Analysis and manipulation of “actagardine” gene clusters from *Actinoplanes*

Robert Andrew Bell

December 2010

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

Acknowledgments

First, I would like to thank my supervisor, Prof. Mervyn Bibb. His continued help and support over the last four years have been invaluable to me. I would also like to thank him for allowing me the opportunity to work on this project and allowing me to work at the John Innes Centre and attend and present my work at a number of national and international venues.

I would also like to thank the many people that I have met and worked with whilst studying at the John Innes Centre. Particular thanks go to members of my supervisory team, Dr Stephen Bornemann and Dr Brain Rudd. I would also like to thank Dr Sean O'Rourke and Nick Bird, for their help getting me settled into my PhD, and their initial involvement in this work. I would also like to thank all members of the Molecular Microbiology Department, especially Prof. Mark Buttner, Dr Govind Chandra and Maureen Bibb for their continued assistance throughout this project.

I would also like to mention a number of students and colleagues that have shared this experience with me and have offered much support over the last four years. These include Jan Claesen, Lucy Foulston, Emma Sherwood, Juan-Pablo Gomez Escribano, Matt Bush, Richard Little, Peter Slavny, Marij Claesen, Antje Hempel and Jane Moore.

I would like to thank the members of Novacta Biosystems that helped guide this project. These include Dr Brain Rudd, Dr Steven Boakes, Dr Jesus Cortes, Dr Anthony Appleyard and Dr Mike Dawson.

Lastly, I would like to thank all my family for their unwavering support though my PhD. I add special thanks to my wife, who has shown me compassion and understanding over the last four years.

This thesis is dedicated to the loving memory of

Beryl Anne Batey

1950 - 2010

Abstract

Lantibiotics are post-translationally modified antimicrobial peptides and are synthesised by a number of Gram-positive bacteria. Lantibiotics are defined by the inclusion of lanthionine residues and often contain other modified amino acids within their structures. Actagardine, produced by *Actinoplanes garbadinensis,* was originally discovered in 1976 and identified as a potential clinically useful antibiotic. In recent years, discovery of an actagardine homolog, deoxyactagardine B, produced by *Actinoplanes liguriae* has prompted further investigation into the biosynthesis of these lantibiotics.

The biosynthetic gene clusters for actagardine and deoxyactagardine B biosynthesis were identified by Novacta Biosystems from cosmid libraries of the natural producers. Heterologous expression of these cosmids in various *Streptomyces* hosts has confirmed that they contained all the genes required for the biosynthesis of these lantibiotics. An in-depth computational analysis of the cosmid has allowed the identification of a number of lantibiotic-associated biosynthesis genes, including the unique lantibiotic monooxygenase GarO, responsible for the oxygen moiety on the C-terminal methylanthionine bridge of actagardine. Using this information, a directed reduction in the cosmid size identified a likely set of genes that are essential for the biosynthesis of actagardine.

With these data, a deletion series of all the putative lantibiotic biosynthesis genes within the actagardine gene cluster demonstrated that a number of these genes were essential for its biosynthesis. It also demonstrated that the *Streptomyces* hosts are apparently able to compensate for the loss of a number of genes.

Computational analysis identified the actagardine homologs michiganin A and gnavucin, an undiscovered lantibiotic prepropeptide. Whilst the structural peptides of these four lantibiotics are very similar; they are encoded by very different putative biosynthetic gene clusters, requiring further investigations for an understanding of their biosynthesis.

Abbreviations

A	Adenine
ATP	Adenosine triphosphate
AviCys	S-[(Z)-2-aminovinyl]-D-cysteine
Apra	Apramycin
BLAST	Basic local alignment search tool
C	Cytosine
Carb	Carbenicillin
Chlor	Chloramphenicol
C-terminal	Carboxy-terminal
Da	Dalton
DAA	Deoxyactagardine A
DAB	Deoxyactagardine B
Dha	2,3-didehydroalanine
Dhb	(Z)-2,3-didehydrobutyrine
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
DNA	Difco Nutrient Agar
DNAse	Deoxyribonuclease
dNTP	Deoxynucleoside triphosphate
G	Guanine
gDNA	Genomic DNA
HPLC	High Pressure Liquid Chromatography

Hyg	Hygromycin
Kan	Kanamycin
kDa	Kilodalton
lacZ	β-galactosidase
Lan	Lanthionine
LC-MS	Liquid Chromatography-Mass Spectrometry
MALDI-ToF	Matrix-Assisted Laser Desorption Ionisation Time-of-Flight Mass Spectroscopy
MeLan	Methyl-lanthionine
MGS	Minimal Gene Set
MIC	Minimum Inhibitory Concentation
Nal	Nalidixic acid
NMR	Nuclear Magnetic Resonance
N-terminal	Amino-terminal
OD	Optical Density
PCR	Polymerase Chain Reaction
qRT-PCR	Quantitative Reverse Transcriptase Polymerase Chain Reaction
Spec	Spectinomycin
Strep	Streptomycin
Tet	Tetracycline
Thi	Thiamine
UV	Ultra Violet

Contents

Index of figures

Index of tables

1 Introduction

The discovery of penicillin in 1928 by Alexander Fleming sparked a continuing battle between human infections and man's ability to fight them. The identification, industrial scale production and wide-spread use of a variety of antibiotics have inevitably led to the evolution of drug-resistant pathogens. Unfortunately, there is now no antibiotic in clinical use to which resistance has not developed (Weigel *et al.*, 2003), and while frequently hailed as the "drug of last resort", vancomycin is now ineffective against many enterococcal infections that are also resistant to other clinically used antibiotics. Consequently there is a pressing need for new antimicrobial agents to fight infectious diseases. Lantibiotics have the potential to reinforce our ability to combat bacterial infections.

Lantibiotics are post-translationally modified, ribosomally synthesised antimicrobial peptides produced by Gram-positive bacteria and are effective against a wide variety of bacteria. The term lantibiotic was first introduced in 1988 to describe these lanthionine (and/or methyl-lanthionine)-containing antibiotics (Schnell *et al.*, 1988). The amino acid lanthionine, created from a thioether cross-link between the *β*-carbons of two alanine residues, was first isolated from sodium carbonate treated wool in 1941 (Horn *et al.*, 1941). The lanthionine and methyl-lanthionine bridges found in lantibiotics arise from 2,3-didehydroalanine (Dha) and (Z)-2,3-didehydrobutyrine (Dhb), which are formed by the enzymatic dehydration of serine and threonine residues, respectively. The dehydrated amino acids are then subjected to nucleophilic attack by SH (thiol) groups of cysteine residues located elsewhere in the peptide to form the characteristic lanthionine and methyl-lanthionine ring structures (see figure 1.1) (Sahl & Bierbaum, 1998). As well as these residues, lantibiotics may also contain other unusual modified amino acids and modifications such as lysinoalanine, S-aminovinyl-D-cysteine, ß-hydroxy-aspartate, D-alanine, 2-oxobutyrate, 2-oxopropionate (pyruvate), 2-hydroxypropionate (lactate), and S-aminovinyl D-methylcysteine (Chatterjee *et al.*, 2005). Some of the more common post-translational modifications and representative lantibiotic structures found in lantibiotics are shown in figure 1.1 and 1.2.

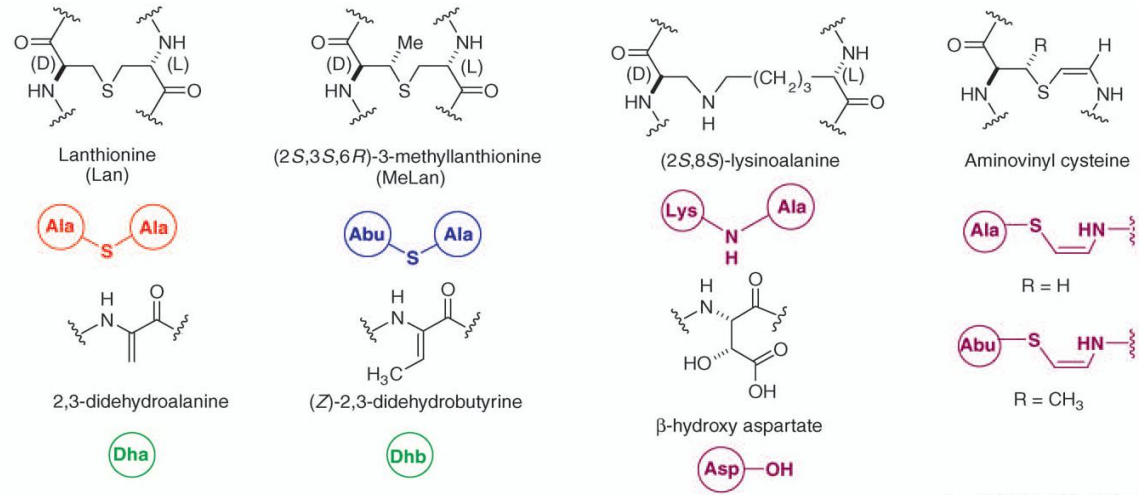


Figure 1.1: Common lantibiotic structural motifs produced from post-translational modification of amino acids via lantibiotic biosynthetic enzymes. The coloured pictograms are used henceforth to represent the individual modifications in lantibiotic structures. Structures taken from Patton and van der Donk (2005).

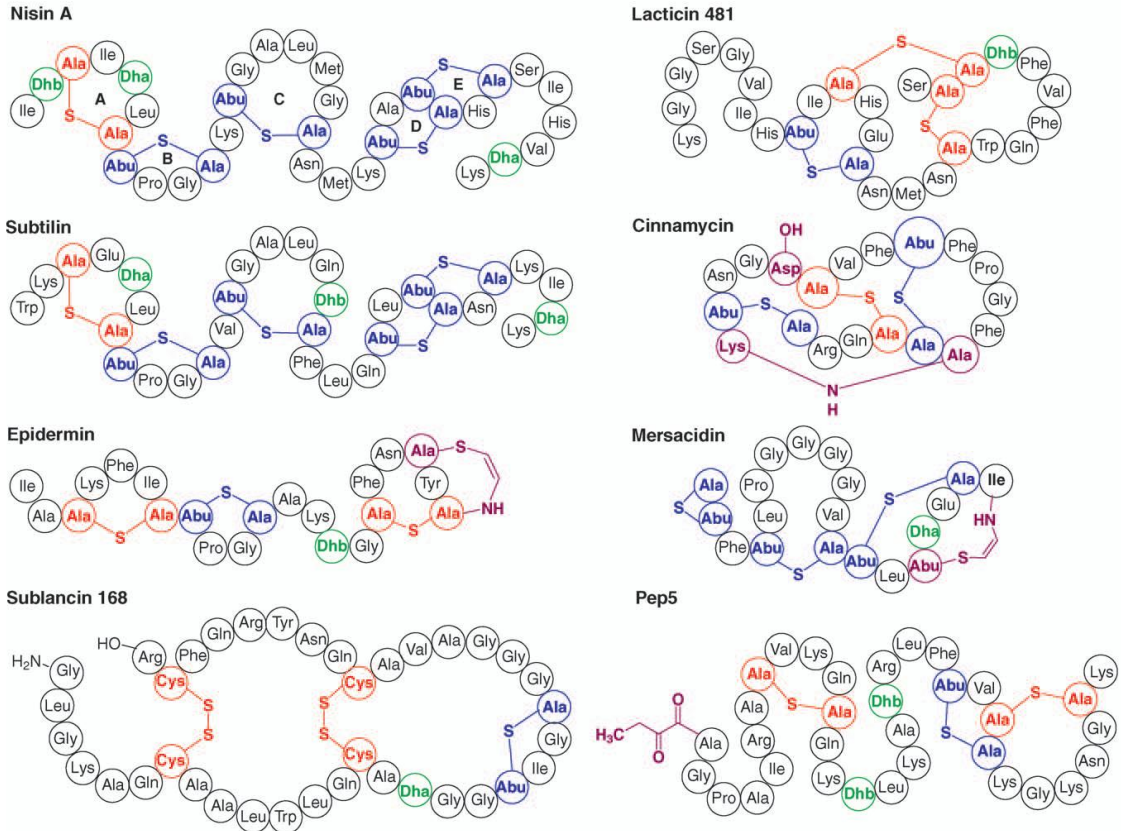


Figure 1.2: Representative structures of various lantibiotics. The common structural motifs and other post-translational modifications are shown in more detail in figure 1.1 and are represented in the same colours in the structures above. Many of these structures are discussed in the subsequent text. Figure taken from Patton and van der Donk (2005).

The unusual modifications found in lantibiotics result from the unique modification enzymes encoded by their respective biosynthetic gene clusters. Lantibiotic biosynthetic gene clusters can be found on conjugative transposable elements (nisin), on the chromosome of the host (subtilin) and on plasmids (epidermin and lacticin 481) (Chatterjee *et al.*, 2005). Genes within many lantibiotic clusters are arranged in operons, consisting of several transcriptional units, often with a weak terminator structure, usually a stem-loop, in the intergenic sequence between the structural gene that encodes the unmodified peptide and downstream genes encoding the modification enzymes (McAuliffe *et al.*, 2001). This can be seen in figure 1.3 that shows a number of representative biosynthetic gene clusters.

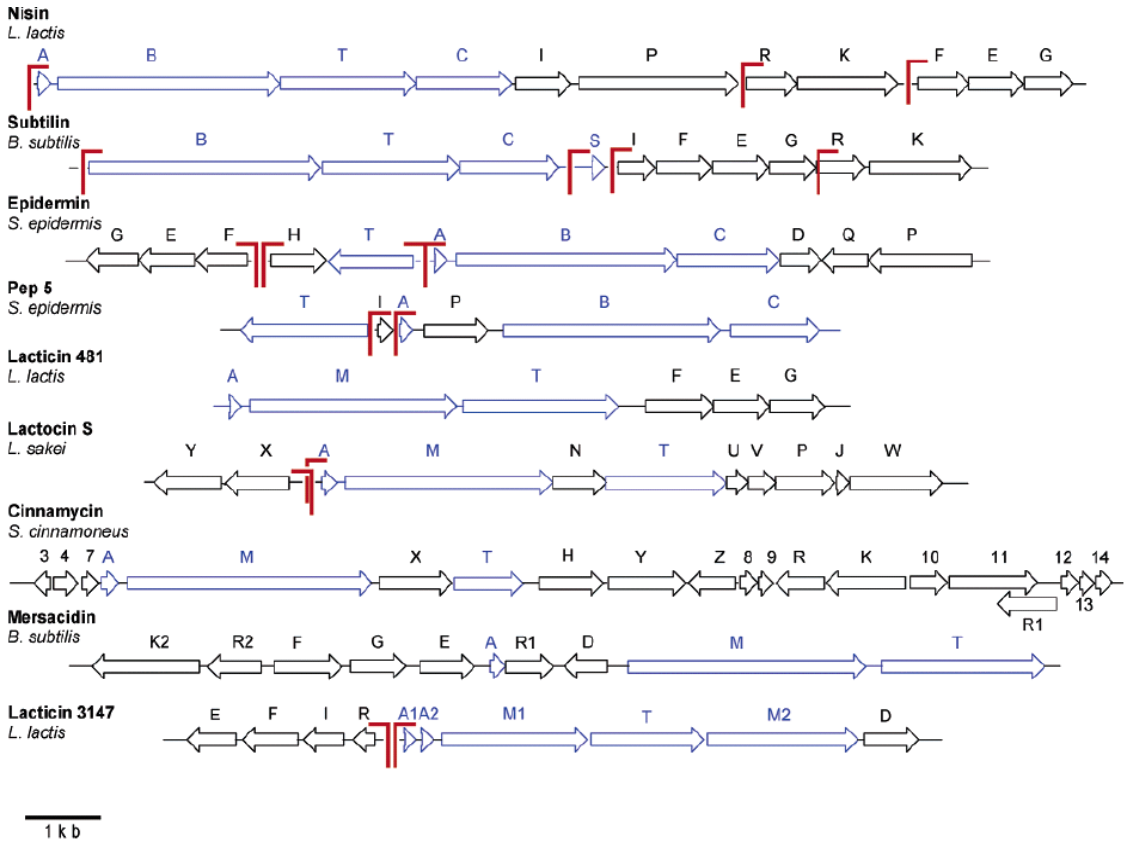


Figure 1.3: The biosynthetic gene clusters of nisin, subtilin, epidermin, Pep5, lacticin 481, lacticin S, cinnamycin, mersacidin and lacticin 3147. In blue are genes that encode functions that are common to all lantibiotic gene clusters. The red bars indicate promoters for known transcriptional units. Individual genes are discussed in the text below. Figure taken from Chatterjee, Paul *et al.* (2005).

