the genes in the cluster were reassigned with the standard lantibiotic nomenclature: *venA* for the structural gene and *venTH* for the two downstream transporter genes. The modification enzyme was called VenL, since its activity has now been shown and it belongs to a novel class of lanthionine synthetases designated LanL. The three distinct domains in VenL (a small N-terminal phosphoSer/Thr lyase domain, a central Ser/Thr kinase domain and a C-terminal LanC-like cyclase) were expressed individually and shown to be catalytically active, providing insight into the mechanism of LanL enzymes. The kinase domain initiates the modification of the propeptide by phosphorylating Ser and Thr residues. A subsequent elimination of the phosphate groups by the lyase domain results in the formation of Dha and Dhb residues. Finally, the cyclase domain catalyses the addition of Cys residues onto the dehydrated residues, forming (Me)Lan bridges. A series of VenA analogs was generated in which one Cys was replaced by Ala. Interpretation of the fragmentation patterns of tandem mass spectrometry analysis allowed determination of the venezuelin ring topology (Goto *et al.* 2010). Venezuelin had no antibacterial activity against three different indicator strains that are typically highly sensitive to various lantibiotics.

This could explain why venezuelin was not detected in the bio-assays described in this chapter. No antibiotic activity could not be demonstrated for the Sap lantipeptides, which are generated by related modification enzymes (Kodani *et al.* 2004; Kodani *et al.* 2005). The biological function of venezuelin – and by extension, the products of many other cryptic lantipeptide clusters – may not be antibacterial. Apart from a possible morphological function, the compounds could, for example, serve as signalling molecules. Indeed, for some lantibiotics, a quorum sensing activity has been demonstrated involving the detection of the mature compound by a dedicated LanRK two-component system resulting in regulation of the biosynthetic and immunity genes (Kuipers *et al.* 1995; Stein *et al.* 2002; Kleerebezem *et al.* 2004).

VII.5. Bullet point summary

- Bioinformatic analysis identified many cryptic lantipeptide gene clusters in sequenced streptomycete genomes. These clusters were classified into distinct families based on their gene organisation and their LanA sequences.

- A novel type of lantipeptide modification enzyme containing an N-terminal phosphoSer/Thr lyase, a Ser/Thr kinase and a LanC-like cyclase domain was identified in one of the cryptic gene clusters from *S. venezuelae*.

- Various *in vivo* approaches were unsuccessfully taken to obtain products from the three lantipeptide gene clusters in *S. venezuelae*.

- An *in vitro* approach by collaborators identified the lantipeptide venezuelin as the product from the cryptic gene cluster with the unusual LanL enzyme.

Chapter VIII – General discussion

Starting from its discovery in 1993 (Komiyama *et al.*), cypemycin has always been hard to classify on a structural basis. The occurrence of four Dhb residues and an AviCys in the mature peptide led to its understandable classification as a lantibiotic (Chatterjee *et al.* 2005). However, cypemycin does not contain the (Me)Lan residues that define the lantibiotic family. The genetic (and some preliminary biochemical) analysis performed in this thesis surprisingly revealed that cypemycin is not a lantibiotic, but the first representative of a novel class of post-translationally modified peptides we named 'linaridins'.

Genome scanning of *S. sp.* OH-4156 was of key importance for the identification of the cypemycin biosynthetic gene cluster, as discussed in Chapter III. A logical first step to take with this data at hand was a bioinformatic search for lantibiotic modification genes. At that time, CinM was the only characterised streptomycete lanthionine synthetase (Widdick *et al.* 2003), although several additional cryptic lanthionine synthetase-like genes had been identified in the sequenced genomes of *S. coelicolor*, *S. avermitilis*, *S. scabies* and *S. venezuelae* (Mervyn Bibb, unpublished data). Deletion and heterologous expression of the cryptic *S. venezuelae* lantipeptide clusters (Chapter VII) had just been completed and the genes encoding their respective lanthionine synthetases were used as the first BLAST queries to search the *S. sp.* OH-4156 genome scanning data. Encouragingly, several contigs were identified in this manner, but closer inspection of their nucleotide sequences revealed that they were about 95 % identical to the *S. venezuelae* sequences. This led initially to the incorrect conclusion that the genomic DNA submitted for sequencing was extracted from a culture that had been contaminated and usurped by *S. venezuelae*. Luckily, anxiety was soon replaced by renewed excitement when the *cypA* gene was identified in an additional tBLASTn search. The initial inspection of this ORF, which encodes a prepropeptide with two Cys precursors for the AviCys, taken together with the lack of apparent lanthionine synthetases in the *S. sp* OH-4156 sequence, hinted that cypemycin would indeed be a very special peptide. Full appreciation of this took shape after the biosynthetic gene cluster was sequenced and the function of the individual genes was confirmed by a deletion analysis (Chapter IV).

Comparative genomic studies of bacterial strains belonging to the same species have identified so-called species-specific 'core genomes' (Medini *et al.* 2008). These are arbitrary collectives containing the genes that are shared by every member of the species under analysis. Comparison of the preliminary *S. sp.* OH-4156 genome scanning data with the full genome sequence of *S. venezuelae* revealed a remarkable degree of sequence homology (Figure III.7.A.). There appears to be a shared genomic 'scaffold' that differs only in the complement of secondary metabolite gene clusters. This might well be a common theme among different streptomycetes and could have implications for the use of genome sequencing for natural product discovery. When looking for the genes responsible for production of a known metabolite, genome scanning is probably the best starting point, as exemplified with cypemycin in this thesis and with the recent discovery of the ribosomal origin for thiopeptide antibiotics (Figure I.1.H.). A potentially useful lesson that can be learned from the core or 'scaffold' genes is that they could be used as guidelines for cloning predicted secondary metabolite gene clusters. In various cases it is probably safe to define a cluster's boundaries based on where the core genome stops and additional genes begin, as exemplified by the chloramphenicol cluster (Figure III.7.B.).

Bacterial genomes are full of small ORFs that are embedded in clusters containing putative biosynthetic genes. At the moment it is hard to estimate what proportion of these produce (post-translationally modified) peptides. There are potentially dozens of unusual peptide families or post-translational modifications that await identification with potentially novel biological functions. Most compounds are identified in activity-based screens, for example by looking for antibiotic activity. However, various types of modified peptides have been described that do not primarily have an antibiotic function. An example of this is the lasso peptides, which have an intriguing structure and underlying biosynthesis, but only two out of over ten structurally characterised compounds have significant antibacterial activity. Similarly, cypemycin, which was identified because of its cytocidal activity against mouse leukemia cells (Komiyama *et al.* 1993) would almost certainly have been discarded in an antibacterial screen. It would be interesting to investigate different bacteria (Figure IV.7.) for the production of linaridins as this could help in the identification of this family's biological function, which might well be regulatory.

Indeed, Gram-positive bacteria are known to use diverse types of modified peptides specifically for quorum sensing. This does not exclude a supplementary function, since some lantibiotics have been shown to serve as auto-inducers for their own synthesis in addition to possessing antimicrobial activity. Another possibility is that the linaridins could serve a specific structural and/or developmental role like, for example, the Sap morphogens or the unmodified chaplins and rodlins in streptomycetes. No phenotypic differences could be observed upon deletion of *cypA*, so if cypemycin has a structural/developmental function, it is not involved in the processes required for growth under the laboratory conditions tested. Moreover, if the linaridins' function was strictly developmental, their gene clusters would be expected to occur more commonly and be less wide-spread among different bacterial lineages. Cypemycin's hydrophobic nature and resulting affinity for the cytoplasmid membrane might provide a clue. Hydrophobic peptides have been identified that modulate the activity and/or stability of certain proteins inserted into the inner cell membrane of *E. coli* and *Salmonella typhimurium* (Alix and Blanc-Potard 2009).

Recent advances in the understanding of post-translationally modified peptides cannot be attributed solely to genome scanning. Improvements in mass spectrometric techniques has also contributed greatly and not only to the identification of novel peptides and post-translational modifications, but also to the biochemical characterisation of their biosynthetic machinery. For this project, Q-ToF MS has been extremely useful for the structural analysis of cypemycin and its variants, particularly the non-dehydrated version, which was shown to have its internal Cys dethiolated. The linear (non-cyclised) nature of cypemycin lends itself well to this type of analysis and the need for only small sample quantities makes it compatible with biochemical assays. It would be interesting to determine how the Thr dehydration and the Cys dethiolation reactions are catalysed. However, this could turn out to be quite challenging because the candidate enzymes lack homology to functionally characterised proteins and thus far the biochemical characterisation of LanB enzymes has proved particularly difficult. None-the-less, the N-methyltransferase CypM and the decarboxylase CypD are significant contributions to the ever-growing 'enzyme toolbox' that could be used for rational

peptide engineering. Modification of a peptide's termini is expected to provide increased resistance against amino- and carboxypeptidases (Rink *et al.* 2010).

In the future, it will become more straightforward to identify and characterise post-translationally modified peptide natural products sourced from different organisms or ecological niches. Molecular genetic techniques will be developed to manipulate the producers and/or to design systems for heterologous production. Several genetic and biochemical strategies have already proven their merit in the activation of cryptic secondary metabolite gene clusters (Challis 2008). Two nice examples of peptides from Eukaryotes are the conopeptides and bombinins. Conopeptides are a diverse collection of post-translationally modified peptides that are produced by the conus snail as a venom to paralyse its prey (Lewis 2009). Bombinins are a family of hydrophobic peptides that contain D-*allo*-Ile and are secreted by frogs of the *Bombina* genus (Simmaco *et al.* 2009). The patellamides (Figure I.1.J. and I.1.K.) were originally identified in extracts from marine ascidians, but turned out to be produced by symbiotic cyanobacteria. Different niches that are currently being mined for natural products include marine organisms, ants that use bacteria to destroy pathogens in their cultivated fungus gardens, and human-associated microbiota. These efforts will of course identify compounds with all sorts of different chemical scaffolds and structures, but there will almost certainly be some (modified) peptides among those. To conclude, many more post-translational modifications and biochemical mechanisms remain to be discovered in the exciting and expanding peptide world.

References

Alix, E. and A. B. Blanc-Potard (2009). Hydrophobic peptides: novel regulators within bacterial membrane. *Mol Microbiol*, **72**(1): 5-11.

Allgaier, H., G. Jung, R. G. Werner, U. Schneider and H. Zahner (1986). Epidermin: sequencing of a heterodetic tetracyclic 21-peptide amide antibiotic. *Eur J Biochem*, **160**(1): 9-22.

Altena, K., A. Guder, C. Cramer and G. Bierbaum (2000). Biosynthesis of the lantibiotic mersacidin: organization of a type B lantibiotic gene cluster. *Appl Environ Microbiol*, **66**(6): 2565-71.

Anderssen, E. L., D. B. Diep, I. F. Nes, V. G. Eijsink and J. Nissen-Meyer (1998). Antagonistic activity of *Lactobacillus plantarum* C11: two new two-peptide bacteriocins, plantaricins EF and JK, and the induction factor plantaricin A. *Appl Environ Microbiol*, **64**(6): 2269-72.

Ansaldi, M., D. Marolt, T. Stebe, I. Mandic-Mulec and D. Dubnau (2002). Specific activation of the *Bacillus* quorum-sensing systems by isoprenylated pheromone variants. *Mol Microbiol*, **44**(6): 1561-73.

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. *Chem Biol*, **16**(5): 490-8.

Arndt, H. D., S. Schoof and J. Y. Lu (2009). Thiopeptide antibiotic biosynthesis. *Angew Chem Int Ed Engl*, **48**(37): 6770-3.

Bagley, M. C., J. W. Dale, E. A. Merritt and X. Xiong (2005). Thiopeptide antibiotics. *Chem Rev*, **105**(2): 685-714.

Baquero, F., D. Bouanchaud, M. C. Martinez-Perez and C. Fernandez (1978). Microcin plasmids: a group of extrachromosomal elements coding for low-molecular-weight antibiotics in *Escherichia coli*. *J Bacteriol*, **135**(2): 342-7.

Bayro, M. J., J. Mukhopadhyay, G. V. Swapna, J. Y. Huang, L. C. Ma, E. Sineva, P. E. Dawson, G. T. Montelione and R. H. Ebright (2003). Structure of antibacterial peptide microcin J25: a 21-residue lariat protoknot. *J Am Chem Soc*, **125**(41): 12382-3.

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. *Appl Environ Microbiol*, **75**(17): 5451-60.

Bellomio, A., P. A. Vincent, B. F. de Arcuri, R. N. Farias and R. D. Morero (2007). Microcin J25 has dual and independent mechanisms of action in *Escherichia coli*: RNA polymerase inhibition and increased superoxide production. *J Bacteriol*, **189**(11): 4180-6.

Bentley, S. D., K. F. Chater, A. M. Cerdeño-Tárraga, G. L. Challis, N. R. Thomson, K. D. James, D. E. Harris, M. A. Quail, H. Kieser, D. Harper, A. Bateman, S. Brown, G. Chandra, C. W. Chen, M. Collins, A. Cronin, A. Fraser, A. Goble, J. Hidalgo, T. Hornsby, S. Howarth, C. H. Huang, T. Kieser, L. Larke, L. Murphy, K. Oliver, S. O'Neil, E. Rabbinowitsch, M. A. Rajandream, K. Rutherford, S. Rutter, K. Seeger, D. Saunders, S. Sharp, R. Squares, S. Squares, K. Taylor, T. Warren, A. Wietzorrek, J. Woodward, B. G. Barrell, J. Parkhill and D. A. Hopwood (2002). Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature*, **417**(6885): 141-7.

Betschel, S. D., S. M. Borgia, N. L. Barg, D. E. Low and J. C. De Azavedo (1998). Reduced virulence of group A streptococcal Tn916 mutants that do not produce streptolysin S. *Infect Immun*, **66**(4): 1671-9.

Bibb, M. J. (2005). Regulation of secondary metabolism in streptomycetes. *Curr Opin Microbiol*, **8**(2): 208-15.

Bierbaum, G., M. Reis, C. Szekat and H. G. Sahl (1994). Construction of an expression system for engineering of the lantibiotic Pep5. *Appl Environ Microbiol*, **60**(12): 4332-8.

Bierbaum, G. and H. G. Sahl (2009). Lantibiotics: mode of action, biosynthesis and bioengineering. *Curr Pharm Biotechnol*, **10**(1): 2-18.

Bierbaum, G., C. Szekat, M. Josten, C. Heidrich, C. Kempter, G. Jung and H. G. Sahl (1996). Engineering of a novel thioether bridge and role of modified residues in the lantibiotic Pep5. *Appl Environ Microbiol*, **62**(2): 385-92.

Blaesse, M., T. Kupke, R. Huber and S. Steinbacher (2000). Crystal structure of the peptidyl-cysteine decarboxylase EpiD complexed with a pentapeptide substrate. *Embo J*, **19**(23): 6299-310.

Blaesse, M., T. Kupke, R. Huber and S. Steinbacher (2003). Structure of MrsD, an FAD-binding protein of the HFCD family. *Acta Crystallogr D Biol Crystallogr*, **59**(Pt 8): 1414-21.

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. *J Antibiot (Tokyo)*, **63**(7): 351-8.

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. *Mol Microbiol*, **72**(5): 1126-36.

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. *Antimicrob Agents Chemother*, **50**(4): 1449-57.

Breukink, E. (2006). A lesson in efficient killing from two-component lantibiotics. *Mol Microbiol*, **61**(2): 271-3.

Breukink, E., H. E. van Heusden, P. J. Vollmerhaus, E. Swiezewska, L. Brunner, S. Walker, A. J. Heck and B. de Kruijff (2003). Lipid II is an intrinsic component of the pore induced by nisin in bacterial membranes. *J Biol Chem*, **278**(22): 19898-903.

Breukink, E., I. Wiedemann, C. van Kraaij, O. P. Kuipers, H. Sahl and B. de Kruijff (1999). Use of the cell wall precursor lipid II by a pore-forming peptide antibiotic. *Science*, **286**(5448): 2361-4.

Brötz, H., G. Bierbaum, P. E. Reynolds and H. G. Sahl (1997). The lantibiotic mersacidin inhibits peptidoglycan biosynthesis at the level of transglycosylation. *Eur J Biochem*, **246**(1): 193-9.

Brurberg, M. B., I. F. Nes and V. G. Eijsink (1997). Pheromone-induced production of antimicrobial peptides in *Lactobacillus*. *Mol Microbiol*, **26**(2): 347-60.

Busso, D., B. Delagoutte-Busso and D. Moras (2005). Construction of a set Gateway-based destination vectors for high-throughput cloning and expression screening in Escherichia coli. *Anal Biochem*, **343**(2): 313-21.

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): 5884-95.

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. *Chem Biol*, **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. *J Biol Chem*, **270**(40): 23533-9.

Chakraburtty, R. and M. Bibb (1997). The ppGpp synthetase gene (*relA*) of *Streptomyces coelicolor* A3(2) plays a conditional role in antibiotic production and morphological differentiation. *J Bacteriol*, **179**(18): 5854-61.

Challis, G. L. (2008). Mining microbial genomes for new natural products and biosynthetic pathways. *Microbiology*, **154**(Pt 6): 1555-69.

Chatterjee, C., L. M. Miller, Y. L. Leung, L. Xie, M. Yi, N. L. Kelleher and W. A. van der Donk (2005). Lacticin 481 synthetase phosphorylates its substrate during lantibiotic production. *J Am Chem Soc*, **127**(44): 15332-3.

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 lacticin 481 synthetase. *Chem Biol*, **13**(10): 1109-17.

Chatterjee, C., M. Paul, L. Xie and W. A. van der Donk (2005). Biosynthesis and mode of action of lantibiotics. *Chem Rev*, **105**(2): 633-84.

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 Microbiol Lett*, **195**(2): 139-44.

Cherepanov, P. P. and W. Wackernagel (1995). Gene disruption in *Escherichia coli*: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene*, **158**(1): 9-14.

Cheung, W. L., S. J. Pan and A. J. Link (2010). Much of the microcin J25 leader peptide is dispensable. *J Am Chem Soc*, **132**(8): 2514-5.

Chikindas, M. L., M. J. Garcia-Garcera, A. J. Driessen, A. M. Ledeboer, J. Nissen-Meyer, I. F. Nes, T. Abee, W. N. Konings and G. Venema (1993).

Pediocin PA-1, a bacteriocin from *Pediococcus acidilactici* PAC1.0, forms hydrophilic pores in the cytoplasmic membrane of target cells. *Appl Environ Microbiol*, **59**(11): 3577-84.

Choung, S. Y., T. Kobayashi, J. Inoue, K. Takemoto, H. Ishitsuka and K. Inoue (1988). Hemolytic activity of a cyclic peptide Ro09-0198 isolated from *Streptoverticillium*. *Biochim Biophys Acta*, **940**(2): 171-9.

Claesen, J. and M. Bibb (2010). Genome mining and genetic analysis of cypemycin biosynthesis reveal an unusual class of posttranslationally modified peptides. *Proc Natl Acad Sci U S A*, **107**(37): 16297-302.

Claessen, D., I. Stokroos, H. J. Deelstra, N. A. Penninga, C. Bormann, J. A. Salas, L. Dijkhuizen and H. A. Wosten (2004). The formation of the rodlet layer of streptomycetes is the result of the interplay between rodlins and chaplins. *Mol Microbiol*, **53**(2): 433-43.

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. *Chem Biol*, **15**(10): 1035-45.

Corsini, G., M. Baeza, O. Monasterio and R. Lagos (2002). The expression of genes involved in microcin maturation regulates the production of active microcin E492. *Biochimie*, **84**(5-6): 539-44.

Cortés, J., A. N. Appleyard and M. J. Dawson (2009). Chapter 22. Whole-cell generation of lantibiotic variants. *Methods Enzymol*, **458**(559-74.

Corvey, C., T. Stein, S. Düsterhus, M. Karas and K. D. Entian (2003). Activation of subtilin precursors by *Bacillus subtilis* extracellular serine proteases subtilisin (AprE), WprA, and Vpr. *Biochem Biophys Res Commun*, **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 lacticin 3147: generating a blueprint for rational drug design. *Mol Microbiol*, **62**(3): 735-47.

Cotter, P. D., C. Hill and R. P. Ross (2005). Bacterial lantibiotics: strategies to improve therapeutic potential. *Curr Protein Pept Sci*, **6**(1): 61-75.

Cotter, P. D., C. Hill and R. P. Ross (2005). Bacteriocins: developing innate immunity for food. *Nat Rev Microbiol*, **3**(10): 777-88.

Cotter, P. D., P. M. O'Connor, L. A. Draper, E. M. Lawton, L. H. Deegan, C. Hill and R. P. Ross (2005). Posttranslational conversion of L-serines to D-alanines is vital for optimal production and activity of the lantibiotic lacticin 3147. *Proc Natl Acad Sci U S A*, **102**(51): 18584-9.

Craik, D. J., N. L. Daly, I. Saska, M. Trabi and K. J. Rosengren (2003). Structures of naturally occurring circular proteins from bacteria. *J Bacteriol*, **185**(14): 4011-21.

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. *J Bacteriol*, **192**(4): 1131-42.

Datsenko, K. A. and B. L. Wanner (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A*, **97**(12): 6640-5.

Datta, V., S. M. Myskowski, L. A. Kwinn, D. N. Chiem, N. Varki, R. G. Kansal, M. Kotb and V. Nizet (2005). Mutational analysis of the group A streptococcal operon encoding streptolysin S and its virulence role in invasive infection. *Mol Microbiol*, **56**(3): 681-95.

de Jong, A., A. J. van Heel, J. Kok and O. P. Kuipers (2010). BAGEL2: mining for bacteriocins in genomic data. *Nucleic Acids Res*, **38 Suppl**(W647-51.

de Lorenzo, V. (1984). Isolation and characterization of microcin E492 from *Klebsiella pneumoniae*. *Arch Microbiol*, **139**(1): 72-5.