Lantibiotic biosynthetic genes have been given the generic locus symbol *lan*, which is replaced by the specific name for each lantibiotic (e.g. *nis* for nisin) (Patton & van der Donk, 2005). The *lanA* gene encodes the lantibiotic precursor prepropeptide, LanA. The LanA prepropeptide contains a C-terminal region, the propeptide, that contains the unmodified structural region which will undergo post translational modification by either the LanBC or LanM modification machinery and become the mature lantibiotic. The LanA prepropeptide also contains the relatively long (23 to 59 amino acids) N-terminal leader sequence ("pre") which does not undergo modification and which is eventually cleaved away from the mature lantibiotic. The leader sequences is thought to be responsible for substrate recognition by the various biosynthetic enzymes and may be involved in sequestration of the lantibiotic before it is exported from the cell (Patton & van der Donk, 2005, Chatterjee *et al.*, 2005). It is interesting to note that while both serine and threonine residues are found in propeptide and leader sequences, cysteine residues are found only in propeptide sequences.

Lantibiotics were classified as group I bacteriocins by Klaenhammer (1993) but the activity of many lantibiotics is much wider than that usually associated with bacteriocins, which generally inhibit only closely related bacteria (Klaenhammer, 1993). Lantibiotics can be classified based on the structure of the mature lantibiotic, and/or by the genes that are present within their respective cluster. Type A lantibiotics are characteristically described as elongated, cationic peptides that can also be amphipathic and screw shaped in solution. Type A lantibiotics are usually between 20 and 34 amino acids in length (Chatterjee *et al.*, 2005, Patton & van der Donk, 2005). Type AI lantibiotics have the modification enzymes LanB and LanC (see later for a detailed description) as two distinct enzymes, whereas the Type AII and B lantibiotics contain a single bifunctional LanM type modification enzyme. Type B lantibiotics differ from the type A class by being globular in shape, and generally have no net charge, or a negative charge at pH 7, and have a single LanM modification enzyme (Chatterjee *et al.*, 2005). Actagardine, the subject of this thesis, belongs to this group of lantibiotics. The most studied lantibiotic to date is the Type AI lantibiotic nisin, produced by *Lactococcus lactis*. Nisin was first discovered in 1928 and has been used extensively in the food industry as a preservative for nearly 40 years, with little development of resistance in susceptible organisms (Delves-Broughton *et al.*, 1996, Rogers, 1928, Rogers & Whittier, 1928).

1.1 Lantibiotic Biosynthesis

The biosynthesis of a particular lantibiotic is determined by the genes present in its corresponding gene cluster. The gene clusters encode all of the necessary components to produce a fully modified, functional peptide, and the ability to resist the effects of the lantibiotic after it has been exported. The gene clusters must include genes for the lantibiotic precursor, for the dehydration and cross-linking reactions, and for secretion of the modified peptide (figure 1.4) (McAuliffe *et al.*, 2001).

1.1.1 Leader Sequence

The leader sequence present at the N-terminus of the prepropeptide has many proposed functions, including signalling for transport, providing host immunity and allowing the recognition and correct modification of the pre-lantibiotic by the post-translational modification machinery, acting as both the signal and scaffold (Chatterjee *et al.*, 2005). Unlike the N-terminal leader sequences of many secreted proteins, those for lantibiotics (with the possible exception of cinnamycin) do not contain the usual characteristics of a *sec*-dependent transport signal sequence and are not likely to be exported by general transporter complexes (Chatterjee *et al.*, 2005). However lantibiotic gene clusters contain *lanT* and sometimes *lanH* genes, encoding LanT and LanH transporter proteins, respectively, that are responsible for export of the peptide. The leader sequence is also thought to be important in sequestration of the lantibiotic, protecting the producing organism from the effects of the mature lantibiotic before it has been exported from the cell. Studies with nisin (van der Meer *et al.*, 1994, Qiao *et al.*, 1995), subtilin (Corvey *et al.*, 2003, Stein *et al.*, 2002), lacticin 481 (Xie *et al.*, 2004) and mutacin II (Chen *et al.*, 2001) have shown that the presence of the leader sequence attached to the fully modified lantibiotic either abolished or severely reduced biological activity. NMR studies demonstrated that the N-terminal leader sequence interfered with the interaction of mature nisin with the cell membrane (van den Hooven *et al.*, 1997). The leader sequence is also crucial for recognition of the prepropeptide by the modification machinery of the lantibiotic cluster (Chakicherla & Hansen, 1995).

Leader sequences often contain conserved motifs that are found only in certain lantibiotic classes, and this characteristic has also been used as a means of classification (see figure 1.4). The class I leader sequences recognised by the LanB and LanC modification enzymes responsible for the dehydration and cyclisation reactions, respectively, contain a high proportion of highly charged amino acids, some of which are highly conserved, giving the leader peptide either a net negative or slightly positive charge. The most notable conserved region within this group is the FNLDV box which is located between positions -15 and -20 (in respect to the cleavage site); there is also a conserved proline residue located at the -2 position (in red in figure 1.4) (Chatterjee *et al.*, 2005, McAuliffe *et al.*, 2001). Class II lantibiotics that are modified by a single LanM enzyme have leader sequences that contain the conserved motifs EL/EV and EL/EM (in red in figure 1.4) (Patton & van der Donk, 2005). The class II lantibiotic leader sequences also have a GG or GA cleavage site (known as the double Gly motif), and typically contain an unusually high concentration of Asp and Glu residues that results in high negative net charge (Chatterjee *et al.*, 2005, McAuliffe *et al.*, 2001). The existence of such conserved motifs in the leader sequences of both classes of lantibiotic suggests that they are important for leader sequence function, and many studies have focused on these conserved motifs in an effort to discover their role in lantibiotic biosynthesis (Patton *et al.*, 2008, Neis *et al.*, 1997, van der Meer *et al.*, 1994).

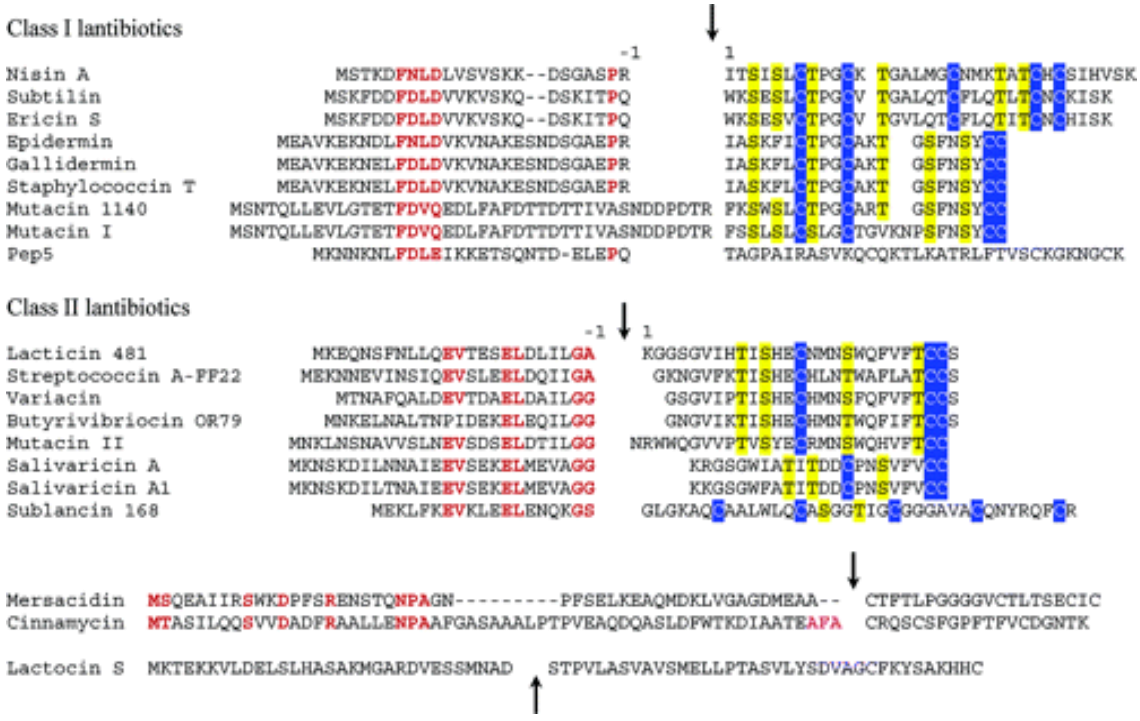


Figure 1.4: Sequence alignments of a number of leader sequences and lantibiotics. The conserved residues discussed in the text are shown in red. Arrows indicate

cleavage sites. The residues in yellow and blue denote those involved in ring formation. Figure taken from Chatterjee, Paul *et al.* (2005).

Several lantibiotics have been shown to tolerate mutations in nearly all of the residues outside of the FN/DLDV box in the class I group leader sequences. Several amino acid substitutions were made in the leader sequence of nisin, including the conserved FN/DLD and P residues. Fully modified and functional nisin resulted from nearly all of the amino acid substitutions that were made outside of the FNLD motif and the -2 position, indicating that these residues are not essential for lantibiotic biosynthesis. In contrast, nisin biosynthesis was abolished when substitutions were made in the FNLD motif, suggesting that these conserved residues may be required for appropriate interaction with the respective modification enzymes and therefore essential for biosynthesis (van der Meer *et al.*, 1994). Chatterjee *et al.* (2005) suggested that that the leader sequences are recognised by the modification machinery as a consequence of the tertiary structure of the leader sequence and propeptide, rather than by recognition of specific amino acids in the leader sequence.

Class II leader sequences are also amenable to numerous amino acid exchanges and appear to be just as intolerant to substitutions within the conserved motifs, particularly at the cleavage site (see figure 1.4). Experiments carried out on mutacin demonstrated that the conversion of either one or both of the double Gly motif residues to Ala at the -1 and -2 positions completely abolished lantibiotic production (Chen *et al.*, 2001).

Strain	Biological activity	Mutacin Prepropeptide	Mutacin peptide
Wild-type	++++	−	+
G-1A	−	+	−
G-2A	−	+	−
I-4V	++++	np	np
I-4D	++++	−	−
L-7M	++++	np	np
L-7K	−	−	−
E-8D	++++	np	np
E-8K	+++	−	+
S-11A	++++	np	np
S-11T	++++	np	np
V-12A	++++	np	np
V-12I	++++	np	np
V-12L	++	−	+
E-13D	++++	np	np
E-13K	+	−	−

[a]The mutacin production level of the wild-type *Streptococcus mutans* T8 was assigned as ++++.

[b]+, signal corresponding to correct mass was identified ; −, no signal corresponding to correct mass was identified;

np, not performed

Figure 1.5: Effects of amino acid substitutions within the leader sequence of mutacin (Chen *et al.*, 2001).

Figure 1.5 also demonstrates that the mutacin leader sequence can tolerate a large number of other amino acid substitutions. Only two substitutions abolished the interaction of the mutacin prepropeptide with its modification enzymes: the conversion of Ile to Asp and Leu to Lys at -4 and -7, respectively. Glu8 to Lys, Val12 to Leu and Glu13 to Lys reduced the level of mutacin production to approximately 75%, 50% and 10% of the wild-type level, respectively. This demonstrates that while not all of the amino acids in the leader sequence are essential for biosynthesis of the lantibiotic, some are important for optimal biosynthesis (Chen *et al.*, 2001).

Further studies on the lantibiotic lacticin 481 have shown that the length of the prepropeptide is not crucial for correct modification (Xie *et al.*, 2004). C-terminally

truncated prepropeptide was modified and complete removal of the leader sequence from lacticin 481 did not abolish dehydration by LctM (Levengood *et al.*, 2007). Prepropeptides that lacked the first three, eight or eleven amino acids from the leader sequence were still modified correctly; though removal of the first 14 amino acids abolished modification (Chatterjee and van der Donk unpublished data) (Chatterjee *et al.*, 2005). In lacticin 481, residues 17 to 24 (in red in figure 1.4) are suggested to be essential for enzyme recognition, while the GA sequence is essential for correct cleavage of leader peptide to yield the mature lantibiotic and are not required for recognition by the modification machinery (Chatterjee *et al.*, 2005).

1.1.2 Enzymes and proteins involved in lantibiotic biosynthesis

As mentioned previously, the *lanA* gene encodes the unmodified prepropeptide that is modified by either a LanM or LanBC modification enzymes. Lantibiotics are synthesised by the dehydration and cyclisation of serine and threonine residues present within the sequence of the propeptide. Lantibiotic gene clusters contain either a single modification gene, *lanM*, or two separate genes *lanBC*. The first step in lantibiotic biosynthesis is the dehydration of serine and threonine residues, followed by cyclisation of the dehydrated peptides with cysteine thiols. In type AI lantibiotics, dehydration is carried out by the LanB dehydratase, followed by ring formation mediated by the cyclase LanC. In type AII and B clusters, both reactions are performed by a single LanM enzyme.

Mutations in the nisin propeptide have given some insight into the organisation and order of lantibiotic biosynthesis. Lubelshi *et al.* (2009) hypothesise that the nisin prepropeptide is processed from the N-terminus to the C-terminus, and that NisB and NisC function alternatively to dehydrate and cyclise the lantibiotic.

1.1.3 LanB

lanB genes encode dehydratase enzymes that are responsible for the selective dehydration of serine and threonine residues to Dha and Dhb, respectively

(Chatterjee *et al.*, 2005, McAuliffe *et al.*, 2001, Patton & van der Donk, 2005). *lanB* genes generally encode large (about 1000 amino acids) generally hydrophilic molecules, with some hydrophobic domains, suggesting that they are membrane associated, as proposed for the nisin biosynthesis complex (see regulation of nisin biosynthesis, figure 1.10) (McAuliffe *et al.*, 2001, Siegers *et al.*, 1996). There is little homology between the LanB proteins of different Class AI lantibiotic gene clusters (generally around 30% amino acid sequence identity), although when comparing structurally similar lantibiotics, homology can be much higher (83% in the case of SpaB and EriB involved in subtilin and ericin S biosynthesis, respectively (Mulders *et al.*, 1991)). LanB activity was first confirmed when dehydrated Pep5 was identified after deletion of *pepC* in a Pep5-producing strain of *Staphylococcus epidermidis* (Meyer *et al.*, 1995). The dehydratase activity of LanB was also confirmed for nisin when *nisABT* was placed in a non-producing strain of *L. lactis*, resulting in the production of dehydrated prenisin lacking thioether rings (Kuipers *et al.*, 2004). His-tagged nisin precursor peptides were expressed in a strain that lacked *nisB*, resulting in accumulation of unmodified prenisin. In contrast removal of *nisC* from *L. lactis* resulted in the production of dehydrated peptides, providing further confirmation that the LanB is responsible for the dehydration of the prepropeptide (Okeley *et al.*, 2003, Patton & van der Donk, 2005). The LanB dehydratases perform the dehydration reaction in the same way as the LanM synthetases (discussed later).

1.1.4 LanC

The next step in modifying the lantibiotic after dehydration in the type AI lantibiotic biosynthetic pathway is cyclisation of the dehydrated prepropeptide to form the uncleaved fully processed lantibiotic. This is performed by the LanC cyclase. LanCs catalyse the addition of cysteine thiols to the dehydrated serine and threonine residues (figure 1.6) (Relyea & van der Donk, 2006). The amount of energy needed to form rings within the lantibiotic structure is thought to be less than that required to dehydrate the prepropeptide. Biomimetic studies of lantibiotic biosynthesis have demonstrated that dehydrated peptides are able to spontaneously undergo cyclisation reactions in the absence of a LanC cyclase. Not only were these reactions rapid, but the resulting peptides contained thioether rings with the correct stereochemistry; questioning the need for a dedicated cyclase (Burrage *et al.*, 2000, Chatterjee *et al.*, 2005). In contrast, Meyer *et al.* (1995)

proposed that the PepC cyclase was essential for the correct and efficient cyclisation of dehydrated Pep5. Disruption of *pepC* led to the production of dehydrated peptides with none of the lanthionine bridges formed correctly (Meyer *et al.*, 1995). Further experiments on nisin have provided additional direct evidence for the important role played by LanCs in lantibiotic biosynthesis. His-tagged nisin precursor peptides were expressed in mutant strains of *L. lactis* lacking *nisB* or *nisC* (Koponen *et al.*, 2002). In the *nisB* deficient strain, no dehydrated or cyclised peptides were recovered. In the *nisC* deficient strain, only dehydrated peptides were isolated, demonstrating that NisC was responsible for the cyclisation of thioether rings, or indirectly by inducing cyclisation activity in NisB (Koponen *et al.*, 2002, Chatterjee *et al.*, 2005). For both NisC and SpaC, heterologous expression and purification in *E. coli* demonstrated that these two enzymes were monomers and metal analysis showed that they contained zinc (Okeley *et al.*, 2003, Patton & van der Donk, 2005). It is thought that the zinc is involved in activation of the thiol substrates in LanC enzymes, lowering the pK_a of the cysteine thiols that are added to the dehydrated Dha and Dhb residues (Chatterjee *et al.*, 2005, Relyea & van der Donk, 2006), similar to a number of enzymes that catalyse thiol alkylation, figure 1.6 (Matthews & Goulding, 1997, Chatterjee *et al.*, 2005). Generally, LanC cyclases share low amino acid identity (~20-30%) and only a few conserved residues. They contain two conserved cysteine residues (Cys284 and Cys330, NisC numbering) and two histidine residues (His212 and His331) that are also conserved in the C-terminal part of the LanM proteins and are thought to act as zinc ligands (Chatterjee *et al.*, 2005, Siezen *et al.*, 1996, Okeley *et al.*, 2003). Conversion of the two conserved cysteines to alanine in SpaC limited its ability to bind zinc and perform its cyclisation activity (Okeley *et al.*, 2003). Thus, while these enzymes are not essential for the cyclisation of all lantibiotics, they are responsible for the efficient and generally correct stereo- and regio-specific formation of lanthionine rings (Chatterjee *et al.*, 2005).

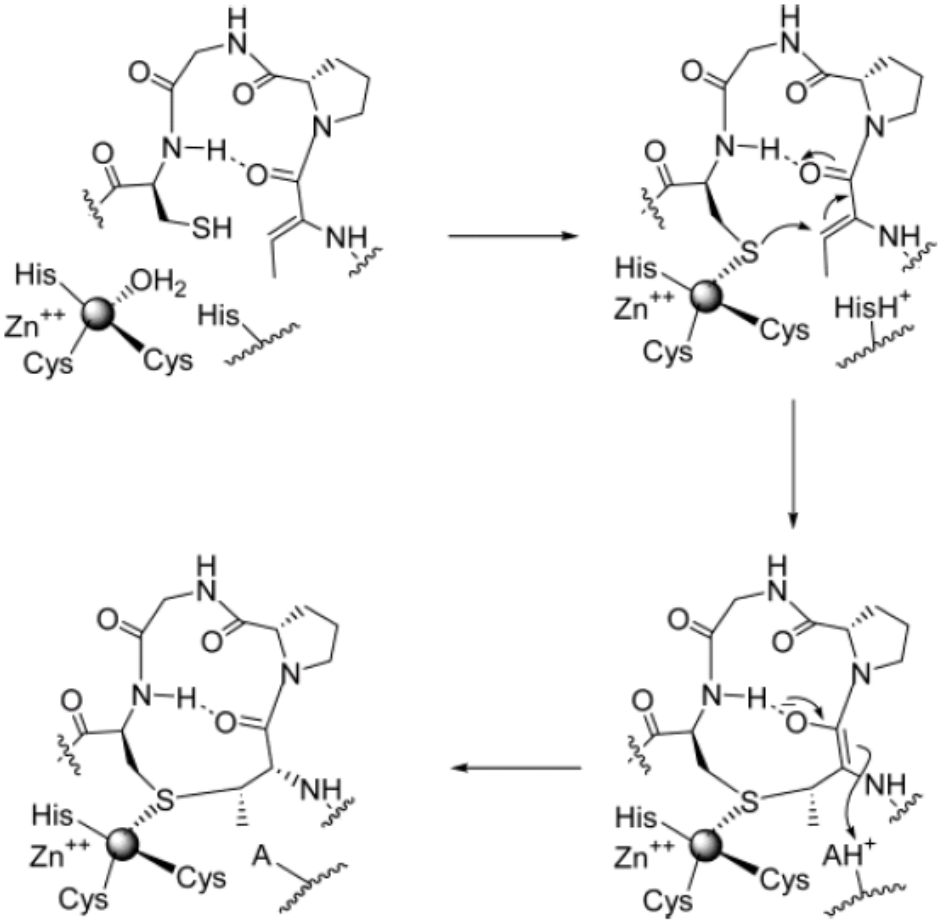


Figure 1.6: Cyclisation by LanC enzymes via the addition of cysteine thiols to the dehydrated peptides, in this case in the formation of the B-ring of subtilin and nisin. Figure taken from Okelely *et al*. (2003).

1.1.5 LanM

In contrast to Type AI lantibiotics, Type II and B lantibiotics are processed by large (generally 900-1000 amino acids) bifunctional modification enzymes encoded by *lanM* genes. These enzymes are able to carry out both the dehydration and cyclisation reactions (Chatterjee *et al.*, 2005, Xie *et al.*, 2004). The mechanism of dehydration is believed to differ from that of LanBs involved in the synthesis of Type IA lantibiotics, and involves the addition of a phosphate group to the alcohol functionalities of Ser and Thr followed by the elimination of these phosphates to yield Dha and Dhb (figure 1.7) (Relyea & van der Donk, 2006). *In vitro* studies on lacticin 481 demonstrated that reconstituted LctM required both Mg^{2+} and ATP to fully process the lantibiotic (Xie *et al.*, 2004, Patton & van der Donk, 2005).

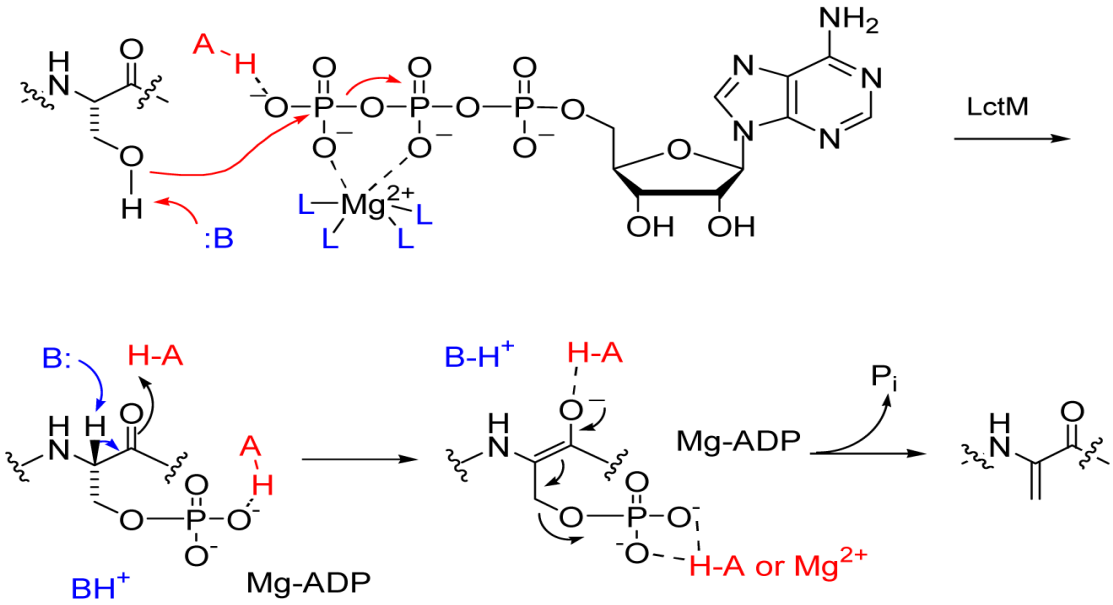


Figure 1.7: Schematic of LanM dehydration via the addition of a phosphate group to the alcohol functionalities of Ser. Figure taken from You and van der Donk (2007).

Amino acid sequence identity between the C-terminal region of LanMs and LanCs is around 20-27%, and there is no apparent homology between LanM and LanB proteins. This suggests that *lanM* genes have not arisen by the fusion of *lanB* and *lanC* genes, and that the two modification machineries have evolved independently of each other (Chatterjee *et al.*, 2005). The homology shared with LanCs includes the conserved residues that are involved in binding zinc.

Recent studies have shown that disruption of the zinc ligands Cys781 and Cys836 in the lacticin synthetase LctM did not affect its dehydration activity. The mutations did however prevent cyclisation of full-length and truncated substrates by LctM. Substitution of another conserved *lanC* and *lanM* residue, His725, with Asn did not abolish the ability of LctM to catalyse dehydration and cyclisation reactions, but it decreased the level of activity compared to the wild-type enzyme (figure 1.8) (You & van der Donk, 2007). You *et al.* (2007) proposed that the C-terminal domain of the peptide, containing the zinc ligands, is responsible for cyclisation and that dehydration takes place independently.

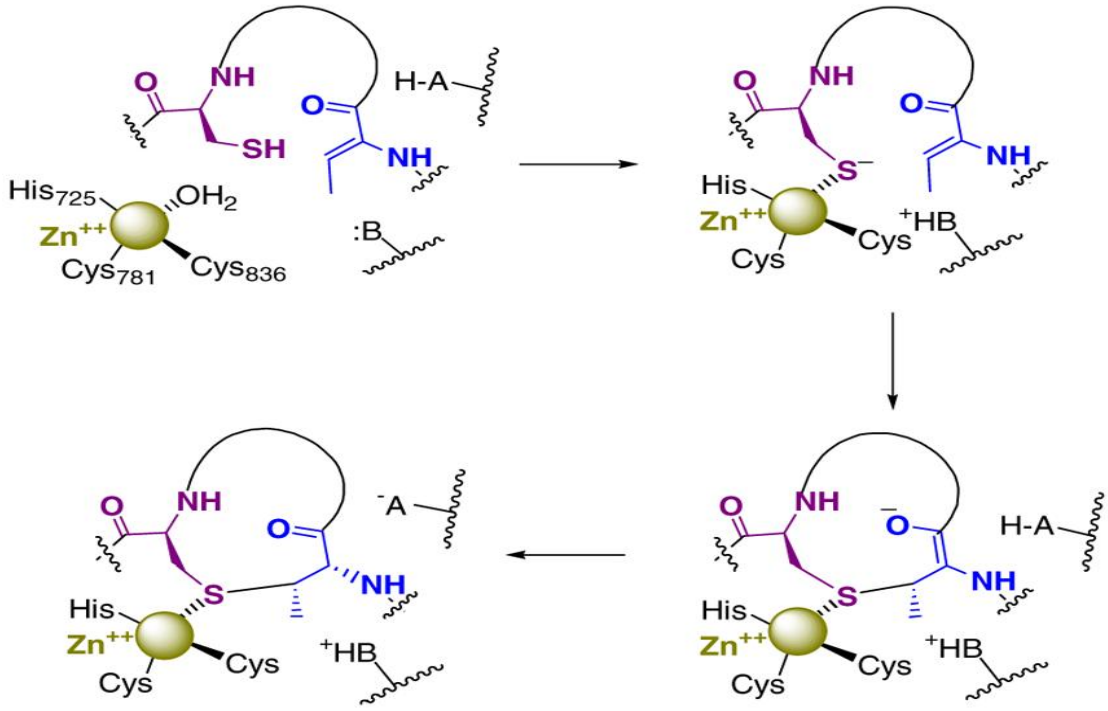


Figure 1.8: The proposed mechanism for the LctM-catalysed formation of a methylanthionine ring involving a number of conserved residues (discussed in the above text). Figure taken from You and van der Donk (2007).

1.1.6 LanD

There are a number of lantibiotic genes found in a number of clusters that encode enzymes able to perform other modifications on lantibiotic prepropeptides. Figure 1.1 illustrates many of the modifications performed by these enzymes, including the production of aminovinyl cysteine, lysinoalanine, D-alanine and ß-hydroxy aspartate (Patton and van der Donk, 2005). Enzymes encoded by *lanD* genes are able to modify the lantibiotic peptides and create C-terminal S-[(Z)-aminovinyl]-D-cysteine (AviCys) residues. Both EpiD and MrsD are able modify the prepropeptides of epidermin and mersacidin by oxidative decarboxylation of C-terminal cysteines to produce AviCys. The crystal structures of these enzymes revealed that they contain a flavin mononucleotide (Schmid *et al.*, 2002, Patton & van der Donk, 2005). Unlike the LanBC and LanM enzymes, EpiD does not have an absolute requirement for the leader sequence as a substrate, and is able to decarboxylate both the EpiA prepropeptide and the peptide without the leader sequence (the propeptide) (Chatterjee *et al.*, 2005). The ability of EpiD to decarboxylate an even wider range of substrates was demonstrated by a study that used a library of heptapeptides with varying amino acid substitutions in the C-terminal region of EpiA. This study concluded that only the C-terminal cysteine was necessary for decarboxylation. As modification requires a free thiol and carboxylate for activity, it is suggested that oxidative decarboxylation and dehydration occur simultaneously, which is then followed by ring formation (Kupke *et al.*, 1995, Schmid *et al.*, 2002).

1.1.7 Other lantibiotic modifications

As mentioned previously (see figure 1.1), a number of other amino acid modifications are found in many lantibiotics. Chatterjee *et al.* (2005) propose that after the removal of the leader sequence, the non-enzymatic hydrolysis of Dhb at position 1 of lactocin S and epilancin K7 leads to the formation of 2-oxobutyryl and 2-oxopropionyl groups, respectively. In epicidin 280, the reduction of 2-oxopropionyl to 2-hydroxypropionyl is thought to be carried out by an oxidoreductase, EciO, encoded by the epicidin 280 cluster, but it is not known how these enzymes function (Heidrich *et al.*, 1998, van de Kamp *et al.*, 1995, Chatterjee *et al.*, 2005).

The lantibiotic cinnamycin contains two modifications that are unique to the duramycin family of lantibiotics to which it belongs. A lysinoalanine bridge is formed by the dehydration of Ser to Dhb followed by the conjugate addition of Lys19. While potentially analogous to lanthionine bridge formation, it is not known whether *cinM* or one of the genes of unknown function in the *cin* cluster is responsible for this unusual modification (Chatterjee *et al.*, 2005). Cinnamycin (and the other duramycins) also contain an *erythro*-3-hydroxy-L-aspartic acid resulting from the hydroxylation of L-Asp15 and is essential for recognition of its target phosphatidyl ethanolamine (Fredenhagen *et al.*, 1990, Chatterjee *et al.*, 2005). A mammalian ß-hydroxlase is responsible for this modification in a number of mammalian proteins, and it is thought that a homolog in the cinnamycin cluster may perform the same function (Chatterjee *et al.*, 2005).

1.1.8 LanP and LanT

Lantibiotics are generally transported out of the cell by a dedicated lantibiotic ATP-dependent transporter LanT (Chatterjee *et al.*, 2005). NisT, the transporter encoded by the nisin gene cluster, is essential for extracellular transport; its deletion from *L. lactis* N8 led to the accumulation of processed nisin in the cytoplasm (Qiao & Saris, 1996). In other clusters, a dedicated lantibiotic transporter is not absolutely essential. For example, deletion of *pepT* resulted in a 90% reduction in Pep5 transport; presumably in its absence, transporter(s) encoded elsewhere in the genome are responsible for the residual ~10% of lantibiotic export (Meyer *et al.*, 1995). Indeed, some clusters do not appear to encode a transporter at all, as is the case for epicidin 280 (Meyer *et al.*, 1995, Heidrich *et al.*, 1998). Many lantibiotic transporters exhibit relaxed substrate specificity. Fusion of the nisin leader sequence with unmodified forms of nisin, as well as with peptide fragments of the non-lantibiotic peptides enkephalin and angiotensin, were transported by NisT (Kuipers *et al.*, 2004). In another study, an alkaline phosphatase was attached to the subtilin leader sequence and was exported in *B. subtilis*. The amount of alkaline phosphatase exported was increased when SpaT was present, demonstrating that LanT transporter systems may possess marked substrate flexibility, and that the leader sequence is important in substrate recognition (Izaguirre & Hansen, 1997, Chatterjee *et al.*, 2005).

The type AI ATP cassette (ABC) transporters represented by NisT and SpaT (for export of nisin and subtilin, respectively) are roughly 600 amino acids in length and have significant homology to haemolysin B-like ATP-dependent transport proteins found in a number of other organisms (Chung *et al.*, 1992, Gentschev & Goebel, 1992, Chatterjee *et al.*, 2005). They contain a hydrophobic N-terminal domain and a six-helix membrane-spanning domain. They also possess a C-terminal ATP-binding domain (Chatterjee *et al.*, 2005).

The type AII clusters contain slightly larger, 700 amino acid LanTs, such as LctT, and MutT (Rince *et al.*, 1994, Chen *et al.*, 1999). Transporters from these clusters contain an additional N-terminal peptidase domain besides the C-terminal ATP-binding and membrane-spanning domains that has been implicated in cleaving the leader sequences from the mature lantibiotic (Chatterjee *et al.*, 2005). The absence of a dedicated LanP protease in these gene clusters, with the exception of lactocin

S, which contains a truncated LanT and a LanP protease, suggests that the N-terminal peptidase domain is responsible for the final lantibiotic processing step.

Type B lantibiotic gene clusters contain two component transport systems consisting of a permease component and ATP-binding subunit. These are present in the gene clusters of cinnamycin, actagardine, deoxyactagardine B and microbisporicin (Boakes *et al.*, 2010, Boakes *et al.*, 2009, Foulston & Bibb, 2010, Widdick *et al.*, 2003). The transporters for microbisporicin have high homology to those involved in cinnamycin export (Foulston & Bibb, 2010). In cinnamycin, the ATP binding subunit is a member of the pfam family of ABC transporters, and has 36% identify to MrsF (part of a putative MrsFEG immunity transporter in mersacidin) and the permease subunit is an integral membrane subunit and is a member of the pfam family of ABC-2 type transporters usually involved in drug efflux and resistance, or carbohydrate export (Widdick *et al.*, 2003).