Delves-Broughton, J., P. Blackburn, R. J. Evans and J. Hugenholtz (1996). Applications of the bacteriocin, nisin. *Antonie Van Leeuwenhoek*, **69**(2): 193-202.

Destoumieux-Garzon, D., J. Peduzzi, X. Thomas, C. Djediat and S. Rebuffat (2006). Parasitism of iron-siderophore receptors of Escherichia coli by the siderophore-peptide microcin E492m and its unmodified counterpart. *Biometals*, **19**(2): 181-91.

Destoumieux-Garzon, D., X. Thomas, M. Santamaria, C. Goulard, M. Barthelemy, B. Boscher, Y. Bessin, G. Molle, A. M. Pons, L. Letellier, J. Peduzzi and S. Rebuffat (2003). Microcin E492 antibacterial activity:

evidence for a TonB-dependent inner membrane permeabilization on *Escherichia coli*. *Mol Microbiol*, **49**(4): 1031-41.

Diep, D. B. and I. F. Nes (2002). Ribosomally synthesized antibacterial peptides in Gram positive bacteria. *Curr Drug Targets*, **3**(2): 107-22.

Diep, D. B., M. Skaugen, Z. Salehian, H. Holo and I. F. Nes (2007). Common mechanisms of target cell recognition and immunity for class II bacteriocins. *Proc Natl Acad Sci U S A*,

Dischinger, J., M. Josten, C. Szekat, H. G. Sahl and G. Bierbaum (2009). Production of the novel two-peptide lantibiotic lichenicidin by *Bacillus licheniformis* DSM 13. *PLoS One*, **4**(8): e6788.

Dodd, H., M. Gasson, M. Mayer and A. Narbad (2006). Identifying lantibiotic gene clusters and novel lantibiotic genes. Patent number: WO 2006/111743 A2 PCT/GB2006/001429.26.

Dodd, H. M., N. Horn, Z. Hao and M. J. Gasson (1992). A lactococcal expression system for engineered nisins. *Appl Environ Microbiol*, **58**(11): 3683-93.

Donia, M. S., B. J. Hathaway, S. Sudek, M. G. Haygood, M. J. Rosovitz, J. Ravel and E. W. Schmidt (2006). Natural combinatorial peptide libraries in cyanobacterial symbionts of marine ascidians. *Nat Chem Biol*, **2**(12): 729-35.

Donia, M. S., J. Ravel and E. W. Schmidt (2008). A global assembly line for cyanobactins. *Nat Chem Biol*, **4**(6): 341-3.

Dorenbos, R., T. Stein, J. Kabel, C. Bruand, A. Bolhuis, S. Bron, W. J. Quax and J. M. Van Dijl (2002). Thiol-disulfide oxidoreductases are essential for the production of the lantibiotic sublancin 168. *J Biol Chem*, **277**(19): 16682-8.

Doull, J. L., S. W. Ayer, A. K. Singh and P. Thibault (1993). Production of a novel polyketide antibiotic, jadomycin B, by *Streptomyces venezuelae* following heat shock. *J Antibiot (Tokyo)*, **46**(5): 869-71.

Doull, J. L., A. K. Singh, M. Hoare and S. W. Ayer (1994). Conditions for the production of jadomycin B by *Streptomyces venezuelae* ISP5230: effects of heat shock, ethanol treatment and phage infection. *J Ind Microbiol*, **13**(2): 120-5.

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 lacticin 3147. *Mol Microbiol*, **71**(4): 1043-54.

Duquesne, S., D. Destoumieux-Garzon, J. Peduzzi and S. Rebuffat (2007). Microcins, gene-encoded antibacterial peptides from enterobacteria. *Nat Prod Rep*, **24**(4): 708-34.

Duquesne, S., D. Destoumieux-Garzon, S. Zirah, C. Goulard, J. Peduzzi and S. Rebuffat (2007). Two enzymes catalyze the maturation of a lasso peptide in *Escherichia coli*. *Chem Biol*, **14**(7): 793-803.

Ehrlich, J., D. Gottlieb and et al. (1948). *Streptomyces venezuelae*, n. sp., the source of chloromycetin. *J Bacteriol*, **56**(4): 467-77.

Eijsink, V. G., M. B. Brurberg, P. H. Middelhoven and I. F. Nes (1996). Induction of bacteriocin production in *Lactobacillus sake* by a secreted peptide. *J Bacteriol*, **178**(8): 2232-7.

Eijsink, V. G., M. Skeie, P. H. Middelhoven, M. B. Brurberg and I. F. Nes (1998). Comparative studies of class IIa bacteriocins of lactic acid bacteria. *Appl Environ Microbiol*, **64**(9): 3275-81.

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. *Appl Environ Microbiol*, **58**(11): 3730-43.

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 lacticin 3147: analysis of mutants producing reduced antibacterial activities. *J Mol Microbiol Biotechnol*, **13**(4): 226-34.

Field, D., P. M. Connor, P. D. Cotter, C. Hill and R. P. Ross (2008). The generation of nisin variants with enhanced activity against specific gram-positive pathogens. *Mol Microbiol*, **69**(1): 218-30.

Fischbach, M. A. and C. T. Walsh (2006). Assembly-line enzymology for polyketide and nonribosomal Peptide antibiotics: logic, machinery, and mechanisms. *Chem Rev*, **106**(8): 3468-96.

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 Microbiol Lett*, **155**(2): 223-9.

Floriano, B. and M. Bibb (1996). *afsR* is a pleiotropic but conditionally required regulatory gene for antibiotic production in *Streptomyces coelicolor* A3(2). *Mol Microbiol*, **21**(2): 385-96.

Floriano, B., J. L. Ruiz-Barba and R. Jimenez-Diaz (1998). Purification and genetic characterization of enterocin I from *Enterococcus faecium* 6T1a, a novel antilisterial plasmid-encoded bacteriocin which does not belong to the pediocin family of bacteriocins. *Appl Environ Microbiol*, **64**(12): 4883-90.

Foulston, L. C. and M. J. Bibb (2010). Microbisporicin gene cluster reveals unusual features of lantibiotic biosynthesis in actinomycetes. *Proc Natl Acad Sci U S A*, **107**(30): 13461-6.

Franz, C. M., M. J. van Belkum, W. H. Holzapfel, H. Abriouel and A. Galvez (2007). Diversity of enterococcal bacteriocins and their grouping in a new classification scheme. *FEMS Microbiol Rev*, **31**(3): 293-310.

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. *J Antibiot (Tokyo)*, **43**(11): 1403-12.

Fujita, K., S. Ichimasa, T. Zendo, S. Koga, F. Yoneyama, J. Nakayama and K. Sonomoto (2007). Structural analysis and characterization of lacticin Q, a novel bacteriocin belonging to a new family of unmodified bacteriocins of gram-positive bacteria. *Appl Environ Microbiol*, **73**(9): 2871-7.

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): 7352-63.

Gálvez, A., M. Maqueda, E. Valdivia, A. Quesada and E. Montoya (1986). Characterization and partial purification of a broad spectrum antibiotic AS-48 produced by *Streptococcus faecalis*. *Can J Microbiol*, **32**(10): 765-71.

Garrido, M. C., M. Herrero, R. Kolter and F. Moreno (1988). The export of the DNA replication inhibitor Microcin B17 provides immunity for the host cell. *Embo J*, **7**(6): 1853-62.

Genilloud, O., F. Moreno and R. Kolter (1989). DNA sequence, products, and transcriptional pattern of the genes involved in production of the DNA replication inhibitor microcin B17. *J Bacteriol*, **171**(2): 1126-35.

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. *J Bacteriol*, **176**(23): 7335-44.

Gonzalez-Pastor, J. E., J. L. San Millan, M. A. Castilla and F. Moreno (1995). Structure and organization of plasmid genes required to produce the translation inhibitor microcin C7. *J Bacteriol*, **177**(24): 7131-40.

Gonzalez-Pastor, J. E., J. L. San Millan and F. Moreno (1994). The smallest known gene. *Nature*, **369**(6478): 281.

Gonzalez, C., G. M. Langdon, M. Bruix, A. Galvez, E. Valdivia, M. Maqueda and M. Rico (2000). Bacteriocin AS-48, a microbial cyclic polypeptide structurally and functionally related to mammalian NK-lysin. *Proc Natl Acad Sci U S A*, **97**(21): 11221-6.

Gonzalez, D. J., S. W. Lee, M. E. Hensler, A. L. Markley, S. Dahesh, D. A. Mitchell, N. Bandeira, V. Nizet, J. E. Dixon and P. C. Dorrestein (2010). Clostridiolysin S, a Post-translationally Modified Biotoxin from *Clostridium botulinum*. *J Biol Chem*, **285**(36): 28220-8.

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 Biol*, **8**(3): e1000339.

Grasemann, H., F. Stehling, H. Brunar, R. Widmann, T. W. Laliberte, L. Molina, G. Doring and F. Ratjen (2007). Inhalation of Moli1901 in patients with cystic fibrosis. *Chest*, **131**(5): 1461-6.

Guijarro, J. I., J. E. Gonzalez-Pastor, F. Baleux, J. L. San Millan, M. A. Castilla, M. Rico, F. Moreno and M. Delepierre (1995). Chemical structure and translation inhibition studies of the antibiotic microcin C7. *J Biol Chem*, **270**(40): 23520-32.

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*. *Adv Appl Microbiol*, **54**(107-28.

Haas, W., B. D. Shepard and M. S. Gilmore (2002). Two-component regulator of *Enterococcus faecalis* cytolysin responds to quorum-sensing autoinduction. *Nature*, **415**(6867): 84-7.

Halami, P. M., T. Stein, A. Chandrashekar and K. D. Entian (2010). Maturation and processing of SpaI, the lipoprotein involved in subtilin immunity in *Bacillus subtilis* ATCC 6633. *Microbiol Res*, **165**(3): 183-9.

Hara, H., Y. Ohnishi and S. Horinouchi (2009). DNA microarray analysis of global gene regulation by A-factor in *Streptomyces griseus*. *Microbiology*, **155**(Pt 7): 2197-210.

Hasper, H. E., B. de Kruijff and E. Breukink (2004). Assembly and stability of nisin-lipid II pores. *Biochemistry*, **43**(36): 11567-75.

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): 1636-7.

Havarstein, L. S., D. B. Diep and I. F. Nes (1995). A family of bacteriocin ABC transporters carry out proteolytic processing of their substrates concomitant with export. *Mol Microbiol*, **16**(2): 229-40.

He, J., N. Magarvey, M. Piraee and L. C. Vining (2001). The gene cluster for chloramphenicol biosynthesis in *Streptomyces venezuelae* ISP5230 includes novel shikimate pathway homologues and a monomodular non-ribosomal peptide synthetase gene. *Microbiology*, **147**(Pt 10): 2817-29.

Healy, F. G., K. P. Eaton, P. Limsirichai, J. F. Aldrich, A. K. Plowman and R. R. King (2009). Characterization of gamma-butyrolactone autoregulatory signaling gene homologs in the angucyclinone polyketide WS5995B producer *Streptomyces acidiscabies*. *J Bacteriol*, **191**(15): 4786-97.