The final step of lantibiotic maturation is the removal of the leader sequence from the fully processed lantibiotic and as implied above, this may be carried out by one of two lantibiotic enzymes, either by the serine-type protease LanP, or by the protease domain of the ATP-binding cassette LanT (Patton & van der Donk, 2005). Some Type B lantibiotic gene clusters (such as cinnamycin, actagardine and microbisporicin) contain no LanP protease and no protease domains within the LanT (Boakes *et al.*, 2010, Boakes *et al.*, 2009, Foulston & Bibb, 2010, Widdick *et al.*, 2003). Subtilin is unusual in that its gene cluster encodes both a LanT and LanP, neither of which is responsible for cleavage of the leader sequence from the lantibiotic (Patton & van der Donk, 2005). Instead, three extracellular serine proteases were identified (subtilisin, WprA and Vpr) that are able to process the subtilin precursor peptide (Corvey *et al.*, 2003). Interestingly, while the nisin transporter NisT was able to secrete prenisin, dehydrated NisA, dehydrated and cyclised NisA, and the NisA leader sequence fused to non-lantibiotic peptides (Kuipers *et al.*, 2004), the NisP protease would only cleave the leader peptide attached to fully matured nisin (Kuipers *et al.*, 2004, Patton & van der Donk, 2005).

1.1.9 LanFEG / LanI

Once exported from the cell and cleaved from the leader sequence, the cell must then protect itself from the mature lantibiotic. To date, two systems have been identified that are thought to be involved in host cell immunity. These are the LanFEG transporter system and LanI peptide system. Host producers can have one or both of these systems. For nisin and subtilin, both systems act independently to confer immunity (Chatterjee *et al.*, 2005). The *lanFEG* genes encode an ABC transporter which is able to remove nisin and subtilin from the cytoplasmic membrane and export it out into the extracellular space (Nagao *et al.*, 2006). NisA and SpaI are lipoproteins which bind to and sequester their respective lantibiotics on the outside of the cytoplasmic membrane, preventing them from reaching their target. NisI has no homology with other non-lantibiotic proteins and has a hydrophobic N-terminus containing a consensus lipoprotein sequence (Kuipers *et al.*, 1993). It is membrane anchored by post-translational removal of the first 19 amino acids as well as by lipid modification of the Cys residue at the new N-terminus (Chatterjee *et al.*, 2005). The LanI lipoproteins orientate to the outside of the cytoplasmic membrane and sequester the lantibiotic by binding to it, but most are unable to provide cross-resistance against other lantibiotics (Nagao *et al.*, 2006). NukH is an exception; this unique immunity protein consists of three transmembrane domains and is able to interact with and provide immunity towards both nukasin and lacticin 481, but not nisin (Nagao *et al.*, 2006, Stein *et al.*, 2005, Stein *et al.*, 2003).

Interestingly, some non-lantibiotic producing *L. lactis* strains are able to resist high levels of nisin. These strains express a 35 kDa nisin resistance protein (NSR) that is able to proteolytically cleave six amino acids from the carboxyl “tail” of nisin, reducing its bacteriocidal activity 100-fold (Sun *et al.*, 2009). *Listeria monocytogenes* is also able to resist high levels of nisin. Enhanced resistance to nisin and other class IIa bacteriocins was associated with the increased expression of three genes encoding a penicillin-binding-protein (PBP), a histidine kinase and a protein of unknown function (Gravesen *et al.*, 2004).

1.1.10 Regulation of Lantibiotic Biosynthesis

Lantibiotic production must be tightly regulated to ensure that there is an appropriate balance between production and resistance and that the peptides are produced at an appropriate time, usually during late exponential growth (Chatterjee *et al.*, 2005). To aid in tight regulation, lantibiotic biosynthetic genes are located in operons within their gene clusters, where a few promoters are able regulate biosynthesis (see figure 1.9).

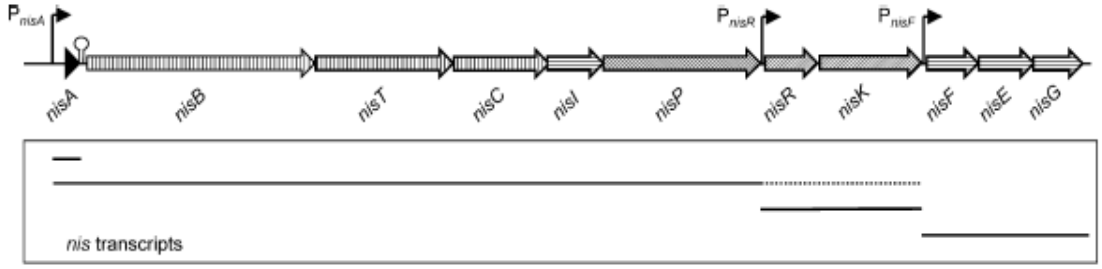


Figure 1.9: The transcriptional operons of nisin. The *nisA* promoter drives expression of *nisABTCIP*, with a transcriptional attenuator situated between the structural gene *nisA* and those encoding modification enzymes and a transporter. The *nisR* promoter drives *nisRK* and the *nisF* promoter drives *nisFEG*. Figure taken from Kleerebezem (2004).

The majority of lantibiotics studied to date use a two component regulatory system to regulate the production on the lantibiotic in response to extracellular signals. The most studied two component regulatory systems are those of nisin and subtilin. There is a reported 26.6% amino acid sequence identity between NisK and SpaK, both of which belong to the EnvZ-like subgroup of sensor kinase proteins (Chatterjee *et al.*, 2005, Kleerebezem, 2004). These proteins are predicted to contain two hydrophobic transmembrane helices at their N-terminal “input” domain, and conserved histidine residues at the C-terminal “transmitter” domain. The N-terminal domains have been shown to interact with their respective lantibiotic and transmit the signal to the appropriate response regulator via the transmitter domain (Kleerebezem, 2004). The response regulators NisR and SpaR have even higher homology, 41.3% amino acid identity, and belong to the OmpR-like subgroup of response regulators. Both proteins contain a conserved aspartate residue that is targeted by sensor kinase mediated phosphotransfer in the N-terminal “receiver” domain. The C-terminal “output” domain of both proteins contains a typical helix-turn-helix motif that is involved in binding the phosphorylated response regulators to

their target DNA sequences (Kleerebezem, 2004). The response regulators are able to bind to promoters within each lantibiotic biosynthetic gene cluster. In the case of nisin, the response regulator binds to two sites, P_{nisA} and P_{nisF}. The NisRK operon is expressed constitutively (see figure 1.9).

Nisin is produced late in the exponential growth phase, similar to many other lantibiotics (Chatterjee *et al.*, 2005). Nisin is able to auto regulate its own transcription acting as a quorum sensor, and as little as five nisin molecules are sufficient to activate transcription (de Ruyter *et al.*, 1996). The NisK sensor kinase is able to bind nisin and transmit the signal to the response regulator NisR. This is then able to promote the transcription of the P_{nisA} and P_{nisF} promoters and increase production of nisin (Kleerebezem, 2004). Galactose and lactose are also able to stimulate *nisA* transcription without externally applied nisin, in a NisRK independent manner (Chandrapati & O'Sullivan, 1999). An overview of the entire process can be seen in figure 1.10. Transcription of the subtilin *lanRK* genes depends on RNA polymerase subunit σH, which is derepressed at the end of exponential growth (Stein *et al.*, 2002).

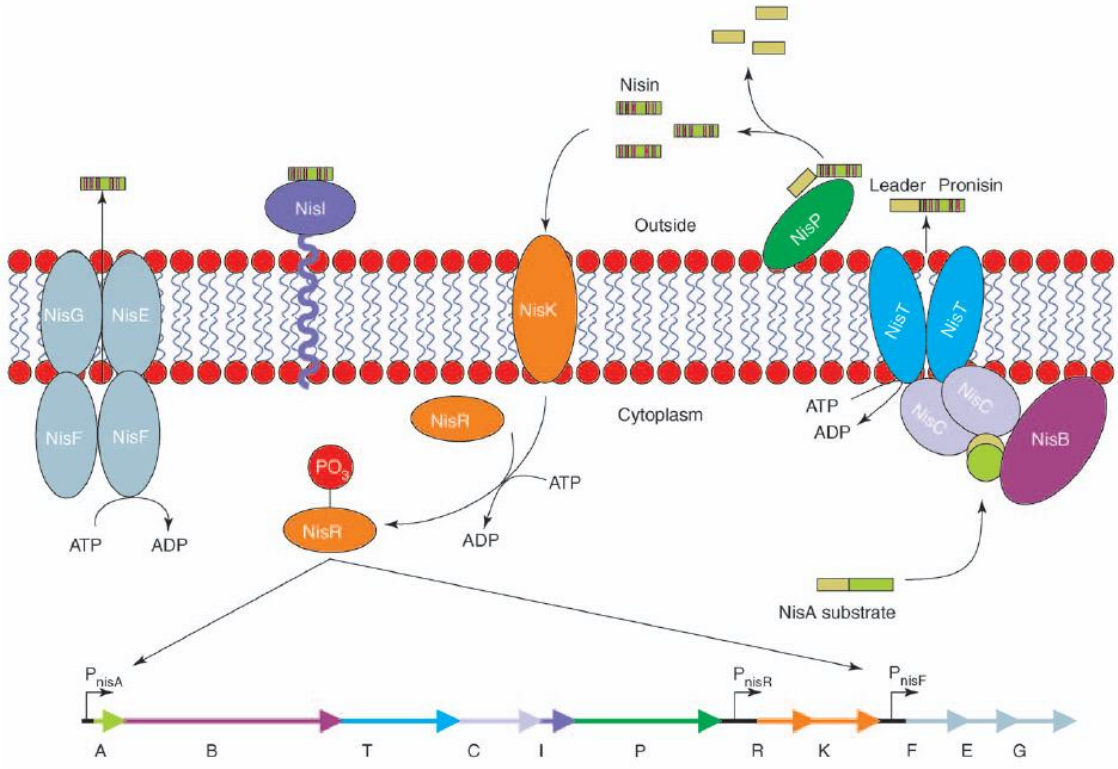


Figure 1.10: Overview of nisin biosynthesis. Extracellular changes in nisin concentration lead to the autophosphorylation of NisK which then transfers the phosphate to the Asp residue of NisR. This induces the expression of the *nisABTCIP* and *nisFEG* operons. Prepronisin is then processed by the NisBCT

multi-enzyme complex and exported from the cell, where the leader sequence is cleaved by NisP, producing mature nisin. NisI acts to sequester nisin, while NisFEG removes the lantibiotic into the extracellular environment. Figure taken from Patton and van der Donk (2005).

Regulation of the Type B lantibiotic mersacidin is different. While it does contain a two component regulatory system, MrsR2 and MrsK2, that shares homology with other LanRKs, it also contains an additional MrsR1 regulator (Guder *et al.*, 2002). Unlike other lantibiotic two component regulatory systems, MrsRK does not regulate mersacidin production; instead it is responsible for the regulation of the MrsFEG immunity transporter. The single regulator MrsR1 is believed to be involved in regulating mersacidin biosynthesis. Currently, the stimulus for its activation is unknown, nor is it clear whether a thus far unidentified sensor kinase is required for its phosphorylation (Chatterjee *et al.*, 2005). Regulation of two other Type B lantibiotics involves a SARP (*Streptomyces* antibiotic regulatory protein) for cinnamycin production in *Streptomyces cinnamoneus* (Widdick *et al.*, 2003), and an extracytoplasmic function σ factor–anti-σ factor complex for microbisporicin production in *Microbispora corallina* (Widdick *et al.*, 2003, Foulston & Bibb, 2010).

1.2 Lantibiotic modes of action

Lantibiotics commonly exert a combined set of cell killing mechanisms. These generally involve either binding to the cell membrane followed by insertion, or the use of a receptor/docking molecule to exert a structural based activity (Asaduzzaman & Sonomoto, 2009). In model membranes systems using cytoplasmic and artificial membrane vesicles, pore formation was determined as the mode of action of nisin and other cationic type AI lantibiotics (Chatterjee *et al.*, 2005). The first step in nisin activity was binding of the peptide to the cell membrane. In nisin, the positive charges that are found at the N- and C- termini attract the molecule to the membrane (Giffard *et al.*, 1997). The electrostatic differences between nisin and the membrane encourage the initial reaction (Asaduzzaman & Sonomoto, 2009). The C-terminus of nisin holds a strong positive charge, and interacts with anionic lipids in the membrane; it is also the C-terminus of nisin that first translocates across the membrane (Breukink *et al.*, 1997, van Kraaij *et al.*, 1998). Once this initial contact is made, the N-terminal rings of nisin bind to the disaccharide-polyphosphate of lipid II, and the C-terminus binds to the head-groups of the lipids within the bilayer to form a pore (Asaduzzaman & Sonomoto, 2009). There are two well-established models to explain how the pores form within the bacterial cell membrane. These are the barrel-stave and wedge models (Asaduzzaman & Sonomoto, 2009). In the barrel-stave model, nisin monomers (the stave) are bound to the cell membrane via their electrostatic charge, eventually forming a preaggregate. The pores are formed when a certain membrane potential is reached when the nisin molecules become perpendicular to the membrane and a pore (the barrel) is formed (Jung & Sahl, 1991). In the wedge model, surface bound nisin (in parallel) generates local strain on the lipid molecules, forcing the formation a pore (see figure 1.11) (Driessen *et al.*, 1995). The pores last for several hundred milliseconds, are between 0.2 to 1-2 nm in size and promote the breakdown of membrane potential and efflux of metabolites (Chatterjee *et al.*, 2005).

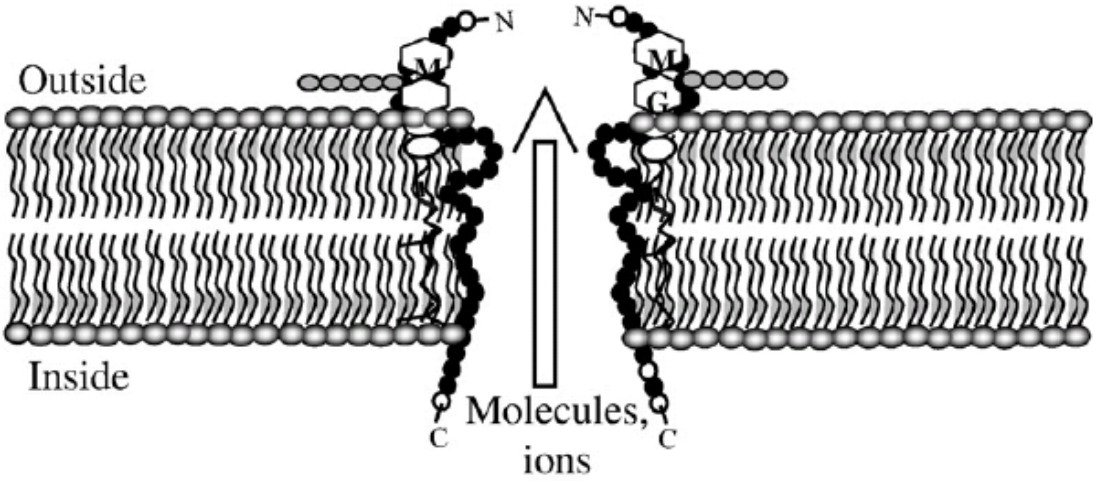


Figure 1.11: The wedge model of nisin pore formation. Nisin is electrostatically bound in parallel to the membrane, and local strain forces a pore through the cell wall. Figure taken from Asaduzzaman and Sonomoto (2009).

Many lantibiotics bind to the essential cell wall precursors Lipid I and Lipid II and interfere with peptidoglycan biosynthesis (Chatterjee *et al.*, 2005). This prevents the utilisation of these molecules by transpeptidase and transglcosylase enzymes in building the cross-linked network of the cell wall (Asaduzzaman & Sonomoto, 2009). Vancomycin is able to kill bacteria by binding to the D-ala-D-ala moiety of lipid II. Nisin, under normal circumstances, binds to lipid II via its disaccharides-pyrophosphate region, making it effective against strains that are resistant to vancomycin (Hsu *et al.*, 2003). The unique ability of nisin to be effective at nanomolar concentrations is due to its ability to bind to lipid II and to form a pore using the latter as a docking molecule (Breukink *et al.*, 1999). Recent NMR data has shown that nisin interacts with lipid II using the backbone amides of rings A and B via 6 hydrogen bonds to the pyrophosphate moiety of lipid II (Hsu *et al.*, 2004). The A and B rings of nisin are conserved in subtilin, epidermin, gallidermin and plantaricin C (Asaduzzaman & Sonomoto, 2009). A schematic for this interaction is shown in figure 1.12. The structure of the lantibiotic is key to its bioactivity. For example, the specific substitution of the dehydro amino acid Dha from position 5 of nisin resulted in a complete loss of sporicidal activity although bacteriolytic activity remained (Pag & Sahl, 2002, Chan *et al.*, 1996, Dodd *et al.*, 1995).