Héchard, Y. and H. G. Sahl (2002). Mode of action of modified and unmodified bacteriocins from Gram-positive bacteria. *Biochimie*, **84**(5-6): 545-57.

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. *Appl Environ Microbiol*, **64**(9): 3140-6.

Hille, M., S. Kies, F. Götz and A. Peschel (2001). Dual role of GdmH in producer immunity and secretion of the Staphylococcal lantibiotics gallidermin and epidermin. *Appl Environ Microbiol*, **67**(3): 1380-3.

Hindré, T., S. Didelot, J. P. Le Pennec, D. Haras, A. Dufour and K. Vallee-Rehel (2003). Bacteriocin detection from whole bacteria by matrix-assisted

laser desorption ionization-time of flight mass spectrometry. *Appl Environ Microbiol*, **69**(2): 1051-8.

Holck, A., L. Axelsson, S. E. Birkeland, T. Aukrust and H. Blom (1992). Purification and amino acid sequence of sakacin A, a bacteriocin from *Lactobacillus sake* Lb706. *J Gen Microbiol*, **138**(12): 2715-20.

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): 643-51.

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*. *Appl Environ Microbiol*, **72**(9): 5814-21.

Hong, H. J., M. I. Hutchings, L. M. Hill and M. J. Buttner (2005). The role of the novel Fem protein VanK in vancomycin resistance in *Streptomyces coelicolor*. *J Biol Chem*, **280**(13): 13055-61.

Hopwood, D. A., T. Kieser, H. M. Wright and M. J. Bibb (1983). Plasmids, recombination and chromosome mapping in *Streptomyces lividans* 66. *J Gen Microbiol*, **129**(7): 2257-69.

Horn, N., S. Swindell, H. Dodd and M. Gasson (1991). Nisin biosynthesis genes are encoded by a novel conjugative transposon. *Mol Gen Genet*, **228**(1-2): 129-35.

Hosoda, K., M. Ohya, T. Kohno, T. Maeda, S. Endo and K. Wakamatsu (1996). Structure determination of an immunopotentiator peptide, cinnamycin, complexed with lysophosphatidylethanolamine by 1H-NMR1. *J Biochem (Tokyo)*, **119**(2): 226-30.

Hyink, O., M. Balakrishnan and J. R. Tagg (2005). *Streptococcus rattus* strain BHT produces both a class I two-component lantibiotic and a class II bacteriocin. *FEMS Microbiol Lett*, **252**(2): 235-41.

Ibrahim, M., A. Guillot, F. Wessner, F. Algaron, C. Besset, P. Courtin, R. Gardan and V. Monnet (2007). Control of the transcription of a short gene encoding a cyclic peptide in *Streptococcus thermophilus*: a new quorum-sensing system? *J Bacteriol*, **189**(24): 8844-54.

Ikeda, H., J. Ishikawa, A. Hanamoto, M. Shinose, H. Kikuchi, T. Shiba, Y. Sakaki, M. Hattori and S. Omura (2003). Complete genome sequence and comparative analysis of the industrial microorganism Streptomyces avermitilis. *Nat Biotechnol*, **21**(5): 526-31.

In 't Veld, G., A. J. Driessen, J. A. Op den Kamp and W. N. Konings (1991). Hydrophobic membrane thickness and lipid-protein interactions of the leucine transport system of *Lactococcus lactis*. *Biochim Biophys Acta*, **1065**(2): 203-12.

Iwatsuki, M., H. Tomoda, R. Uchida, H. Gouda, S. Hirono and S. Omura (2006). Lariatins, antimycobacterial peptides produced by *Rhodococcus sp.* K01-B0171, have a lasso structure. *J Am Chem Soc*, **128**(23): 7486-91.

Jack, R. W., J. R. Tagg and B. Ray (1995). Bacteriocins of gram-positive bacteria. *Microbiol Rev*, **59**(2): 171-200.

Jacob, F., A. Lwoff, A. Siminovitch and E. Wollman (1953). Définition de quelques termes relatifs à la lysogénie. *Ann Inst Pasteur (Paris)*, **84**(1): 222-4.

Ji, G., R. Beavis and R. P. Novick (1997). Bacterial interference caused by autoinducing peptide variants. *Science*, **276**(5321): 2027-30.

Ji, G., R. C. Beavis and R. P. Novick (1995). Cell density control of staphylococcal virulence mediated by an octapeptide pheromone. *Proc Natl Acad Sci U S A*, **92**(26): 12055-9.

Ji, G., W. Pei, L. Zhang, R. Qiu, J. Lin, Y. Benito, G. Lina and R. P. Novick (2005). *Staphylococcus intermedius* produces a functional agr autoinducing peptide containing a cyclic lactone. *J Bacteriol*, **187**(9): 3139-50.

Jung, G. (1991). Lantibiotics - ribosomally synthesized biologically active polypeptides containing sulfide bridges and α,β-didehydroamino acids. *Angew Chem Int Ed Engl*, **30**(9): 1051-1068.

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. *J Appl Microbiol*,

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. *Eur J Biochem*, **199**(2): 411-5.

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*. *J Bacteriol*, **186**(7): 2206-11.

Kawulka, K., T. Sprules, R. T. McKay, P. Mercier, C. M. Diaper, P. Zuber and J. C. Vederas (2003). Structure of subtilosin A, an antimicrobial peptide from *Bacillus subtilis* with unusual posttranslational modifications linking cysteine sulfurs to alpha-carbons of phenylalanine and threonine. *J Am Chem Soc*, **125**(16): 4726-7.

Keijser, B. J., G. P. van Wezel, G. W. Canters and E. Vijgenboom (2002). Developmental regulation of the *Streptomyces lividans ram* genes: involvement of RamR in regulation of the *ramCSAB* operon. *J Bacteriol*, **184**(16): 4420-9.

Kelemen, G. H. and M. J. Buttner (1998). Initiation of aerial mycelium formation in *Streptomyces*. *Curr Opin Microbiol*, **1**(6): 656-62.

Kellner, R., G. Jung, T. Horner, H. Zahner, N. Schnell, K. D. Entian and F. Gotz (1988). Gallidermin: a new lanthionine-containing polypeptide antibiotic. *Eur J Biochem*, **177**(1): 53-9.

Kelly, W. L., L. Pan and C. Li (2009). Thiostrepton biosynthesis: prototype for a new family of bacteriocins. *J Am Chem Soc*, **131**(12): 4327-34.

Kemperman, R., A. Kuipers, H. Karsens, A. Nauta, O. Kuipers and J. Kok (2003). Identification and characterization of two novel clostridial bacteriocins, circularin A and closticin 574. *Appl Environ Microbiol*, **69**(3): 1589-97.

Kido, Y., T. Hamakado, T. Yoshida, M. Anno, Y. Motoki, T. Wakamiya and T. Shiba (1983). Isolation and characterization of ancovenin, a new inhibitor of angiotensin I converting enzyme, produced by actinomycetes. *J Antibiot (Tokyo)*, **36**(10): 1295-9.

Kieser, T., M. J. Bibb, M. J. Buttner, K. F. Chater and D. A. Hopwood (2000). Practical *Streptomyces* genetics. *The John Innes Foundation, Norfolk, U.K.*

Kimura, K., F. Kanou, H. Takahashi, Y. Esumi, M. Uramoto and M. Yoshihama (1997). Propeptin, a new inhibitor of prolyl endopeptidase produced by *Microbispora*. I. Fermentation, isolation and biological properties. *J Antibiot (Tokyo)*, **50**(5): 373-8.

Klaenhammer, T. R. (1993). Genetics of bacteriocins produced by lactic acid bacteria. *FEMS Microbiol Rev*, **12**(1-3): 39-85.

Kleerebezem, M. (2004). Quorum sensing control of lantibiotic production; nisin and subtilin autoregulate their own biosynthesis. *Peptides*, **25**(9): 1405-14.

Kleerebezem, M., R. Bongers, G. Rutten, W. M. de Vos and O. P. Kuipers (2004). Autoregulation of subtilin biosynthesis in *Bacillus subtilis*: the role of the spa-box in subtilin-responsive promoters. *Peptides*, **25**(9): 1415-24.

Kluskens, L. D., A. Kuipers, R. Rink, E. de Boef, S. Fekken, A. J. 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): 12827-34.

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*. *Proc Natl Acad Sci U S A*, **101**(31): 11448-53.

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. *Mol Microbiol*, **58**(5): 1368-80.

Koglin, A. and C. T. Walsh (2009). Structural insights into nonribosomal peptide enzymatic assembly lines. *Nat Prod Rep*, **26**(8): 987-1000.

Komiyama, K., K. Otoguro, 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. *J Antibiot (Tokyo)*, **46**(11): 1666-71.

Koponen, O., M. Tolonen, M. Qiao, G. Wahlström, J. Helin and P. E. Saris (2002). NisB is required for the dehydration and NisC for the lanthionine formation in the post-translational modification of nisin. *Microbiology*, **148**(Pt 11): 3561-8.

Kordel, M., R. Benz and H. G. Sahl (1988). Mode of action of the staphylococcinlike peptide Pep 5: voltage-dependent depolarization of bacterial and artificial membranes. *J Bacteriol*, **170**(1): 84-8.

Kozbial, P. Z. and A. R. Mushegian (2005). Natural history of S-adenosylmethionine-binding proteins. *BMC Struct Biol*, **5**(19.

Krull, R. E., P. Chen, J. Novak, M. Kirk, S. Barnes, J. Baker, N. R. Krishna and P. W. Caufield (2000). Biochemical structural analysis of the lantibiotic mutacin II. *J Biol Chem*, **275**(21): 15845-50.

Kuipers, A., E. de Boef, R. Rink, S. Fekken, L. D. Kluskens, A. J. 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. *J Biol Chem*, **279**(21): 22176-82.

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. *J Biol Chem*, **270**(45): 27299-304.

Kuipers, O. P., H. S. Rollema, W. M. de Vos and R. J. Siezen (1993). 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 Lett*, **330**(1): 23-7.

Kuipers, O. P., H. S. Rollema, W. M. Yap, H. J. Boot, R. J. Siezen and W. M. de Vos (1992). Engineering dehydrated amino acid residues in the antimicrobial peptide nisin. *J Biol Chem*, **267**(34): 24340-6.

Kupke, T., P. Hernandez-Acosta, S. Steinbacher and F. A. Culianez-Macia (2001). Arabidopsis thaliana flavoprotein AtHAL3a catalyzes the decarboxylation of 4'-Phosphopantothenoylcysteine to 4'-phosphopantetheine, a key step in coenzyme A biosynthesis. *J Biol Chem*, **276**(22): 19190-6.

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. *J Biol Chem*, **269**(8): 5653-9.

Kupke, T., C. Kempter, G. Jung and F. Gotz (1995). Oxidative decarboxylation of peptides catalyzed by flavoprotein EpiD. Determination of substrate specificity using peptide libraries and neutral loss mass spectrometry. *J Biol Chem*, **270**(19): 11282-9.

Kupke, T., S. Stevanovic, B. Ottenwalder, J. W. Metzger, G. Jung and F. Gotz (1993). Purification and characterization of EpiA, the peptide substrate for

post-translational modifications involved in epidermin biosynthesis. *FEMS Microbiol Lett*, **112**(1): 43-8.

Kupke, T., S. Stevanovic, H. G. Sahl and F. Gotz (1992). Purification and characterization of EpiD, a flavoprotein involved in the biosynthesis of the lantibiotic epidermin. *J Bacteriol*, **174**(16): 5354-61.

Kupke, T., M. Uebele, D. Schmid, G. Jung, M. Blaesse and S. Steinbacher (2000). Molecular characterization of lantibiotic-synthesizing enzyme EpiD reveals a function for bacterial Dfp proteins in coenzyme A biosynthesis. *J Biol Chem*, **275**(41): 31838-46.

Lagos, R., M. Baeza, G. Corsini, C. Hetz, E. Strahsburger, J. A. Castillo, C. Vergara and O. Monasterio (2001). Structure, organization and characterization of the gene cluster involved in the production of microcin E492, a channel-forming bacteriocin. *Mol Microbiol*, **42**(1): 229-43.

Lawen, A. and R. Zocher (1990). Cyclosporin synthetase. The most complex peptide synthesizing multienzyme polypeptide so far described. *J Biol Chem*, **265**(19): 11355-60.

Lawton, E. M., P. D. Cotter, C. Hill and R. P. Ross (2007). Identification of a novel two-peptide lantibiotic, Haloduracin, produced by the alkaliphile *Bacillus halodurans* C-125. *FEMS Microbiol Lett*, **267**(1): 64-71.

Lawton, E. M., R. P. Ross, C. Hill and P. D. Cotter (2007). Two-peptide lantibiotics: a medical perspective. *Mini Rev Med Chem*, **7**(12): 1236-47.

Lazzarini, A., L. Gastaldo, P. Candiani, I. Ciciliato, D. Losi, F. Marinelli, E. Selva and F. Parenti (2005). Antibiotics 107891, its factors A1 and A2, pharmaceutically acceptable salts and compositions, and use thereof. Patent number: WO 2005/014628 A.

Lee, S. W., D. A. Mitchell, A. L. Markley, M. E. Hensler, D. Gonzalez, A. Wohlrab, P. C. Dorrestein, V. Nizet and J. E. Dixon (2008). Discovery of a widely distributed toxin biosynthetic gene cluster. *Proc Natl Acad Sci U S A*, **105**(15): 5879-84.

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*. *Proc Natl Acad Sci U S A*, **88**(6): 2461-5.

Lewis, R. J. (2009). Conotoxin venom peptide therapeutics. *Adv Exp Med Biol*, **655**(44-8.

Li, B., L. E. Cooper and W. A. van der Donk (2009). Chapter 21. In vitro studies of lantibiotic biosynthesis. *Methods Enzymol*, **458**(533-58.

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. *Proc Natl Acad Sci U S A*, **107**(23): 10430-5.

Li, B. and W. A. van der Donk (2007). Identification of essential catalytic residues of the cyclase NisC involved in the biosynthesis of nisin. *J Biol Chem*, **282**(29): 21169-75.

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): 1464-7.

Li, C. and W. L. Kelly (2010). Recent advances in thiopeptide antibiotic biosynthesis. *Nat Prod Rep*, **27**(2): 153-64.

Li, Y. M., J. C. Milne, L. L. Madison, R. Kolter and C. T. Walsh (1996). From peptide precursors to oxazole and thiazole-containing peptide antibiotics: microcin B17 synthase. *Science*, **274**(5290): 1188-93.

Liao, R., L. Duan, C. Lei, H. Pan, Y. Ding, Q. Zhang, D. Chen, B. Shen, Y. Yu and W. Liu (2009). Thiopeptide biosynthesis featuring ribosomally synthesized precursor peptides and conserved posttranslational modifications. *Chem Biol*, **16**(2): 141-7.

Liu, G., J. Zhong, J. Ni, M. Chen, H. Xiao and L. Huan (2009). Characteristics of the bovicin HJ50 gene cluster in *Streptococcus bovis* HJ50. *Microbiology*, **155**(Pt 2): 584-93.

Liu, W. and J. N. Hansen (1992). Enhancement of the chemical and antimicrobial properties of subtilin by site-directed mutagenesis. *J Biol Chem*, **267**(35): 25078-85.

Lombardi, A., B. D'Agostino, F. Nastri, L. D. D'Andrea, A. Filippelli, M. Falciani, F. Rossi and V. Pavone (1998). A novel super-potent neurokinin A receptor antagonist containing dehydroalanine. *Bioorg Med Chem Lett*, **8**(10): 1153-6.

Ma, H. and K. Kendall (1994). Cloning and analysis of a gene cluster from *Streptomyces coelicolor* that causes accelerated aerial mycelium formation in *Streptomyces lividans*. *J Bacteriol*, **176**(12): 3800-11.

Machaidze, G. and J. Seelig (2003). Specific binding of cinnamycin (Ro 09-0198) to phosphatidylethanolamine. Comparison between micellar and membrane environments. *Biochemistry*, **42**(43): 12570-6.

Machaidze, G., A. Ziegler and J. Seelig (2002). Specific binding of Ro 09-0198 (cinnamycin) to phosphatidylethanolamine: a thermodynamic analysis. *Biochemistry*, **41**(6): 1965-71.

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): 61-8.

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. *J Nat Prod*, **72**(4): 605-7.

Magnuson, R., J. Solomon and A. D. Grossman (1994). Biochemical and genetic characterization of a competence pheromone from *B. subtilis*. *Cell*, **77**(2): 207-16.

Majchrzykiewicz, J. A., J. Lubelski, G. N. Moll, A. Kuipers, J. J. Bijlsma, O. P. Kuipers and R. Rink (2010). Production of a class II two-component lantibiotic of *Streptococcus pneumoniae* using the class I nisin synthetic machinery and leader sequence. *Antimicrob Agents Chemother*, **54**(4): 1498-505.

Majer, F., D. G. Schmid, K. Altena, G. Bierbaum and T. Kupke (2002). The flavoprotein MrsD catalyzes the oxidative decarboxylation reaction involved in formation of the peptidoglycan biosynthesis inhibitor mersacidin. *J Bacteriol*, **184**(5): 1234-43.

Maqueda, M., M. Sanchez-Hidalgo, M. Fernandez, M. Montalban-Lopez, E. Valdivia and M. Martinez-Bueno (2008). Genetic features of circular bacteriocins produced by Gram-positive bacteria. *FEMS Microbiol Rev*, **32**(1): 2-22.

Marcille, F., A. Gomez, P. Joubert, M. Ladire, G. Veau, A. Clara, F. Gavini, A. Willems and M. Fons (2002). Distribution of genes encoding the trypsin-

dependent lantibiotic ruminococcin A among bacteria isolated from human fecal microbiota. *Appl Environ Microbiol*, **68**(7): 3424-31.

Martin-Visscher, L. A., X. Gong, M. Duszyk and J. C. Vederas (2009). The three-dimensional structure of carnocyclin A reveals that many circular bacteriocins share a common structural motif. *J Biol Chem*, **284**(42): 28674-81.

Martin-Visscher, L. A., M. J. van Belkum, S. Garneau-Tsodikova, R. M. Whittal, J. Zheng, L. M. McMullen and J. C. Vederas (2008). Isolation and characterization of carnocyclin a, a novel circular bacteriocin produced by *Carnobacterium maltaromaticum* UAL307. *Appl Environ Microbiol*, **74**(15): 4756-63.

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): 3049-56.

Martinez-Bueno, M., M. Maqueda, A. Galvez, B. Samyn, J. Van Beeumen, J. Coyette and E. Valdivia (1994). Determination of the gene sequence and the molecular structure of the enterococcal peptide antibiotic AS-48. *J Bacteriol*, **176**(20): 6334-9.

McAuliffe, O., C. Hill and R. P. Ross (2000). Each peptide of the two-component lantibiotic lacticin 3147 requires a separate modification enzyme for activity. *Microbiology*, **146 (Pt 9)**(2147-54.

McAuliffe, O., R. P. Ross and C. Hill (2001). Lantibiotics: structure, biosynthesis and mode of action. *FEMS Microbiol Rev*, **25**(3): 285-308.

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. *Proc Natl Acad Sci U S A*, **103**(46): 17243-8.

McGhee, C. N. and C. N. Anastas (1996). Widespread ocular use of topical chloramphenicol: is there justifiable concern regarding idiosyncratic aplastic anaemia? *Br J Ophthalmol*, **80**(2): 182-4.

McIntosh, J. A., M. S. Donia and E. W. Schmidt (2009). Ribosomal peptide natural products: bridging the ribosomal and nonribosomal worlds. *Nat Prod Rep*, **26**(4): 537-59.

McIntosh, J. A., M. S. Donia and E. W. Schmidt (2010). Insights into heterocyclization from two highly similar enzymes. *J Am Chem Soc*, **132**(12): 4089-91.

Medini, D., D. Serruto, J. Parkhill, D. A. Relman, C. Donati, R. Moxon, S. Falkow and R. Rappuoli (2008). Microbiology in the post-genomic era. *Nat Rev Microbiol*, **6**(6): 419-30.

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. *Angew Chem Int Ed Engl*, **49**(6): 1151-4.

Metlitskaya, A., T. Kazakov, A. Kommer, O. Pavlova, M. Praetorius-Ibba, M. Ibba, I. Krasheninnikov, V. Kolb, I. Khmel and K. Severinov (2006). Aspartyl-tRNA synthetase is the target of peptide nucleotide antibiotic Microcin C. *J Biol Chem*, **281**(26): 18033-42.

Meyer, C., G. Bierbaum, C. Heidrich, M. Reis, J. Süling, 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. *Eur J Biochem*, **232**(2): 478-89.

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. *Anal Biochem*, **223**(2): 185-90.

Miller, L. M., C. Chatterjee, W. A. van der Donk and N. L. Kelleher (2006). The dehydratase activity of lacticin 481 synthetase is highly processive. *J Am Chem Soc*, **128**(5): 1420-1.

Milne, J. C., A. C. Eliot, N. L. Kelleher and C. T. Walsh (1998). ATP/GTP hydrolysis is required for oxazole and thiazole biosynthesis in the peptide antibiotic microcin B17. *Biochemistry*, **37**(38): 13250-61.

Milne, J. C., R. S. Roy, A. C. Eliot, N. L. Kelleher, A. Wokhlu, B. Nickels and C. T. Walsh (1999). Cofactor requirements and reconstitution of microcin B17 synthetase: a multienzyme complex that catalyzes the formation of oxazoles and thiazoles in the antibiotic microcin B17. *Biochemistry*, **38**(15): 4768-81.

Minami, Y., K. Yoshida, R. Azuma, A. Urakawa, T. Kawauchi, T. Otani, K. Komiyama and S. Omura (1994). Structure of cypemycin, a new peptide antibiotic. *Tetrahedron Lett*, **35**(43): 8001-8004.