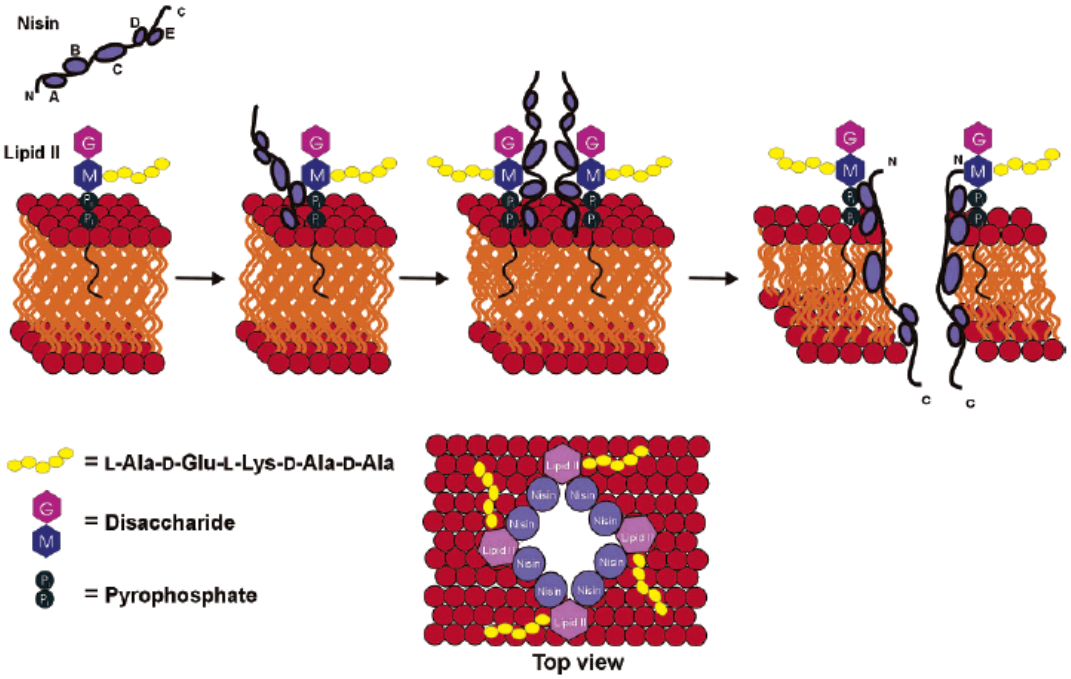


Figure 1.12: Proposed model for nisin-lipid II interaction. The N-terminus of nisin binds to lipid II and forms a pore within the membrane composed of four lipid II and eight nisin molecules. Figure taken from Chatterjee, Paul *et al.* (2005).

Type B lantibiotics such as mersacidin and actagardine are also able to interact with Lipid II but have no structural similarity with nisin. By binding to lipid II they are able to disrupt cell wall biosynthesis by blocking the transglycosylation step of peptidoglycan biosynthesis (Brotz *et al.*, 1997, Boakes *et al.*, 2010, Boakes *et al.*, 2009). These compounds do not work at nanomolar concentrations because they do not form pores, but as mersacidin is effective against MRSA, it does have potential clinical applications.

The two component lantibiotic lacticin 3147 works at nanomolar concentration, and requires a 1:1 stoichiometry of its subunits (LtnA1:LtnA2). LtnA1 first binds to lipid II where it recruits LtnA2 which inhibits cell wall biosynthesis and creates pores in the membrane (Wiedemann *et al.*, 2006).

Cinnamycin and the duramycins are also unique among lantibiotics, exerting their antibacterial effect by binding to phosphatidylethanolamine (PE) and disrupting membrane function (Fredenhagen *et al.*, 1990, Machaidze & Seelig, 2003). In mammalian cells, and probably also in bacteria, PE is predominantly found on the inner leaflet of the plasma membrane. Upon binding to PE, cinnamycin induces transbilayer lipid movement, flipping the PE to the outer leaflet (Makino *et al.*, 2003).

Binding of cinnamycin induces reorganisation of the membrane into highly deformed curved structures that disrupt normal membrane function (Iwamoto *et al.*, 2007).

1.3 Lantibiotic Engineering

Lantibiotics have several key features that make them targets for selective engineering. All lantibiotics are ribosomally synthesised, and the amino acids encoded by the *lanA* gene are directly incorporated into the lantibiotic. Consequently the structures of these antibiotics can be readily manipulated using well-understood and established DNA manipulation techniques such as site-directed mutagenesis, and oligonucleotide synthesis. These approaches could be used, for example, in attempts to improve lantibiotic activity, stability and pharmacokinetic properties (Nagao *et al.*, 2006). In the future, the lantibiotic modification machinery could also be used to introduce dehydrated amino acids into other peptides and to generate novel cyclic peptides. This is made possible by the relative plasticity of the lantibiotic synthetases, and their ability to recognise non-lantibiotic peptides attached to leader sequences.

The generation of lantibiotic derivatives is most efficiently achieved by varying the sequence of the *lanA* gene. The first systematic mutational analysis of a lantibiotic, replacing targeted amino acids with alanine residues, was performed on the two-component lantibiotic lacticin 3147. The results suggested that various parts of the lantibiotic prepropeptide were amenable to amino acid changes (Cotter *et al.*, 2006). Mersacidin (figure 1.4) was the first lantibiotic to undergo systematic exchange of all of the amino acids in MrsA (except those involved in bridge formation) (Appleyard *et al.*, 2009). Appleyard *et al.* (2009) generated 228 mersacidin mutants (figure 1.13) of which 80 were produced at good levels. Variants with insertions of amino acids into ring A and flanking regions lost all antimicrobial activity, but ring B was highly amenable to insertions and substitutions. This study concluded that the geometry of the molecule was highly evolved to bind to lipid II. Most of these variants did not improve the antimicrobial activity of mersacidin, but some did perform better against some Gram-positive pathogens, suggesting this method may be a useful approach in lantibiotic improvement programs (Appleyard *et al.*, 2009).

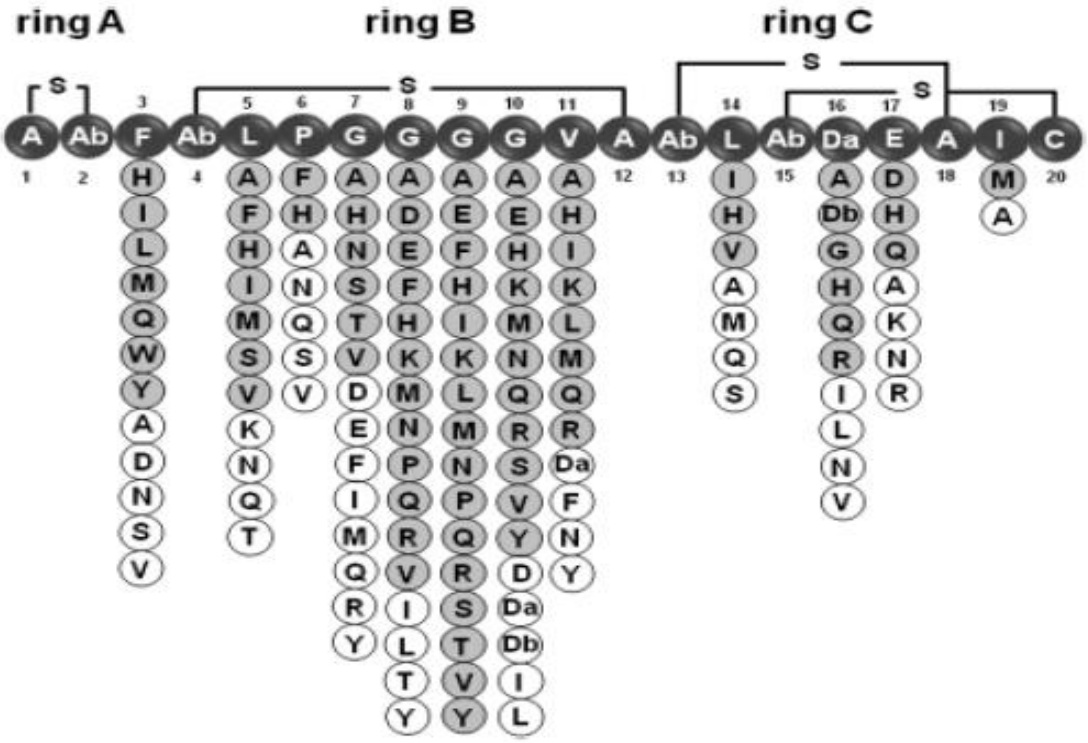


Figure 1.13: Mersacidin substitution library. Circles under the mersacidin peptide indicate variants that were detected by LC-MS analysis. Grey circles represent variants that were produced at high yield (>10% of mersacidin production) and open circles represent those produced at trace levels. Da, Dha; Db, Dhb; Ab, 2-aminobutyrate. Figure taken from Appleyard *et al.* (2009).

1.4 Actagardine

Actagardine was first isolated from *Actinoplanes garbadinensis* and also thought to be made by *Actinoplanes liguriae* in 1976 (Arioli *et al.*, 1976, Coronelli *et al.*, 1976, Parenti *et al.*, 1976). Actagardine was formerly known as gardimycin as the two producer strains of *Actinoplanes* were both isolated from garden soil samples from the locality of Garbady in India (hence *A. garbadinensis*) and the Italian region of Liguria (hence *A. liguriae*) (Parenti *et al.*, 1976). Actagardine was first identified as a peptide antibiotic because of its antimicrobial activity. Treatment of experimental infections in mice using *Streptococcus pneumoniae*, *Streptococcus haemolyticus* and *Staphylococcus aureus* demonstrated that actagardine had high therapeutic effectiveness, as well as a low toxicity, potentially making it a good candidate for clinical development (Arioli *et al.*, 1976). Further experiments indicated that the frequency of *S. haemolyticus* mutants that were resistant to actagardine was less than 10^{-10}. When 10^{9} cells were cultured on Brain-Heart infusion agar containing 5 µg/ml actagardine (normal MIC is 1 µg/ml), no resistant mutants were obtained. Only after training (growth in Brain-Heart infusions with increasing concentrations of actagardine (data not shown)) were mutants resistant to 20 µg/ml actagardine obtained, and these exhibited no cross-resistance to penicillin, bactitracin, ristocetin, vancomycin, rifampicin and erythromycin (Arioli *et al.*, 1976). Recently, Novacta Biosystems have conducted further research into the genes responsible for the biosynthesis of actagardine (Boakes *et al.*, 2010, Boakes *et al.*, 2009).

1.4.1 Structural characterisation and mode of action

Actagardine inhibits cell wall biosynthesis, and a number of experiments were carried out to understand its mechanism of action. Experiments were conducted on *B. subtilis*, where the addition of 100 µg/ml actagardine caused almost immediate cell lysis, which was measured by a reduction in culture absorbence (Somma *et al.*, 1977). In this study, the addition of actagardine to a growing culture immediately suppressed the incorporation of GlcNAc, a specific precursor of peptidoglycan, into the cell wall of *B. subtilis*. Radio-labelled alanine, used to determine whether actagardine disrupted peptidoglycan biosynthesis rather than GlcNac uptake or utilisation, was not incorporated into the cell wall; protein biosynthesis was unaffected, unlike tetracycline-treated controls. The observation that actagardine has a bactericidal effect on actively dividing cells, but not resting cells that were not

actively dividing, was also consistent with inhibition of cell wall synthesis, although the report did not comment on the lytic ability of actagardine on stationary phase cells (Somma *et al.*, 1977).

The initial physicochemical characterisation of actagardine was carried out using a combination of infrared and U.V. spectroscopy, and acid hydrolysis. These procedures identified the majority of amino acids present in the peptide and also detected two unknown amino acids containing sulphur moieties that presumably arose from two of the four lanthionine rings present in actagardine (Coronelli *et al.*, 1976). The first major step forward in structurally characterising the lantibiotic took place in 1990 when Kettenring *et al.* (1990) used homonuclear 2D NMR spectroscopy to establish the first, but incorrect, structural model of actagardine (figure 1.14) (Kettenring *et al.*, 1990).

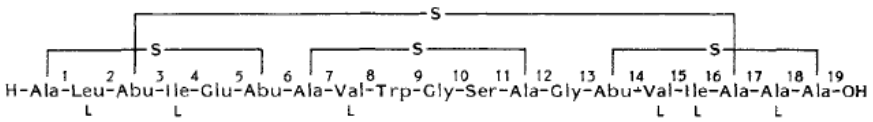


Figure 1.14: The proposed sequence of actagardine based on 2D NMR. Figure taken from Kettenring *et al.* (1990).

This study correctly identified a number of the amino acids as well as the sulphur bridges, but was unable to assemble a complete structure. The structure was eventually determined by a study carried out by Zimmermann *et al*. (1995), and further refined in a later publication (Zimmermann *et al.*, 1995, Zimmermann & Jung, 1997). The complete structure of actagardine (figure 1.15) was obtained by combining a number of techniques including mass spectrometry, chemical modification, Edman degradation and NMR.

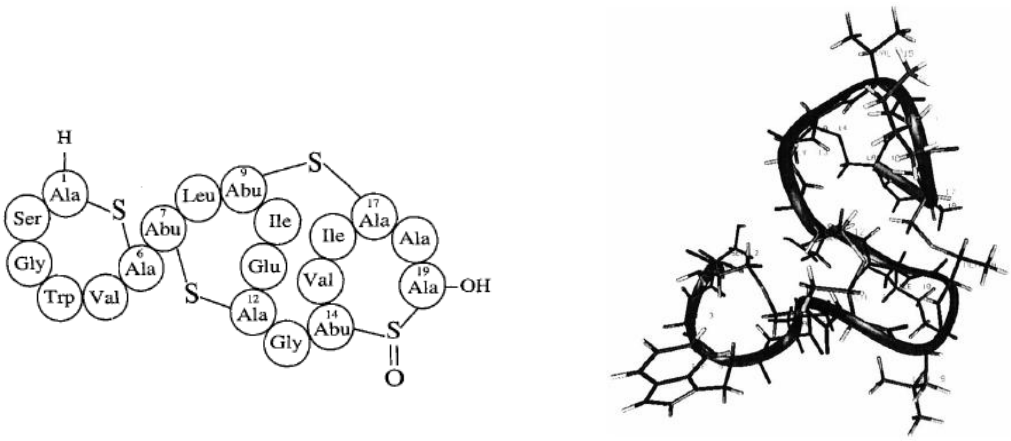


Figure 1.15: The structure of actagardine. Figure adapted from Zimmermann and Jung (1997).

Zimmermann *et al.* (1997) characterised the peptide as having two distinct thioether ring systems, linked by a short linear segment, leading to the formation of two possible binding pockets. The first pocket is lined by amide moieties built from the cyclisation of the C-terminal ring system. The second consists of the N-terminal ring, which is perpendicular to the first, and two hydrophobic residues, Glu11 and Ser2, which are positioned in the centre. This pocket is located within the region that shows sequence similarity with mersacidin. The sequence of actagardine from Gly3 to Ala12 is very similar to the sequence of mersacidin residues Gly9 to Ah18 (see figure 1.15). These data supported the hypothesis that actagardine was able to inhibit the transfer of the disacchardic pentapeptide from the lipid carrier to the acceptor site which is present on the cell wall (Somma *et al.*, 1977, Zimmermann & Jung, 1997). It was also noted that the unusual structure of this molecule made it a good lead structure when trying to design other peptidoglycan biosynthesis inhibitory compounds (Zimmermann & Jung, 1997).

1.4.2 Actagardine derivatives

To further explore the use of actagardine as an antimicrobial agent, Malabarba *et al*. (1990) produced 13 amide derivatives of actagardine using synthetic chemistry. They produced some monocarboxamides (2 – 10), two diamides (11 and 12) and a *N*-acyl derivative. These derivatives were then tested against a number of different bacteria to test their biological activity.

Organism	MIC (μg/ml)												
	2	3	4	5	6	7	8	9	10	11	12	13	1 (actagardine)
Staphylococcus aureus Tour	6.2	6.2	12.5	50	50	25	12.5	25	6.2	12.5	25	100	25
S. aureus Tour[b]	25	25	25	50	50	50	25	50	12.5	25	25	>100	50
S. aureus Tour[c]	3.1	3.1	6.2	12.5	6.2	25	12.5	12.5	6.2	6.2	12.5	>100	25
Streptococcus pyogenes C 203 SKF 13400	0.2	0.2	0.4	0.8	0.8	0.4	0.8	0.4	0.8	0.8	0.8	12.5	1.6
S. pneumoniae UC 41	12.5	12.5	6.2	12.5	25	6.2	6.2	6.2	6.2	12.5	12.5	25	25
S. mitis L 1320[d]	3.1	3.1	3.1	3.1	6.2	3.1	0.8	1.6	0.8	nd	nd	nd	12.5
S. faecalis L 1321[d]	25	25	50	100	100	25	50	25	12.5	nd	nd	nd	100
S. sanguis L 1322[d]	50	50	50	25	100	50	12.5	50	6.2	nd	nd	nd	100
S. sanguis L 1324[d]	50	25	50	50	100	50	12.5	50	6.2	nd	nd	nd	100
S. salivarius L 1323[d]	0.4	0.4	0.8	0.8	6.2	0.4	0.05	0.8	0.05	nd	nd	nd	3.1
S. bovis L 1325[d]	50	50	25	25	25	25	12.5	25	12.5	nd	nd	nd	100

[a] See Experimental section.
[b] Inoculum 10^6 cfu/ml.
[c] In the presence of 30% bovine serum.
[d] Clinical isolates.
nd: Not determined.