Mitchell, D. A., S. W. Lee, M. A. Pence, A. L. Markley, J. D. Limm, V. Nizet and J. E. Dixon (2009). Structural and functional dissection of the heterocyclic peptide cytotoxin streptolysin S. *J Biol Chem*, **284**(19): 13004-12.

Montville, T. J. and Y. Chen (1998). Mechanistic action of pediocin and nisin: recent progress and unresolved questions. *Appl Microbiol Biotechnol*, **50**(5): 511-9.

Morishita, Y., S. Chiba, E. Tsukuda, T. Tanaka, T. Ogawa, M. Yamasaki, M. Yoshida, I. Kawamoto and Y. Matsuda (1994). RES-701-1, a novel and selective endothelin type B receptor antagonist produced by *Streptomyces sp.* RE-701. I. Characterization of producing strain, fermentation, isolation, physico-chemical and biological properties. *J Antibiot (Tokyo)*, **47**(3): 269-75.

Morris, R. P., J. A. Leeds, H. U. Naegeli, L. Oberer, K. Memmert, E. Weber, M. J. LaMarche, C. N. Parker, N. Burrer, S. Esterow, A. E. Hein, E. K. Schmitt and P. Krastel (2009). Ribosomally synthesized thiopeptide antibiotics targeting elongation factor Tu. *J Am Chem Soc*, **131**(16): 5946-55.

Müller, 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. *Angew Chem Int Ed Engl*, **49**(13): 2436-40.

Nagao, J., Y. Aso, K. Shioya, J. Nakayama and K. Sonomoto (2007). Lantibiotic engineering: molecular characterization and exploitation of lantibiotic-synthesizing enzymes for peptide engineering. *J Mol Microbiol Biotechnol*, **13**(4): 235-42.

Nakayama, J., Y. Cao, T. Horii, S. Sakuda, A. D. Akkermans, W. M. de Vos and H. Nagasawa (2001). Gelatinase biosynthesis-activating pheromone: a peptide lactone that mediates a quorum sensing in *Enterococcus faecalis*. *Mol Microbiol*, **41**(1): 145-54.

Nakayama, J., S. Chen, N. Oyama, K. Nishiguchi, E. A. Azab, E. Tanaka, R. Kariyama and K. Sonomoto (2006). Revised model for *Enterococcus*

faecalis fsr quorum-sensing system: the small open reading frame *fsrD* encodes the gelatinase biosynthesis-activating pheromone propeptide corresponding to staphylococcal *agrD*. *J Bacteriol*, **188**(23): 8321-6.

Naruse, N., O. Tenmyo, K. Tomita, M. Konishi, T. Miyaki, H. Kawaguchi, K. Fukase, T. Wakamiya and T. Shiba (1989). Lanthiopeptin, a new peptide antibiotic. Production, isolation and properties of lanthiopeptin. *J Antibiot (Tokyo)*, **42**(6): 837-45.

Navaratna, M. A., H. G. Sahl and J. R. Tagg (1998). Two-component anti-*Staphylococcus aureus* lantibiotic activity produced by *Staphylococcus aureus* C55. *Appl Environ Microbiol*, **64**(12): 4803-8.

Navaratna, M. A., H. G. Sahl and J. R. Tagg (1999). Identification of genes encoding two-component lantibiotic production in *Staphylococcus aureus* C55 and other phage group II *S. aureus* strains and demonstration of an association with the exfoliative toxin B gene. *Infect Immun*, **67**(8): 4268-71.

Nes, I. F. and J. R. Tagg (1996). Novel lantibiotics and their pre-peptides. *Antonie Van Leeuwenhoek*, **69**(2): 89-97.

Netz, D. J., R. Pohl, A. G. Beck-Sickinger, T. Selmer, A. J. Pierik, C. Bastos Mdo and H. G. Sahl (2002). Biochemical characterisation and genetic analysis of aureocin A53, a new, atypical bacteriocin from *Staphylococcus aureus*. *J Mol Biol*, **319**(3): 745-56.

Nissen-Meyer, J., H. Holo, L. S. Havarstein, K. Sletten and I. F. Nes (1992). A novel lactococcal bacteriocin whose activity depends on the complementary action of two peptides. *J Bacteriol*, **174**(17): 5686-92.

Nolan, E. M., M. A. Fischbach, A. Koglin and C. T. Walsh (2007). Biosynthetic tailoring of microcin E492m: post-translational modification affords an antibacterial siderophore-peptide conjugate. *J Am Chem Soc*, **129**(46): 14336-47.

Nolan, E. M. and C. T. Walsh (2008). Investigations of the MceIJ-catalyzed posttranslational modification of the microcin E492 C-terminus: linkage of ribosomal and nonribosomal peptides to form "trojan horse" antibiotics. *Biochemistry*, **47**(35): 9289-99.

Novoa, M. A., L. Diaz-Guerra, J. L. San Millan and F. Moreno (1986). Cloning and mapping of the genetic determinants for microcin C7 production and immunity. *J Bacteriol*, **168**(3): 1384-91.

O'Connor, T. J., P. Kanellis and J. R. Nodwell (2002). The *ramC* gene is required for morphogenesis in *Streptomyces coelicolor* and expressed in a cell type-specific manner under the direct control of RamR. *Mol Microbiol*, **45**(1): 45-57.

Ohnishi, Y., J. Ishikawa, H. Hara, H. Suzuki, M. Ikenoya, H. Ikeda, A. Yamashita, M. Hattori and S. Horinouchi (2008). Genome sequence of the streptomycin-producing microorganism *Streptomyces griseus* IFO 13350. *J Bacteriol*, **190**(11): 4050-60.

Ohnishi, Y., H. Yamazaki, J. Y. Kato, A. Tomono and S. Horinouchi (2005). AdpA, a central transcriptional regulator in the A-factor regulatory cascade that leads to morphological development and secondary metabolism in *Streptomyces griseus*. *Biosci Biotechnol Biochem*, **69**(3): 431-9.

Okada, M., I. Sato, S. J. Cho, H. Iwata, T. Nishio, D. Dubnau and Y. Sakagami (2005). Structure of the *Bacillus subtilis* quorum-sensing peptide pheromone ComX. *Nat Chem Biol*, **1**(1): 23-4.

Okamoto, S., A. Lezhava, T. Hosaka, Y. Okamoto-Hosoya and K. Ochi (2003). Enhanced expression of S-adenosylmethionine synthetase causes overproduction of actinorhodin in *Streptomyces coelicolor* A3(2). *J Bacteriol*, **185**(2): 601-9.

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): 13613-24.

Oman, T. J. and W. A. van der Donk (2010). Follow the leader: the use of leader peptides to guide natural product biosynthesis. *Nat Chem Biol*, **6**(1): 9-18.

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**(12): 3923-33.

Onaka, H., H. Tabata, Y. Igarashi, Y. Sato and T. Furumai (2001). Goadsporin, a chemical substance which promotes secondary metabolism and morphogenesis in streptomycetes. I. Purification and characterization. *J Antibiot (Tokyo)*, **54**(12): 1036-44.

Ottenwalder, B., T. Kupke, S. Brecht, V. Gnau, J. Metzger, G. Jung and F. Gotz (1995). Isolation and characterization of genetically engineered gallidermin and epidermin analogs. *Appl Environ Microbiol*, **61**(11): 3894-903.

Otto, M., A. Peschel and F. Götz (1998). Producer self-protection against the lantibiotic epidermin by the ABC transporter EpiFEG of *Staphylococcus epidermidis* Tu3298. *FEMS Microbiol Lett*, **166**(2): 203-11.

Otto, M., R. Sussmuth, G. Jung and F. Gotz (1998). Structure of the pheromone peptide of the *Staphylococcus epidermidis agr* system. *FEBS Lett*, **424**(1-2): 89-94.

Pag, U. and H. G. Sahl (2002). Multiple activities in lantibiotics--models for the design of novel antibiotics? *Curr Pharm Des*, **8**(9): 815-33.

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). *J Bacteriol*, **181**(1): 204-11.

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. *J Biol Chem*, **273**(36): 23134-42.

Papadelli, M., A. Karsioti, R. Anastasiou, M. Georgalaki and E. Tsakalidou (2007). Characterization of the gene cluster involved in the biosynthesis of macedocin, the lantibiotic produced by *Streptococcus macedonicus*. *FEMS Microbiol Lett*, **272**(1): 75-82.

Parenti, F., H. Pagani and G. Beretta (1976). Gardimycin, a new antibiotic from *Actinoplanes*. I. Description of the producer strain and fermentation studies. *J Antibiot (Tokyo)*, **29**(5): 501-6.

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 lacticin 481. *Biochemistry*, **47**(28): 7342-51.

Philmus, B., G. Christiansen, W. Y. Yoshida and T. K. Hemscheidt (2008). Post-translational modification in microviridin biosynthesis. *Chembiochem*, **9**(18): 3066-73.

Philmus, B., J. P. Guerrette and T. K. Hemscheidt (2009). Substrate specificity and scope of MvdD, a GRASP-like ligase from the microviridin biosynthetic gene cluster. *ACS Chem Biol*, **4**(6): 429-34.

Portmann, C., J. F. Blom, K. Gademann and F. Juttner (2008). Aerucyclamides A and B: isolation and synthesis of toxic ribosomal heterocyclic peptides

from the cyanobacterium *Microcystis aeruginosa* PCC 7806. *J Nat Prod*, **71**(7): 1193-6.

Portmann, C., J. F. Blom, M. Kaiser, R. Brun, F. Juttner and K. Gademann (2008). Isolation of Aerucyclamides C and D and Structure Revision of Microcyclamide 7806A: Heterocyclic Ribosomal Peptides from *Microcystis aeruginosa* PCC 7806 and Their Antiparasite Evaluation. *J Nat Prod*,

Qi, F., P. Chen and P. W. Caufield (1999). Purification of mutacin III from group III *Streptococcus mutans* UA787 and genetic analyses of mutacin III biosynthesis genes. *Appl Environ Microbiol*, **65**(9): 3880-7.

Qi, F., P. Chen and P. W. Caufield (2000). Purification and biochemical characterization of mutacin I from the group I strain of *Streptococcus mutans*, CH43, and genetic analysis of mutacin I biosynthesis genes. *Appl Environ Microbiol*, **66**(8): 3221-9.

Qiao, M., T. Immonen, O. Koponen and P. E. Saris (1995). The cellular location and effect on nisin immunity of the NisI protein from *Lactococcus lactis* N8 expressed in *Escherichia coli* and *L. lactis*. *FEMS Microbiol Lett*, **131**(1): 75-80.

Qiu, R., W. Pei, L. Zhang, J. Lin and G. Ji (2005). Identification of the putative staphylococcal AgrB catalytic residues involving the proteolytic cleavage of AgrD to generate autoinducing peptide. *J Biol Chem*, **280**(17): 16695-704.

Quadri, L. E., M. Sailer, M. R. Terebiznik, K. L. Roy, J. C. Vederas and M. E. Stiles (1995). Characterization of the protein conferring immunity to the antimicrobial peptide carnobacteriocin B2 and expression of carnobacteriocins B2 and BM1. *J Bacteriol*, **177**(5): 1144-51.

Ra, S. R., M. Qiao, T. Immonen, I. Pujana and E. J. Saris (1996). Genes responsible for nisin synthesis, regulation and immunity form a regulon of two operons and are induced by nisin in *Lactoccocus lactis* N8. *Microbiology*, **142 (Pt 5)**(1281-8.

Ramnath, M., M. Beukes, K. Tamura and J. W. Hastings (2000). Absence of a putative mannose-specific phosphotransferase system enzyme IIAB component in a leucocin A-resistant strain of *Listeria monocytogenes*, as shown by two-dimensional sodium dodecyl sulfate-polyacrylamide gel electrophoresis. *Appl Environ Microbiol*, **66**(7): 3098-101.

Rea, M. C., C. S. Sit, E. Clayton, P. M. O'Connor, R. M. Whittal, J. Zheng, J. C. Vederas, R. P. Ross and C. Hill (2010). Thuricin CD, a posttranslationally modified bacteriocin with a narrow spectrum of activity against *Clostridium difficile*. *Proc Natl Acad Sci U S A*, **107**(20): 9352-7.

Reis, M., M. Eschbach-Bludau, M. I. Iglesias-Wind, T. Kupke and H. G. Sahl (1994). Producer immunity towards the lantibiotic Pep5: identification of the immunity gene pepI and localization and functional analysis of its gene product. *Appl Environ Microbiol*, **60**(8): 2876-83.

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 Rep*, **9**(7): 670-5.

Rink, R., A. Arkema-Meter, I. Baudoin, E. Post, A. Kuipers, S. A. Nelemans, M. H. Akanbi and G. N. Moll (2010). To protect peptide pharmaceuticals against peptidases. *J Pharmacol Toxicol Methods*, **61**(2): 210-8.

Rink, R., A. Kuipers, E. de Boef, K. J. Leenhouts, A. J. 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): 8873-82.

Rink, R., J. Wierenga, A. Kuipers, L. D. Kluskens, A. J. Driessen, O. P. Kuipers and G. N. Moll (2007). Dissection and modulation of the four distinct activities of nisin by mutagenesis of rings A and B and by C-terminal truncation. *Appl Environ Microbiol*, **73**(18): 5809-16.

Rink, R., J. Wierenga, A. Kuipers, L. D. Kluskens, A. J. Driessen, O. P. Kuipers and G. N. Moll (2007). Production of dehydroamino acid-containing peptides by *Lactococcus lactis*. *Appl Environ Microbiol*,

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. *Appl Environ Microbiol*, **59**(7): 2014-21.

Roush, R. F., E. M. Nolan, F. Lohr and C. T. Walsh (2008). Maturation of an *Escherichia coli* ribosomal peptide antibiotic by ATP-consuming N-P bond formation in microcin C7. *J Am Chem Soc*, **130**(11): 3603-9.

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, lacticin 3147. *J Biol Chem*, **274**(53): 37544-50.

Ryan, M. P., M. C. Rea, C. Hill and R. P. Ross (1996). An application in cheddar cheese manufacture for a strain of *Lactococcus lactis* producing a novel broad-spectrum bacteriocin, lacticin 3147. *Appl Environ Microbiol*, **62**(2): 612-9.

Ryttersgaard, C., S. C. Griffith, M. R. Sawaya, D. C. MacLaren, S. Clarke and T. O. Yeates (2002). Crystal structure of human L-isoaspartyl methyltransferase. *J Biol Chem*, **277**(12): 10642-6.

Sahl, H. G., M. Kordel and R. Benz (1987). Voltage-dependent depolarization of bacterial membranes and artificial lipid bilayers by the peptide antibiotic nisin. *Arch Microbiol*, **149**(2): 120-4.

Salomon, R. A. and R. N. Farias (1992). Microcin 25, a novel antimicrobial peptide produced by *Escherichia coli*. *J Bacteriol*, **174**(22): 7428-35.

Sambrook, J., P. MacCallum and D. Russell (2001). Molecular cloning: a laboratory manual, third edition. *Cold Spring Harbor Laboratory Press*.

Samyn, B., M. Martinez-Bueno, B. Devreese, M. Maqueda, A. Galvez, E. Valdivia, J. Coyette and J. Van Beeumen (1994). The cyclic structure of the enterococcal peptide antibiotic AS-48. *FEBS Lett*, **352**(1): 87-90.

Schmidt, E. W., J. T. Nelson, D. A. Rasko, S. Sudek, J. A. Eisen, M. G. Haygood and J. Ravel (2005). Patellamide A and C biosynthesis by a microcin-like pathway in *Prochloron didemni*, the cyanobacterial symbiont of *Lissoclinum patella*. *Proc Natl Acad Sci U S A*, **102**(20): 7315-20.

Schmitz, S., A. Hoffmann, C. Szekat, B. Rudd and G. Bierbaum (2006). The lantibiotic mersacidin is an autoinducing peptide. *Appl Environ Microbiol*, **72**(11): 7270-7.

Schneider, T. R., J. Karcher, E. Pohl, P. Lubini and G. M. Sheldrick (2000). *Ab initio* structure determination of the lantibiotic mersacidin. *Acta Crystallogr D Biol Crystallogr*, **56**(Pt 6): 705-13.

Schnell, N., G. Engelke, J. Augustin, R. Rosenstein, V. Ungermann, F. Gotz and K. D. Entian (1992). Analysis of genes involved in the biosynthesis of lantibiotic epidermin. *Eur J Biochem*, **204**(1): 57-68.

Schnell, N., K. D. Entian, U. Schneider, F. Götz, H. Zähner, R. Kellner and G. Jung (1988). Prepeptide sequence of epidermin, a ribosomally synthesized antibiotic with four sulphide-rings. *Nature*, **333**(6170): 276-8.

Schultz, J. (2004). HTTM, a horizontally transferred transmembrane domain. *Trends Biochem Sci*, **29**(1): 4-7.

Scott, J. C., H. G. Sahl, A. Carne and J. R. Tagg (1992). Lantibiotic-mediated anti-lactobacillus activity of a vaginal *Staphylococcus aureus* isolate. *FEMS Microbiol Lett*, **72**(1): 97-102.

Sechi, S. and B. T. Chait (1998). Modification of cysteine residues by alkylation. A tool in peptide mapping and protein identification. *Anal Chem*, **70**(24): 5150-8.

Shenkarev, Z. O., E. I. Finkina, E. K. Nurmukhamedova, S. V. Balandin, K. S. Mineev, K. D. Nadezhdin, Z. A. Yakimenko, A. A. Tagaev, Y. V. Temirov, A. S. Arseniev and T. V. Ovchinnikova (2010). Isolation, structure elucidation, and synergistic antibacterial activity of a novel two-component lantibiotic lichenicidin from *Bacillus licheniformis* VK21. *Biochemistry*, **49**(30): 6462-72.

Sheth, T. R., R. M. Henderson, S. B. Hladky and A. W. Cuthbert (1992). Ion channel formation by duramycin. *Biochim Biophys Acta*, **1107**(1): 179-85.

Shin, S. K., D. Xu, H. J. Kwon and J. W. Suh (2006). S-adenosylmethionine activates *adpA* transcription and promotes streptomycin biosynthesis in *Streptomyces griseus*. *FEMS Microbiol Lett*, **259**(1): 53-9.

Shirling, E. B. and D. Gottlieb (1972). Cooperative descriptions of type strains of *Streptomyces*. V. Additional descriptions. *Int J Syst Bacteriol*, **22**(1): 265-394.

Siezen, R. J., O. P. Kuipers and W. M. de Vos (1996). Comparison of lantibiotic gene clusters and encoded proteins. *Antonie Van Leeuwenhoek*, **69**(2): 171-84.

Simmaco, M., G. Kreil and D. Barra (2009). Bombinins, antimicrobial peptides from *Bombina* species. *Biochim Biophys Acta*, **1788**(8): 1551-5.

Skaugen, M., C. I. Abildgaard and I. F. Nes (1997). Organization and expression of a gene cluster involved in the biosynthesis of the lantibiotic lactocin S. *Mol Gen Genet*, **253**(6): 674-86.

Skaugen, M., J. Nissen-Meyer, G. Jung, S. Stevanovic, K. Sletten, C. Inger, M. Abildgaard and I. F. Nes (1994). *In vivo* conversion of L-serine to D-alanine in a ribosomally synthesized polypeptide. *J Biol Chem*, **269**(44): 27183-5.

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. *Eur J Biochem*, **267**(23): 6810-6.

Solbiati, J. O., M. Ciaccio, R. N. Farias, J. E. Gonzalez-Pastor, F. Moreno and R. A. Salomon (1999). Sequence analysis of the four plasmid genes required to produce the circular peptide antibiotic microcin J25. *J Bacteriol*, **181**(8): 2659-62.

Somma, S., W. Merati and F. Parenti (1977). Gardimycin, a new antibiotic inhibiting peptidoglycan synthesis. *Antimicrob Agents Chemother*, **11**(3): 396-401.

Stein, T., S. Borchert, P. Kiesau, S. Heinzmann, S. Kloss, C. Klein, M. Helfrich and K. D. Entian (2002). Dual control of subtilin biosynthesis and immunity in *Bacillus subtilis*. *Mol Microbiol*, **44**(2): 403-16.

Stein, T., S. Heinzmann, S. Dusterhus, S. Borchert and K. D. Entian (2005). Expression and functional analysis of the subtilin immunity genes *spaIFEG* in the subtilin-sensitive host *Bacillus subtilis* MO1099. *J Bacteriol*, **187**(3): 822-8.

Stein, T., S. Heinzmann, I. Solovieva and K. D. Entian (2003). Function of *Lactococcus lactis* nisin immunity genes *nisI* and *nisFEG* after coordinated expression in the surrogate host *Bacillus subtilis*. *J Biol Chem*, **278**(1): 89-94.

Steinbacher, S., P. Hernandez-Acosta, B. Bieseler, M. Blaesse, R. Huber, F. A. Culianez-Macia and T. Kupke (2003). Crystal structure of the plant PPC decarboxylase AtHAL3a complexed with an ene-thiol reaction intermediate. *J Mol Biol*, **327**(1): 193-202.

Strahsburger, E., M. Baeza, O. Monasterio and R. Lagos (2005). Cooperative uptake of microcin E492 by receptors FepA, Fiu, and Cir and inhibition by the siderophore enterochelin and its dimeric and trimeric hydrolysis products. *Antimicrob Agents Chemother*, **49**(7): 3083-6.

Studier, F. W. and B. A. Moffatt (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol*, **189**(1): 113-30.

Sturme, M. H., J. Nakayama, D. Molenaar, Y. Murakami, R. Kunugi, T. Fujii, E. E. Vaughan, M. Kleerebezem and W. M. de Vos (2005). An agr-like two-component regulatory system in Lactobacillus plantarum is involved in production of a novel cyclic peptide and regulation of adherence. *J Bacteriol*, **187**(15): 5224-35.

Stuttard, C. (1982). Temperate Phages of *Streptomyces venezuelae*: Lysogeny and Host Specificity Shown by Phages SV1 and SV2. *J Gen Microbiol*, **128**(1): 115-121.