Figure 1.16: *In vitro* activity of a number of amide derivatives of actagardine. Figure taken from Malabarba *et al.* (1990).

Figure 1.16 shows that a number of these derivatives have lower MICs against some of the bacteria tested than actagardine. Only one compound (4) was taken forward into a mouse model, where it did comparatively well against a number of different pathogens. It would be interesting to know if the toxicity of these compounds was also tested.

Shortly after the structural elucidation of actagardine, it was reported that *A. liguriae* ATCC 31048 produced a similar lantibiotic, Ala(0)-actagardine (Vertesy *et al.*, 1999). This lantibiotic was produced when *A. liguriae* ATCC 31048 was cultured on mannitol and soya meal, and was identified by a combination of solid phase extraction and two-step chromatographic separation. Ala(0)-actagardine was further analysed by 2D-NMR, amino acid analysis and Edman degradation, as well as by partial synthesis from actagardine (Vertesy *et al.*, 1999). Compared to actagardine, it has an additional N-terminal alanine residue, and was reported to have higher antimicrobial activity and a different spectrum of activity. Using synthetic methods, Vertesy *et al.* were also able to produce Ala(0)-actagardine by

modifying actagardine, dissolving it in anhydrous dimethylformamide and treating it with Boc-Ala-O-*N*-hydroxysuccinimide. Along with Ala(0)-actagardine, two other analogues were made; Lys(0)-actagardine and Ile(0)-actagardine using similar methods. These new derivatives were in some cases reported to have better antimicrobial activity than mersacidin against *Enterococci* (Vertesy *et al.*, 1999).

1.4.3 Actagardine homologues

Homologues of actagardine produced by other bacteria have also been reported (figure 1.17). The tomato pathogen *Clavibacter michiganensis* subsp. *michiganensis*, which is particularly effective at inhibiting the growth of a tomato pathogen *Clavibacter michiganensis* subsp. *sepedonicus*, produces the lantibiotic michiganin A (Holtsmark *et al.*, 2006). The structure proposed for michiganin A is based on and similar to the resolved structure of actagardine with a few key differences. There are six amino acid differences between the two lantibiotics, michiganin A has one less MeLan bridge, and the C-terminal MeLan bridge lacks the oxygen atom present in actagardine. However, the proposed structure may not be correct; the Dhb at position 8 and Cys residue at position 13 would appear to be amenable to Lan formation. The study carried out by Holtsmark *et al.* was unable to find any homology of the leader sequence of michiganin A with other type B lantibiotics. The similarity between the michiganin A and actagardine gene clusters is investigated in detail in chapter 5.

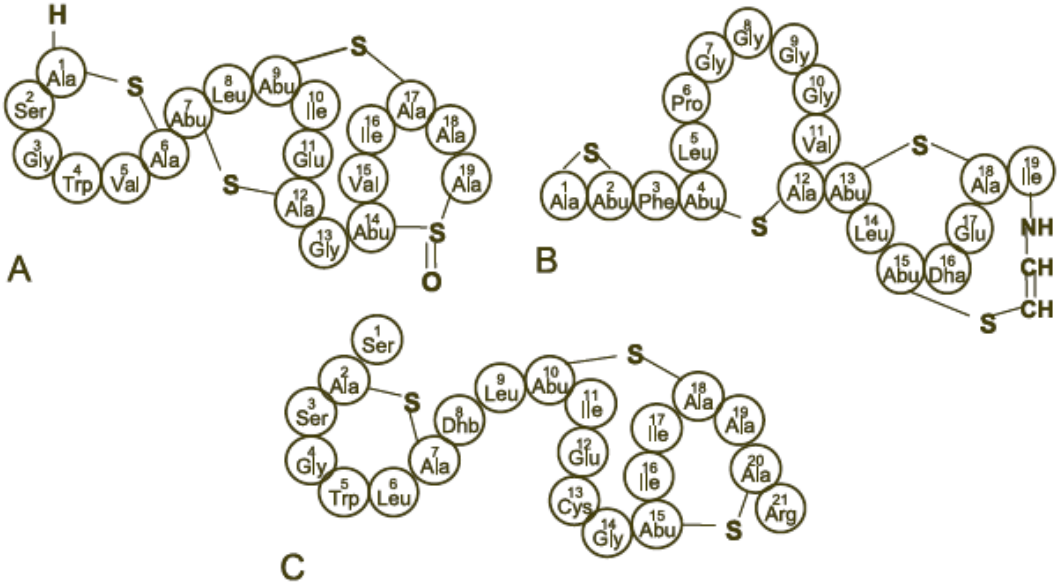


Figure 1.17: The structures of actagardine (A), mersacidin (B) and michiganin (C). Figure taken from Holtsmark *et al.* (2006).

1.4.4 Recent research

In recent research articles, Novacta Biosystems have conducted a number of experiments on the gene clusters of actagardine and DAB, providing the background and parallel work to this study. Below are the structures of the two lantibiotics now known to be made by the two *Actinoplanes* species.

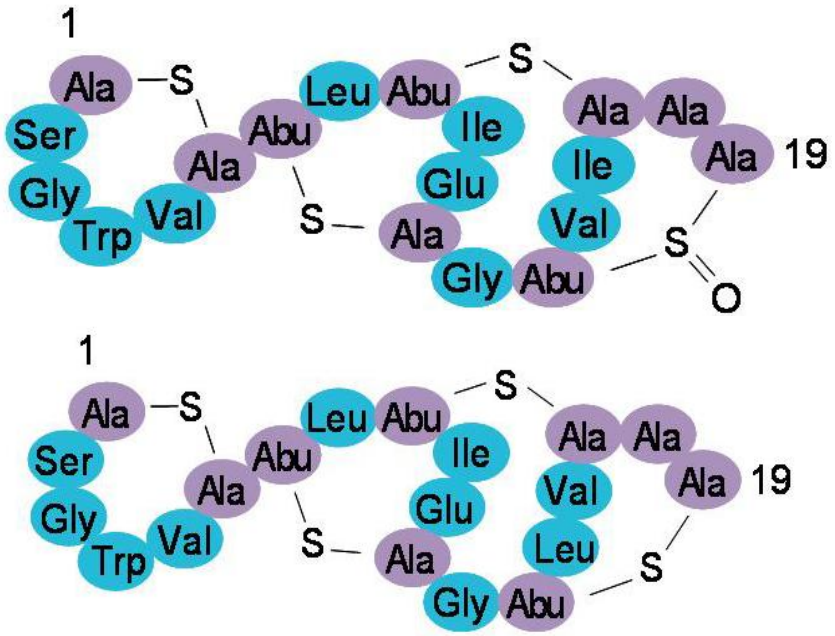


Figure 1.18: The structures of actagardine from *Actinoplanes garbadinensis* (top) and deoxyactagardine B (DAB) from *Actinoplanes liguriae* (bottom) (Boakes *et al.*, 2010, Boakes *et al.*, 2009).

Boakes *et al.* (2010) reported for the first time that *A. liguriae* does not produce actagardine, but instead the related and novel compound DAB, conflicting with earlier reports (Vertesy *et al.*, 1999). DAB differs from actagardine by the substitution of two amino acids; Val15 to Leu and Ile16 to Val, and the lack of a sulfoxide between residues 14 and 19 (Boakes *et al.*, 2010). The DAB gene cluster was discovered during probing of genome libraries of both *A. garbadinensis* and *A. liguriae* for the biosynthetic genes involved in actagardine biosynthesis. The probe hybridised to cosmid AL02, and subsequent sequencing revealed the variant *lanA* gene. This prompted a reinvestigation of the lantibiotics produced by *A. liguriae*, and mass spectrometry and NMR analysis confirmed the presence of the DAB peptide in the culture supernatants of *A. liguriae* (Boakes *et al.*, 2010). Cosmids containing the biosynthetic gene clusters of both actagardine and DAB have been

sequenced and annotated (figures 1.19 and 1.20) (Boakes *et al.*, 2010, Boakes *et al.*, 2009).

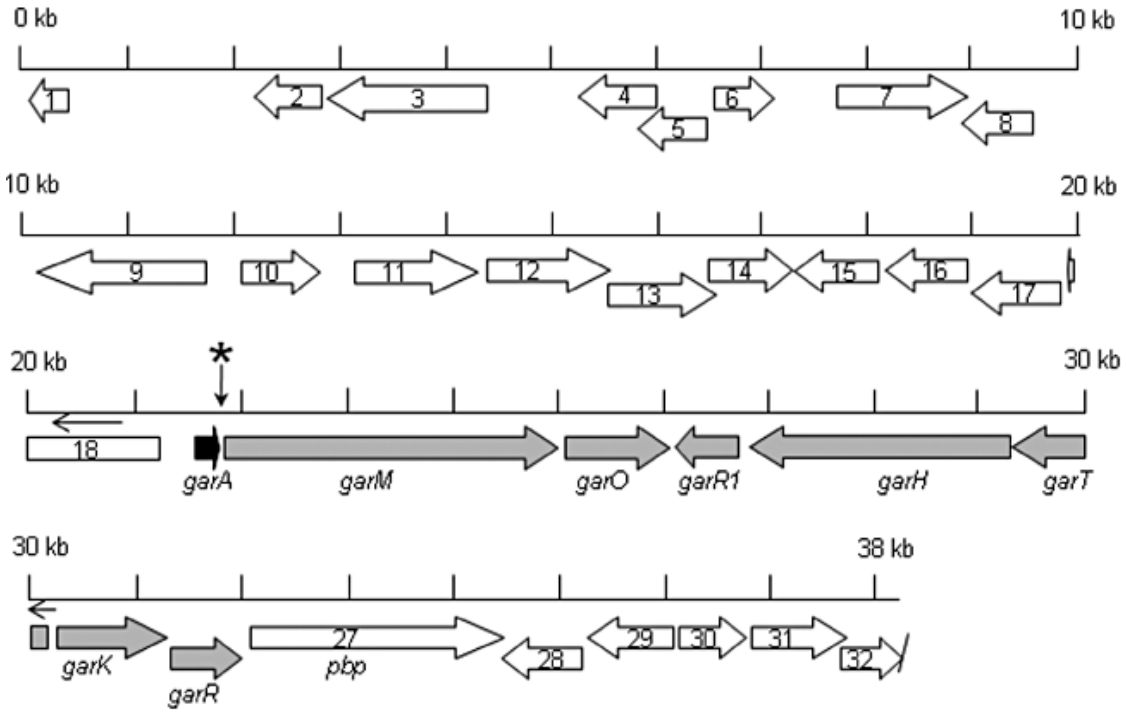


Figure 1.19: Annotation of the AG14 cosmid from *A. garbadinensis* containing the biosynthetic genes required for actagardine biosynthesis (see chapter 3 for detailed description). Figure taken from Boakes *et al.* (2009).

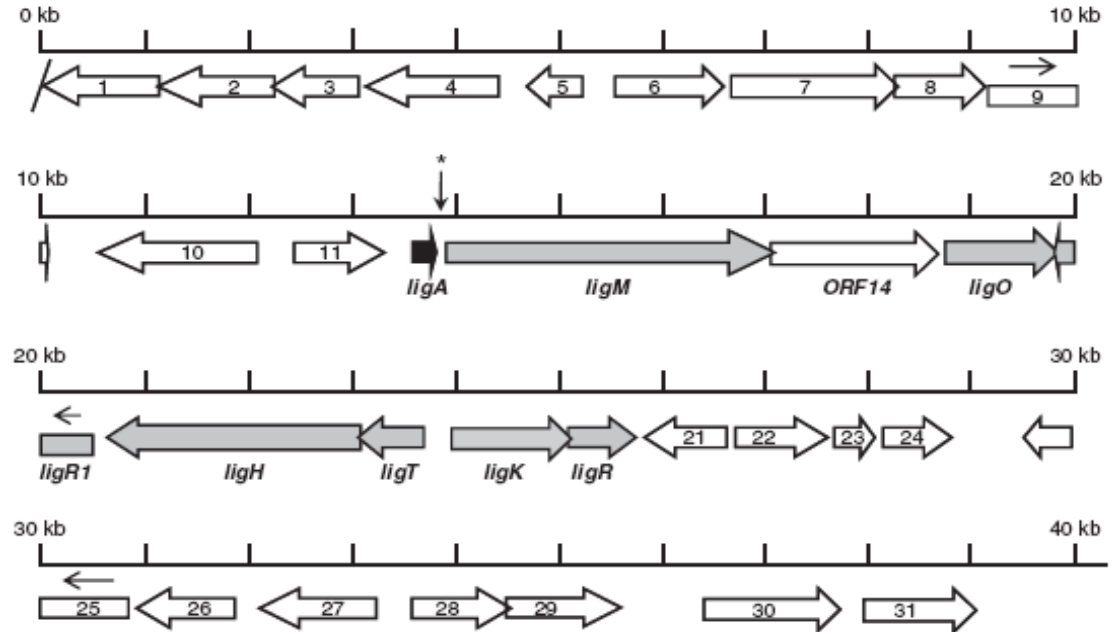


Figure 1.20: Annotation of the AL02 cosmid from *A. liguriae*, containing the biosynthetic genes required for DAB biosynthesis (see chapter 3 for detailed description). Figure taken from Boakes *et al.* (2010).

The gene clusters for actagardine and DAB biosynthesis contain homologues of many of the previously characterised lantibiotic biosynthetic genes described above. There is one exception to this, *garO*, which appears to encode a luciferase-like monooxygenase that was presumed to be responsible for incorporation of the sulfoxide group into the C-terminal methylanthionine bridge of actagardine (Boakes *et al.*, 2009). Deletion of this gene from *A. garbadinensis* resulted in the production of deoxyactagardine A (Boakes *et al.*, 2009). The gene cluster of DAB also contains the *garO* homolog *ligO*, but DAB does not contain the sulfoxide group, indicating that either DAB is not a substrate for LigO, or that the monooxygenase is non-functional. Expression of *ligA*, encoding the DAB prepropeptide, in a Δ*garA* strain of *A. garbadinensis* resulted in the production of the oxidised form of DAB, actagardine B, and the conclusion that DAB is a viable substrate for oxidation by at least GarO, prompting further speculation that LigO is non-functional (Boakes *et al.*, 2010).

The main focus of Novacta Biosystem's research was to generate potentially improved versions of actagardine by variant generation. Boakes *et al.* (2009) developed a plasmid, pAGvarX that enables the rapid generation of libraries of actagardine variants via mutagenesis. This strategy does not require the creation of double-cross over recombinant mutants as the construct is expressed *in trans* in a Δ*garA* mutant of *A. garbadinensis*. Using this plasmid, they constructed an alanine scanning library and the resulting variants were analysed for improved activity; they were also able to assess the flexibility of the actagardine biosynthetic machinery (figure 1.21).

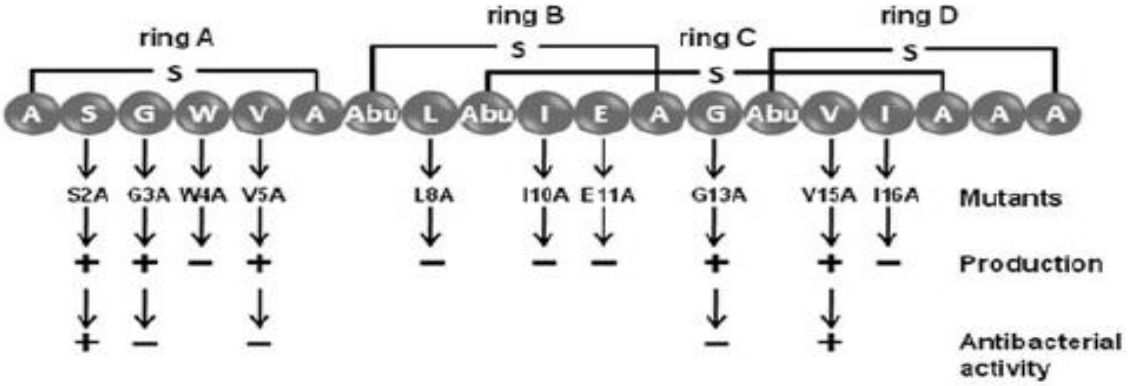


Figure 1.21: The results of alanine scanning of the actagardine propeptide. Abu, 2-aminobutyric acid. Figure taken from Boakes *et al.* (2009).

Figure 1.21 summarises the results of the alanine scanning of actagardine. Only two variants, S2A and V15A, showed any biological activity against the test organism, *Micrococcus luteus*.

The eventual hope is that actagardine and/or one of its variants will be used in a clinical setting to tackle multidrug and vancomycin resistant infections.

1.5 Aims and Objectives

The main aim of this study was to identify experimentally the genes required for the biosynthesis of actagardine and DAB.

Although the gene clusters for actagardine and DAB are very similar, there are some notable differences that warranted further investigation. The AL02 cosmid contains an additional gene, *orf14* (known in this study as *ligQ*) (figure 1.20) that encodes a transport protein. It is hypothesised that this gene may be involved in immunity to DAB. An analogous role might be played by *garPBP*, thought to encode a penicillin binding protein in the AG14 cosmid and which lies next to *garR* in the actagardine gene cluster (figure 1.19). Since homologues of *garPBP* have not been found in other lantibiotic gene clusters, its role, if any, in actagardine biosynthesis had to be determined experimentally. Like mersacidin, these gene clusters contain three regulatory genes, *lanRK* and *lanR1*, but their direct targets for regulation are not known. The role of these genes in lantibiotic production also required experimental verification. These mutational studies could also be complemented by the construction of minimal gene sets derived from both AG14 and AL02 to ensure that other genes within the two cosmids were not required for the biosynthesis of these lantibiotics.

At the beginning of this work, only the AG14 cosmid has been successfully expressed in a heterologous host. Attempts to express the AL02 cosmid in *S. lividans* 1326 had been unsuccessful (Novacta Biosystems, unpublished), warranting further attempts at heterologous expression and subsequent mutational analysis of the DAB gene cluster.

1.6 Summary of Research Aims

- To carry out a mutational analysis of all of the potential lantibiotic regulatory genes carried on the AG14 cosmid.

- To construct minimal gene sets for actagardine and DAB biosynthesis from cosmids AG14 and AL02, respectively, to identify the extent of the biosynthetic gene clusters.

- To attempt to express the DAB gene cluster heterologously.

- To use computational analysis to discover new actagardine homologs and gene clusters, and to carry out a bioinformatic comparison with actagardine and its gene cluster.

2 Materials and methods

2.1 Bacterial strains and plasmids

Table 2.1: Plasmids and cosmids used or created in this work.

Plasmid name	Description	Selection marker	Reference
SuperCos1	Cosmid library vector – pUC*ori cos*	Carb Kan	Stratagene
pIJ10702	*attP int* (ΦC31) *oriT* pUC*ori* cos (pMJCOS1)	Carb Apra	(Yanai *et al.*, 2006)
pSET152	*attP int* (ΦC31) *oriT* pUC*ori*	Apra	(Bierman *et al.*, 1992)
pIJ10257	Expression vector - *oriT* Φ*BT1 int-attB ermEp**	Hyg	(Hong *et al.*, 2005)
pUZ8002	*tra* RP4 derivative	Kan	(Paget *et al.*, 1999)
pIJ773	pBluescript SK+ containing cassette P1-FRT-*oriT*-*aac(3)IV*-FRT-P2	Apra	(Gust *et al.*, 2004)
pIJ790	λ-RED (*gam bet*, *exo*) *araC rep101*^{ts}	Chlor	(Gust *et al.*, 2004)
pIJ782	pBluscript SK+ containing cassette P1-FRT-*tet-oriT*-FRT-P2	Tet	(Gust *et al.*, 2004)
pGem	Cloning vector	Carb	Promega
CosAG14HEapra	pIJ10702 CosAG14	Carb Apra	Novacta Biosystems
pIJ12201	SuperCos1 AG14	Carb Kan	This work
pIJ12202	SuperCos1 AL02	Carb Kan	This work
pIJ12203	SuperCos1 AL03	Carb Kan	This work
pIJ12204	SuperCos1 AL13	Carb Kan	This work
pIJ12205	SuperCos1 AL17	Carb Kan	This work
pIJ12206	pIJ10702 AG14	Carb Apra	This work
pIJ12207	pIJ10702 AL02	Carb Apra	This work
pIJ12208	pIJ10702 AL03	Carb Apra	This work
pIJ12209	pIJ10702 AL13	Carb Apra	This work
pIJ12210	pIJ10702 AL17	Carb Apra	This work
pIJ12211	pIJ12206Δ*garO*::tet	Carb Kan Tet	This work
pIJ12377	pIJ12201Δ*garA*::apra	Carb Kan Apra	This work
pIJ12378	pIJ12201Δ*garM*::apra	Carb Kan Apra	This work
pIJ12379	pIJ12201Δ*garO*::apra	Carb Kan Apra	This work
pIJ12380	pIJ12201Δ*garR1*::apra	Carb Kan Apra	This work
pIJ12381	pIJ12201Δ*garT*::apra	Carb Kan Apra	This work
pIJ12382	pIJ12201Δ*garH*::apra	Carb Kan Apra	This work
pIJ12383	pIJ12201Δ*garK*::apra	Carb Kan Apra	This work
pIJ12384	pIJ12201Δ*garR*::apra	Carb Kan Apra	This work
pIJ12385	pIJ12201Δ*garPBP*::apra	Carb Kan Apra	This work

Plasmid name	Description	Selection marker	Reference
pIJ12386	pIJ12202Δ*ligQ*::apra	Carb Kan Apra	This work
pIJ12387	pIJ12201Δ*garA*::scar	Carb Kan	This work
pIJ12388	pIJ12201Δ*garM*::scar	Carb Kan	This work
pIJ12389	pIJ12201Δ*garO*::scar	Carb Kan	This work
pIJ12390	pIJ12201Δ*garR1*::scar	Carb Kan	This work
pIJ12391	pIJ12201Δ*garT*::scar	Carb Kan	This work
pIJ12392	pIJ12201Δ*garH*::scar	Carb Kan	This work
pIJ12393	pIJ12201Δ*garK*::scar	Carb Kan	This work
pIJ12394	pIJ12201Δ*garR*::scar	Carb Kan	This work
pIJ12395	pIJ12201Δ*garPBP*::scar	Carb Kan	This work
pIJ12396	pIJ12202Δ*ligQ*::scar	Carb Kan	This work
pIJ12397	pIJ12206Δ*garA*::scar	Carb Apra	This work
pIJ12398	pIJ12206Δ*garM*::scar	Carb Apra	This work
pIJ12399	pIJ12206Δ*garO*::scar	Carb Apra	This work
pIJ12500	pIJ12206Δ*garR1*::scar	Carb Apra	This work
pIJ12501	pIJ12206Δ*garT*::scar	Carb Apra	This work
pIJ12502	pIJ12206Δ*garH*::scar	Carb Apra	This work
pIJ12503	pIJ12206Δ*garK*::scar	Carb Apra	This work
pIJ12504	pIJ12206Δ*garR*::scar	Carb Apra	This work
pIJ12505	pIJ12206Δ*garPBP*::scar	Carb Apra	This work
pIJ12506	pIJ12207Δ*ligQ*::scar	Carb Apra	This work
pIJ12507	pIJ12201ΔAG14.2-AG14.14::tet	Tet Carb Kan	This work
pIJ12508	pIJ12201ΔAG14.2-AG14.18::tet	Tet Carb Kan	This work
pIJ12509	pIJ12201ΔAG14.28-AG14.30::tet	Tet Carb Kan	This work
pIJ12510	pIJ12206ΔAG14.2-AG14.14::tet	Tet Carb Apra	This work
pIJ12511	pIJ12206ΔAG14.2-AG14.18::tet	Tet Carb Apra	This work
pIJ12512	pIJ12206ΔAG14.28-AG14.30::tet	Tet Carb Apra	This work
pIJ12212	pIJ10257 *ermE*p* *garA*	Hyg	This work
pIJ12213	pIJ10257 e*rmE*p* *garM*	Hyg	This work
pIJ12214	pIJ10257 *ermE*p* *garO*	Hyg	This work
pIJ12215	pIJ10257 *ermE*p* *garR1*	Hyg	This work
pIJ12216	pIJ10257 *ermE*p* *garK*	Hyg	This work
pIJ12217	pIJ10257 *ermE*p* *garR*	Hyg	This work
pIJ12218	pIJ10257 *ermE*p* *ligQ*	Hyg	This work

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

Strain	Genotype	Antibiotic resistance	Reference
DH5α	recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac	None	(Sambrook & Russell, 2001)
ET12567	*dam13::*Tn*9 dcm6 hsdM hsdR recF143 zjj201::*Tn*10 galK galT22 ara14 lacY1 xyl5 leuB6 thi*l *tonA*31 *rpL*136 *his*G4 *tsx*78 *mtl*i *glnV*44 F^-	Chlor, Tet	(MacNeil *et al.*, 1992)
BW25113	Δ(*araD-araB*)*567* Δ*lacZ4787*(::*rrnB-4*) *lacIp-4000*(lacI^Q) λ^- *rpoS369*(Am) *rph-1* Δ(*rhaD-rhaB*)568 *hsdR514*	None	(Datsenko & Wanner, 2000)
DH5α/BT340	As DH5α with pCP20 (*FLP+, λ* ci857+, λ P_R Rep^{ts}, Ap^R, Cm^R)	Chlor	(Datsenko & Wanner, 2000, Cherepanov & Wackernagel, 1995)
S17-1	*recA*1 *pro thi hsdR* RP4-2-Tc::Mu-Km::Tn7	Strep, Spec	(Simon *et al.*, 1983)

Table 2.3: Actinomycetes used in this work.

Species	Strain	Genotype	Reference
Streptomyces lividans	TK24	Str-6 SLP2$^-$ SLP3$^-$	(Kieser, 2000)
Streptomyces lividans	1326	SLP2 SLP3	(Hopwood *et al.*, 1983)
Streptomyces coelicolor	M1146	M145 Δ*act* Δ*red* Δ*cpk* Δ*cda*	(Gomez-Escribano & Bibb, 2010)
Micrococcus luteus	ATCC4698		Novacta Biosystems

Table 2.4: *Streptomyces lividans* TK24 derivatives created in this work.

Strain	Genotype	Reference
M1201	pIJ12206	This work
M1202	pIJ12207	This work
M1203	pIJ12208	This work
M1204	pIJ12209	This work
M1205	pIJ12210	This work
M1206	pIJ12397	This work
M1207	pIJ12398	This work
M1208	pIJ12399	This work
M1209	pIJ12500	This work

Strain	Genotype	Reference
M1210	pIJ12501	This work
M1211	pIJ12502	This work
M1212	pIJ12503	This work
M1213	pIJ12504	This work
M1214	pIJ12505	This work
M1215	pIJ12506	This work
M1216	pIJ10702 (empty vector control)	This work
M1278	pIJ12211	This work
M1279	pIJ12397 in ΦC31 *attB* (Δ*garA*) and pIJ12212 in ΦBT1 *attB*	This work
M1280	pIJ12398 in ΦC31 *attB* (Δ*garM*) and pIJ12213 in ΦBT1 *attB*	This work
M1281	pIJ12399 in ΦC31 *attB* (Δ*garO*) and pIJ12214 in ΦBT1 *attB*	This work
M1282	pIJ12500 in ΦC31 *attB* (Δ*garR1*) and pIJ12215 in ΦBT1 *attB*	This work
M1283	pIJ12503 in ΦC31 *attB* (Δ*garK*) and pIJ12216 in ΦBT1 *attB*	This work
M1284	pIJ12504 in ΦC31 *attB* (Δ*garR*) and pIJ12217 in ΦBT1 *attB*	This work
M1285	pIJ12506 in ΦC31 *attB* (Δ*liqQ*) and pIJ12218 in ΦBT1 *attB*	This work

Table 2.5: Derivatives of *Streptomyces coelicolor* M1146 and *Streptomyces lividans* 1326 created in this work.

Strain	Genotype	Reference
M1217	M1146 pIJ12206	This work
M1218	M1146 pIJ12397 (Δ*garA*::scar)	This work
M1219	M1146 pIJ10702 (empty vector control)	This work
M1220	1326 pIJ12206	This work
M1276	1326 pIJ12397 (Δ*garA*::scar)	This work
M1277	1326 pIJ10702 (empty vector control)	This work

Table 2.6: *Actinoplanes* strains used in this work.

Species	Strain	Genotype	Reference
Actinoplanes garbadinensis	DSM 4432		(Parenti *et al.*, 1976)
Actinoplanes liguriae	NCIMB41362		(Boakes *et al.*, 2010)

2.2 Culture media and antibiotics

Table 2.7: Concentration of antibiotics used in this work

Antibiotic	Concentration (µg/ml)
Apramycin (Apra)	40
Carbenicillin (Carb)	100
Chloramphenicol (Chlor)	25
Hygromycin (Hyg)	40
Kanamycin (Kan)	50
Nalidixic acid (Nal)	20
Spectinomycin (Spec)	50 (*E. coli*) 200 (*Streptomyces* sp.)
Streptomycin (Strep)	50 (*E. coli*) 10 (*Streptomyces* sp.)
Tetracycline (Tet)	10

Unless stated otherwise, the media used for the culturing actinomycetes and *E. coli* were prepared as previously described (Kieser, 2000).

Table 2.8: Agar media used in this work.

Medium	Composition		Preparation instructions
L-Agar	Agar Difco Bacto tryptone NaCl Glucose Distilled water	10 g 10 g 5 g 1 g Up to 1000 ml	The ingredients, except agar, were dissolved in the distilled water and 200 ml aliquots were dispensed into 250 ml Erlenmeyer flasks containing 2 g agar. The flasks were closed and autoclaved.
Difco nutrient agar (DNA)	Difco Nutrient Agar Distilled water	4.6 g 200 ml	Difco Nutrient Agar was placed in each 250 ml Erlenmeyer flask and distilled water was added. The flasks were closed and autoclaved.
Mannitol soya flour medium (MS)(SFM)	Agar Mannitol Soya Flour Tap water	20 g 20 g 20 g Up to 1000 ml	The mannitol was dissolved in the water and 200 ml aliquots poured into 250 ml Erlenmeyer flasks each containing 2 g agar and 2 g soya flour. The flasks were closed and autoclaved twice, with gentle shaking between the two runs.

Medium	Composition		Preparation instructions
Oatbran Medium (OB-N)	Porridge Oats Lab M Agar Tap Water	40 g 20 g To 1000 ml	The ingredients were mixed and dispensed into 250 ml Erlenmeyer flasks. The flasks were closed and autoclaved twice, with gentle shaking between the two runs.
ABB13	Soytone Peptone Soluble starch $CaCO_3$ MOPS Agar Distilled water Adjust pH to 7.0	5 g 5 g 3 g 2.1 g 20 g To 1000 ml	The ingredients were mixed and dispensed into 250 ml Erlenmeyer flasks. The flasks were closed and autoclaved twice, with gentle shaking between the two runs.

2.3 Liquid Media

Unless stated otherwise, the media used for the culturing of actinomycetes and *E. coli* were prepared as previously described (Kieser, 2000, Sambrook & Russell, 2001).

Table 2.9: Liquid media used in this work.

Medium	Composition		Preparation Instructions
L (Lennox) broth (LB)	Difco Bacto tryptone Difco yeast extract NaCl Glucose Distilled water	10 g 5 g 5 g 1 g Up to 1000 ml	The ingredients were dissolved, in the distilled water and 10 ml aliquots were dispensed into universals and autoclaved.
2 X YT medium	Difco Bacto tryptone Difco yeast extract NaCl Distilled water	16 g 10 g 5 g Up to 1000 ml	The ingredients were dissolved, in the distilled water and 10 ml aliquots were dispensed into universals and autoclaved.
SOB medium	Tryptone Yeast extrace NaCl Distilled water	20 g 5 g 0.5 g Up to 950 ml	After dissolving the solutes in water, 10 ml 250 mM KCl was added and the pH was adjusted to pH 7 with 5 N NaOH. The volume was then made up to 1000 ml with deionised water and autoclaved.

Medium	Composition		Preparation Instructions
SOC medium			SOC medium is identical to SOB medium except that after autoclaving, 20 ml of sterile 1 M solution of glucose and 5 ml of sterile 2 M $MgCl_2$ were added.
AAS seed medium	Soluble starch Glucose Peptone Dry corn steep liquor Yeast extract Distilled water Adjust pH to 6.0	10 g 10 g 5 g 1 g 2 g Up to 1000 ml	The ingredients were dissolved in the distilled water and autoclaved.
GM1	Lab Lemco meat extract Peptone NaCl Yeast extract Soy flour Glucose $CaCO^3$ Distilled water Adjust pH to 7.6	4 g 4 g 2.5 g 1 g 10 g 25 g 5 g Up to 1000 ml	The ingredients were dissolved in the distilled water and autoclaved.
TSB	Tryptic soy broth Distilled water Adjust pH to 7.0	30 g Up to 1000 ml	The ingredients were dissolved in the distilled water and autoclaved.