Subtelny, A. O., M. C. Hartman and J. W. Szostak (2008). Ribosomal synthesis of N-methyl peptides. *J Am Chem Soc*, **130**(19): 6131-6.

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. *Appl Environ Microbiol*, **69**(7): 3777-83.

Tagg, J. R., A. S. Dajani and L. W. Wannamaker (1976). Bacteriocins of gram-positive bacteria. *Bacteriol Rev*, **40**(3): 722-56.

Takano, E., M. Tao, F. Long, M. J. Bibb, L. Wang, W. Li, M. J. Buttner, M. J. Bibb, Z. X. Deng and K. F. Chater (2003). A rare leucine codon in *adpA* is implicated in the morphological defect of *bldA* mutants of *Streptomyces coelicolor*. *Mol Microbiol*, **50**(2): 475-86.

Teng, C. Y., B. Ganem, S. Doktor, B. P. Nichols, R. K. Bhatnagar and L. C. Vining (1985). Total biosynthesis of 4-amino-4-deoxychorismic acid: a key intermediate in the biosynthesis of *p*-aminobenzoic acid and L-*p*-aminophenylalanine. *J Am Chem Soc*, **107**(5008-5009.

Thomas, X., D. Destoumieux-Garzon, J. Peduzzi, C. Afonso, A. Blond, N. Birlirakis, C. Goulard, L. Dubost, R. Thai, J. C. Tabet and S. Rebuffat (2004). Siderophore peptide, a new type of post-translationally modified antibacterial peptide with potent activity. *J Biol Chem*, **279**(27): 28233-42.

Tillotson, R. D., H. A. Wosten, M. Richter and J. M. Willey (1998). A surface active protein involved in aerial hyphae formation in the filamentous fungus *Schizophillum commune* restores the capacity of a bald mutant of the

filamentous bacterium *Streptomyces coelicolor* to erect aerial structures. *Mol Microbiol*, **30**(3): 595-602.

Ueda, K., K. Oinuma, G. Ikeda, K. Hosono, Y. Ohnishi, S. Horinouchi and T. Beppu (2002). AmfS, an extracellular peptidic morphogen in *Streptomyces griseus*. *J Bacteriol*, **184**(5): 1488-92.

Ueda, K., H. Takano, M. Nishimoto, H. Inaba and T. Beppu (2005). Dual transcriptional control of *amfTSBA*, which regulates the onset of cellular differentiation in *Streptomyces griseus*. *J Bacteriol*, **187**(1): 135-42.

Upton, M., J. R. Tagg, P. Wescombe and H. F. Jenkinson (2001). Intra- and interspecies signaling between *Streptococcus salivarius* and *Streptococcus pyogenes* mediated by SalA and SalA1 lantibiotic peptides. *J Bacteriol*, **183**(13): 3931-8.

van Belkum, M. J. and M. E. Stiles (2000). Nonlantibiotic antibacterial peptides from lactic acid bacteria. *Nat Prod Rep*, **17**(4): 323-35.

van der Donk, W. A. (2006). Rings, radicals, and regeneration: the early years of a bioorganic laboratory. *J Org Chem*, **71**(26): 9561-71.

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. *J Bacteriol*, **175**(9): 2578-88.

Vassiliadis, G., D. Destoumieux-Garzon, C. Lombard, S. Rebuffat and J. Peduzzi (2010). Isolation and characterization of two members of the siderophore-microcin family, microcins M and H47. *Antimicrob Agents Chemother*, **54**(1): 288-97.

Vértesy, L., W. Aretz, A. Bonnefoy, E. Ehlers, M. Kurz, A. Markus, M. Schiell, M. Vogel, J. Wink and H. Kogler (1999). Ala(0)-actagardine, a new lantibiotic from cultures of *Actinoplanes liguriae* ATCC 31048. *J Antibiot (Tokyo)*, **52**(8): 730-41.

Walsh, C. T. and E. M. Nolan (2008). Morphing peptide backbones into heterocycles. *Proc Natl Acad Sci U S A*, **105**(15): 5655-6.

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.

Wareham, D. W. and P. Wilson (2002). Chloramphenicol in the 21st century. *Hosp Med*, **63**(3): 157-61.

Weber, W., W. Fischli, E. Hochuli, E. Kupfer and E. K. Weibel (1991). Anantin--a peptide antagonist of the atrial natriuretic factor (ANF). I. Producing organism, fermentation, isolation and biological activity. *J Antibiot (Tokyo)*, **44**(2): 164-71.

Weil, H. P., A. G. Beck-Sickinger, J. Metzger, S. Stevanovic, G. Jung, M. Josten and H. G. Sahl (1990). Biosynthesis of the lantibiotic Pep5. Isolation and characterization of a prepeptide containing dehydroamino acids. *Eur J Biochem*, **194**(1): 217-23.

Welby, M. and J. F. Tocanne (1982). Evidence for the incorporation of a fluorescent anthracene fatty acid into the membrane lipids of *Micrococcus luteus*. *Biochim Biophys Acta*, **689**(1): 173-6.

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. *Proc Natl Acad Sci U S A*, **100**(7): 4316-21.

Wiedemann, I., T. Böttiger, R. R. Bonelli, A. Wiese, S. O. Hagge, T. Gutsmann, U. Seydel, L. Deegan, C. Hill, P. Ross and H. G. Sahl (2006). The mode of action of the lantibiotic lacticin 3147--a complex mechanism involving specific interaction of two peptides and the cell wall precursor lipid II. *Mol Microbiol*, **61**(2): 285-96.

Wieland-Brown, L. C., M. G. Acker, J. Clardy, C. T. Walsh and M. A. Fischbach (2009). Thirteen posttranslational modifications convert a 14-residue peptide into the antibiotic thiocillin. *Proc Natl Acad Sci U S A*, **106**(8): 2549-53.

Willey, J., R. Santamaria, J. Guijarro, M. Geistlich and R. Losick (1991). Extracellular complementation of a developmental mutation implicates a small sporulation protein in aerial mycelium formation by *S. coelicolor*. *Cell*, **65**(4): 641-50.

Willey, J. M. and W. A. van der Donk (2007). Lantibiotics: peptides of diverse structure and function. *Annu Rev Microbiol*, **61**(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. *Appl Environ Microbiol*, **75**(5): 1381-7.

Wilson, K. A., M. Kalkum, J. Ottesen, J. Yuzenkova, B. T. Chait, R. Landick, T. Muir, K. Severinov and S. A. Darst (2003). Structure of microcin J25, a peptide inhibitor of bacterial RNA polymerase, is a lassoed tail. *J Am Chem Soc*, **125**(41): 12475-83.

Wirawan, R. E., K. M. Swanson, T. Kleffmann, R. W. Jack and J. R. Tagg (2007). Uberolysin: a novel cyclic bacteriocin produced by *Streptococcus uberis*. *Microbiology*, **153**(Pt 5): 1619-30.

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): 103-8.

Xie, L., C. Chatterjee, R. Balsara, N. M. Okeley and W. A. van der Donk (2002). Heterologous expression and purification of SpaB involved in subtilin biosynthesis. *Biochem Biophys Res Commun*, **295**(4): 952-7.

Xie, 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): 679-81.

Xue, Y., L. Zhao, H. W. Liu and D. H. Sherman (1998). A gene cluster for macrolide antibiotic biosynthesis in *Streptomyces venezuelae*: architecture of metabolic diversity. *Proc Natl Acad Sci U S A*, **95**(21): 12111-6.

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*. *Mol Microbiol*, **53**(2): 555-72.

Yang, K., L. Han and L. C. Vining (1995). Regulation of jadomycin B production in *Streptomyces venezuelae* ISP5230: involvement of a repressor gene, *jadR2*. *J Bacteriol*, **177**(21): 6111-7.

Yoon, G. S., K. H. Ko, H. W. Kang, J. W. Suh, Y. S. Kim and Y. W. Ryu (2006). Characterisation of *S*-adenosylmethionine synthetase from *Streptomyces avermitilis* NRRL8165 and its effect on antibiotic production. *Enzyme Microb Technol*, **39**(466-473.

Yorgey, P., J. Lee, J. Kordel, E. Vivas, P. Warner, D. Jebaratnam and R. Kolter (1994). Posttranslational modifications in microcin B17 define an additional class of DNA gyrase inhibitor. *Proc Natl Acad Sci U S A*, **91**(10): 4519-23.

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. *J Bacteriol*, **192**(3): 841-60.

Zerikly, M. and G. L. Challis (2009). Strategies for the discovery of new natural products by genome mining. *Chembiochem*, **10**(4): 625-33.

Zhang, L., L. Gray, R. P. Novick and G. Ji (2002). Transmembrane topology of AgrB, the protein involved in the post-translational modification of AgrD in *Staphylococcus aureus*. *J Biol Chem*, **277**(38): 34736-42.

Zhang, L., J. Lin and G. Ji (2004). Membrane anchoring of the AgrD N-terminal amphipathic region is required for its processing to produce a quorum-sensing pheromone in *Staphylococcus aureus*. *J Biol Chem*, **279**(19): 19448-56.

Zheng, G., R. Hehn and P. Zuber (2000). Mutational analysis of the *sbo-alb* locus of *Bacillus subtilis*: identification of genes required for subtilosin production and immunity. *J Bacteriol*, **182**(11): 3266-73.

Zheng, G., L. Z. Yan, J. C. Vederas and P. Zuber (1999). Genes of the *sbo-alb* locus of *Bacillus subtilis* are required for production of the antilisterial bacteriocin subtilosin. *J Bacteriol*, **181**(23): 7346-55.

Zheng, J. T., S. L. Wang and K. Q. Yang (2007). Engineering a regulatory region of jadomycin gene cluster to improve jadomycin B production in *Streptomyces venezuelae*. *Appl Microbiol Biotechnol*, **76**(4): 883-8.

Ziemert, N., K. Ishida, A. Liaimer, C. Hertweck and E. Dittmann (2008). Ribosomal synthesis of tricyclic depsipeptides in bloom-forming cyanobacteria. *Angew Chem Int Ed Engl*, **47**(40): 7756-9.

Ziemert, N., K. Ishida, P. Quillardet, C. Bouchier, C. Hertweck, N. T. de Marsac and E. Dittmann (2008). Microcyclamide biosynthesis in two strains of *Microcystis aeruginosa*: from structure to genes and *vice versa*. *Appl Environ Microbiol*, **74**(6): 1791-7.

Ziemert, N., K. Ishida, A. Weiz, C. Hertweck and E. Dittmann (2010). Exploiting the natural diversity of microviridin gene clusters for discovery of novel tricyclic depsipeptides. *Appl Environ Microbiol*, **76**(11): 3568-74.

Zimmermann, N. and G. Jung (1997). The three-dimensional solution structure of the lantibiotic murein-biosynthesis-inhibitor actagardine determined by NMR. *Eur J Biochem*, **246**(3): 809-19.

Zouhir, A., R. Hammami, I. Fliss and J. B. Hamida (2010). A new structure-based classification of gram-positive bacteriocins. *Protein J*, **29**(6): 432-9.