2.4 Solutions and buffers

Table 2.10: Solutions and buffers used in this study.

Solution/buffer	Compostion and instructions for preparation	
Cosmid isolation solution I	Glucose Tris-HCl (pH 8) EDTA	50 mM 25 mM 10 mM
Cosmid isolation solution II	NaOH SDS	0.2 M 1 %
Cosmid isolation solution III	Sodium acetate (pH 4.8) Acetic acid	3 M
DNA loading buffer	Bromophenol blue Xylene-cyanol blue Sucrose in water	0.25 % (w/v) 0.25 % (w/v) 40 % (w/v)

Solution/buffer	Compostion and instructions for preparation	
Lysozyme solution (genomic DNA extractions)	Sucrose Tris-HCl (pH 8) EDTA (pH 8) Lysozyme (Sigma)	0.3 M 25 mM 25 mM 2 mg/ml
Phenol-chloroform-isoamylalcohol	Phenol-chloroform-isoamylalcohol solutions were prepared by mixing the constituents 25:24:1.	
SDS stock solution	10% (w/v) in H_2O, filtered through a 0.45μm membrane.	
TAE buffer	Tris Acetic acid EDTA	40 mM 1.142 % 1 mM
TES buffer	Tris-HCl (pH 8) EDTA NaCl	10 mM 1 mM 1 M

2.5 Growth conditions

2.5.1 Growth and storage of *E. coli*.

E. coli was grown overnight on solid L-agar, or in liquid LB broth with shaking at 37°C (30°C for temperature sensitive BW25113/pIJ10790). For selection of plasmid- or cosmid-containing cells the appropriate antibiotics were added at the appropriate concentration. For long-term storage, overnight cultures of LB broth cultures were mixed 40 % glycerol and stored at –80°C.

2.5.2 Growth and storage of *Streptomyces.*

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

For liquid culturing of *S. lividans* and *S. coelicolor*, ~10^8 spores were heat shocked at 50°C for 10 min in 5 ml TES buffer. An equal volume of 2 x YT medium was added and this was 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 x *g* on a benchtop centrifuge and then resuspended in 500 µl

of water. The cells were vortexed to disperse the clumps. 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 x *g* in a benchtop centrifuge.

For manipulation and generation of spore stocks, *S. lividans* and *S. coelicolor* were plated onto SFM (with appropriate antibiotics) and incubated at 30°C for up to 7 days. Spores were collected in approximately 2 ml sterile 20% glycerol, using a sterile cotton pad through which the spores were filtered using a sterile 2 ml syringe and transferred to a cryotube. Spore stocks were stored at -20°C. Spore stocks were titred by making serial dilutions in water and plating on SFM. Resulting colonies were counted from at least three dilutions and averaged.

2.5.3 Growth and storage of *Actinoplanes*

Actinoplanes strains were grown and manipulated as described in Boakes *et al.* (2009) and Boakes *et al*. (2010). *A. garbadinensis* and *A. liguriae* were grown on OB-N or ABB13 agar for 7 to 10 days at 30°C. For liquid growth, agar plugs taken from an *Actinoplanes* agar culture were typically placed in 250 ml baffled conical flasks containing 40 ml of TSB and incubated with shaking at 30°C for 6 to 7 days. For long-term storage, culture broth containing mycelium was dispensed into 2 ml cryotubes and stored at -20°C or -80°C.

2.5.4 Growth and storage of *M. luteus*

M. luteus ATCC4698 was typically grown in L broth with shaking at 250rpm or on L agar at 30 or 37°C. For long-term storage, strains were grown over-night in L medium and cells resuspended in 20% glycerol and stored at -20°C or -80°C.

2.6 Genetic manipulations

2.6.1 Plasmid Isolation

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.

2.6.2 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. 200 µl solution II was added and the tubes inverted ten times. A volume of 150 µl solution III was then added and mixed in by inverting the tube five times. The tube was then centrifuged at 13000 rpm in a microcentrifuge for 5 min at room temperature. The supernatant was mixed with 400 µl phenol/chloroform, vortexed briefly to mix and then centrifuged at 13000 rpm in a microcentrifuge for 5 min. The upper phase was then 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 13000 rpm in a microcentrifuge for 5 min. The pellet was washed with 1 ml 70% ethanol and centrifuged at 13000 rpm in a microcentrifuge. The pellet was dried by leaving the tube open for 5 min at room temperature prior to resuspension in 30 µl elution buffer.

2.6.3 DNA extraction from *Streptomyces*

2 µl of *S. coelicolor* spore stock was used to inoculate 10 ml SOC and grown overnight with shaking at 30°C. The mycelium was recovered by centrifugation at 3000 x *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/isoamyl alcohol was added and vortexed briefly until completely mixed. The sample was centrifuged in a microcentrifuge on 16000 x *g* for 8 min. Following centrifugation, the aqueous phase was removed to a fresh Eppendorf tube. To decrease viscosity, 200 µl water was added to the supernatant. The supernatant was extracted twice with phenol/chloroform/isoamyl alcohol. 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 x *g* for 5 min. The DNA pellet was then washed with 70% ethanol before dissolving in 500 µl elution buffer at room temperature for 2 h. DNAse free RNAse was then added to a final concentration of 40 µg/ml and the sample was incubated at 37°C for 30 min. The DNA sample then underwent two phenol/chloroform/isoamyl alcohol extractions. The DNA was precipitated and washed as before and was then dissolved in 100 µl elution buffer. The DNA concentration was determined by spectrophotometry using the ND-1000 spectrophotometer (NanoDrop®).

2.6.4 Digestion of DNA using 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 typically 20 µl for analytical digests and 200 µl for preparative digests for use in targeting cassette construction. Unless otherwise instructed, digests were typically carried out for 1 h at 37°C.

2.6.5 Agarose gel electrophoresis

1% agarose gels were prepared with TAE buffer with 0.5 µg/ml ethidium bromide. Agarose gel electrophoresis was carried out in 1% TAE buffer at 100 V until completion. Hyperladder I (Bioline) or 1 kb ladder (NEB) were used to provide size markers.

2.6.6 Extraction of DNA 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 ultrapure water or elution buffer.

2.6.7 Preparation and transformation of electro-competent cells

A single *E. coli* colony was used to inoculate 10 ml LB broth and grown overnight with shaking at 37°C. 100 µl of this preculture was inoculated into 10 ml SOB and grown at 37°C (30°C for BW25113/pIJ790) for 3-4 h with shaking at 200 rpm to an OD_{600} of ~ 0.6. The cells were recovered by centrifugation at 3000 x 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 was 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 (30°C for BW25113/pIJ790 transformed with cosmid). Transformants were selected by spreading onto LB agar containing the appropriate antibiotic (or on DNA agar for selection with hygromycin).

2.6.8 Transformation of chemically competent cells

For transformation, 50 µl of the suspension of competent cells 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 exactly 90 s. 950 µl of LB (prewarmed to 37°C) was added to the cells and they 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 antibiotics.

2.6.9 FLP-mediated recombination to generate scar mutants

E. coli DH5α/BT340 was grown in 10 ml L broth containing 25 µg/ml chloramphenicol overnight at 30°C with shaking at 250 rpm. 10ml L broth containing chloramphenicol were inoculated 1 in 20 with this overnight culture and incubated with shaking at 30°C for 3-4h until OD_{600} ~0.6. Electrocompetent cells were generated from the culture as described in 2.6.7. 50 µl of the cell suspension were mixed with ~100 ng cosmid containing the FRT-flanked cassette to be removed. Electroporation was carried out as 2.6.7. Typically 150 µl aliquots of the transformation mix were plated onto L agar containing 25 µg/ml chloramphenicol and apramycin and incubated for 2 days at 30°C. Approximately eight single

colonies were picked and streaked on to L agar without antibiotics and incubated at 42°C overnight to promote FLP-recombination. Single colonies from these plates were picked and streaked first on to L agar containing apramycin and then on to L agar containing antibiotics for selection of the backbone of the cosmid (carbenicillin and kanamycin for Supercos or carbenicillin and apramycin for pIJ10702). The plates were incubated at 37°C overnight and compared to identify clones sensitive to the selection for the cassette used for mutagenesis but resistant to the selection for the cosmid backbone. Four to eight such clones were selected and the removal of the cassette confirmed by PCR using flanking primers and the PCR products checked by Sanger sequencing.

2.6.10 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. The purifications were carried out according to the manufacturer's instructions. The DNA was eluted in elution buffer. A ligation reaction volume of 10 µl was carried out with an insert: vector molar ratio of 3:1. Ligations were carried out overnight at 16°C using high concentration T_4 DNA ligase (Promega).

2.6.11 Conjugation of DNA into *Streptomyces*

E. coli ET12567/pUZ8002 or S17-1 containing an *oriT*-containing vector was inoculated into 10 ml LB containing chloramphenicol (25 µg/ml), kanamycin (50 µg/ml) and the appropriate antibiotic for selection of the *oriT*-containing vector and grown overnight at 37°C. 100 µl overnight culture was 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 (10^8) *Streptomyces* spores were added to 500 µl 2 × 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} in a total of 200 µl of water was generated. 200 µl of each dilution was plated out on MS agar + 10 mM $MgCl_2$ (without antibiotics) and incubated at 30°C for 16-20 h. The plate was overlaid with 1 ml of water containing 0.5 mg nalidixic acid and the appropriate plasmid selection for the isolation of exconjugants. Incubation was continued at 30°C.

2.7 Polymerase chain reaction (PCR) and DNA sequencing methods

2.7.1 General PCR

PCR from *Streptomyces* genomic DNA was typically carried out using 10 ng DNA as template (genomic DNA prepared according to section 2.6.3). PCR from *E. coli* plasmid DNA was typically carried out using 1 µl of a standard plasmid or cosmid preparation (section 2.6.1 and 2.6.2) as template. Primers were designed to have an annealing temperature between 55 and 60°C.

Reaction mixture:

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

Cycle conditions (Touchgene Gradient, Techne):

Initial denaturation:	96°C, 3 min	
Denaturation	96°C, 30 s	30 cycles
Primer annealing	56°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 minute was assumed for the DNA polymerase.

2.7.2 *E. coli* colony PCR

Colony PCR allows for the rapid screening of recombinant plasmids. The PCR reaction mix was as in 2.7.1. Single colonies of *E. coli* were picked from agar plates using a sterile toothpick and added to 50 µl of DMSO in a 1.5 ml Eppendorf tube. The tube was centrifuged briefly to pellet cell debris and 2.5 µl of the DMSO was added to the appropriate PCR mix in place of the DMSO and DNA template. The initial denaturation time was increased by 2 minutes.

2.7.3 *Streptomyces* colony PCR

Streptomyces single colonies were patched on to DNA media containing 1% glucose. Plates were incubated at 30°C overnight (20-24h). *Streptomyces* mycelium was scraped from the plates using a sterile toothpick and added to 50 ul 100% DMSO in a 1.5 ml Eppendorf tube. The tube was shaken vigorously for 30-45 mins and centrifuged briefly to pellet cell debris; 2.5 µl of the DMSO was added to the appropriate PCR mix in place of the DMSO and DNA template. The initial denaturation time was increased by 2 minutes.

2.7.4 Purification of PCR products

The QIAquick PCR purification kit (Qiagen) was used to remove unincorporated primers, dNTPs and enzymes from completed PCR reactions. One tenth of the PCR reaction mixture was submitted to agarose gel electrophoresis and stained for visualisation. 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 free of primers, dNTPs and enzymes and the DNA fragment was eluted in elution buffer.

2.7.5 DNA sequencing

DNA sequencing was used to confirm the presence of inserts in recombinant plasmids and cosmids and to confirm the correct sequence of PCR amplified DNA fragments. DNA sequencing was carried out using ABI 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 Genome Analysis Centre at the John Innes Centre. Sequence chromatograms were analysed using the chromatogram trace viewer software FinchTV (version 1.4; http://www.softpedia.com).

2.8 PCR targeting of cosmids

PCR targeting was used to generate mutagenised cosmids in *E. coli* that could be subsequently transferred to *Streptomyces* via conjugation. In general, the strategy is carried out with minor modifications to that described in REDIRECT technology: PCR-targeting system in *Streptomyces coelicolor* by Gust *et al.* (2004).

2.8.1 PCR amplification of disruption cassettes

The pIJ782 (tet) and pIJ773 (apra) templates were used to generate targeting cassettes for the creation of mutant cosmids. 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 2.11.

pIJ10257 was used for the constitutive expression of genes using the *ermE*p* promoter, where the gene of interest was amplified with appropriate primers (table 2.13) that include restriction sites. After cloning of the PCR product into pGem (following manufacturer's instructions), PCR amplified fragments were confirmed by PCR (sections 2.7.4 and 2.7.5), cut from the vector with the appropriate enzymes and ligated with pre-digested pIJ10257 (section 2.6.10) and transferred into *Streptomyces* by conjugation (2.6.11).

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

Reaction mixture:

Forward primer (50 pmoles/µl)	1 µl	50 pmoles
Reverse primer (50 pmoles/µl)	1 µ	50 pmoles
Template DNA	50 ng	
Buffer (10x)	5 µl	1x
dNTPs (40 nmoles/µl)	1 µl	50 µM each
Dimethylsulfoxide (DMSO)	2.5 µl	5%
DNA polymerase (2.5 U/µl)	1µl	2.5µl
Water to a volume of	50 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 2.6.6). To remove enzymes, unincorporated primers and dNTPs, the remaining 45 µl of the PCR product was purified using the Qiagen PCR purification kit according to the manufacturer's instructions (section 2.7.4). The PCR product was eluted from the column with 20 µl of water to give a DNA concentration of approximately 100 ng/µl.

Table 2.11: Primers used for the amplification of gene-specific disruption cassettes in this study.

Primer name	Primer sequence (5' → 3')
garA APR LFP SpeI forward	ggctattcgcccgggaagtccaccgaaaggaagacacaccactagtattccggggatccgtcgacc
garA APR LFP AvrII reverse	gggcgggcgatgcccgccccgggccggaaacgatcgtcgacctaggtgtaggctggagctgcttc
garM APR LFP SpeI forward	cgcccgccccgggccatccatttcgcgcggggagcagccactagtattccggggatccgtcgacc
garM APR LFP AvrII reverse	agcacgctgagcacgatctcttccggcatgggctcctcccctaggtgtaggctggagctgcttc
garO APR LFP SpeI forward	tggcctggctgcagccgccgctgacctgaggaggagcccactagtattccggggatccgtcgacc
garO APR LFP AvrII reverse	gcgcgggacctacgcggagctaccgcggacctgaggccgcctaggtgtaggctggagctgcttc
garR1 APR LFP SpeI forward	tccgcgtaggtcccgcgcaatgccgcggacctgacggccgactagtattccggggatccgtcgacc
garR1 APR LFP AvrII reverse	atataagtgaagtaatgtcacgaatgcggatgcaagttgccctaggtgtaggctggagctgcttc
garH APR LFP SpeI forward	atttttcgcgtatagtcattaacaacacacagcgcgtcgactagtattccggggatccgtcgacc
garH APR LFP AvrII reverse	tcagatcctcgcggcgatggcgtcggtggcggtgtgacgcctaggtgtaggctggagctgcttc
garT APR LFP SpeI forward	aggcgcgcagctgggaccaggcaagcgcgaagatcatcgactagtattccggggatccgtcgacc
garT APR LFP AvrII reverse	agcgtgggtgcggtgacaacgcagacgacacacgccgagcctaggtgtaggctggagctgcttc
garK APR LFP SpeI forward	atgtctgcttaggctgggaccgtggtgcgagggaaccggactagtattccggggatccgtcgacc
garK APR LFP AvrII reverse	atcgtcggcgatcagcacggcgatcggtggccgcgtcatcctaggtgtaggctggagctgcttc
garR APR LFP SpeI forward	accacctggcccctcaccccggaaggagcgcgcgcatgaactagtattccggggatccgtcgacc
garR APR LFP AvrII reverse	tgccaccgtggcaccgtgtcaagggcctgagggtgcggacctaggtgtaggctggagctgcttc
garPBP APR LFP SpeI forward	gtttcgcctcacactggtggtggtgcggtgagttcctggactagtattccggggatccgtcgacc
garPBP APR LFP AvrII reverse	cacgaggcgaacgccggcgacgtctacgccgtcagccggcctaggtgtaggctggagctgcttc
ligQ APR LFP SpeI forward	gtcacctggctggaaccgccgcggggcacgcatgtctgaactagtattccggggatccgtcgacc
ligQ APR LFP AvrII reverse	cctggtccaggacactgagcatggcgtacacctcgggatcctaggtgtaggctggagctgcttc
garO tet LFP SpeI forward	tggcctggctgcagccgccgctgacctgaggaggagcccactagtaataggcgtatcacgaggcc
garO tet LFP AvrII reverse	gcgcgggacctacgcggagctaccgcggacctgaggccgcctaggcgccggcttc
CosAG14.2 tet LFP SpeI forward	ccgtcgggaccgaccgtgacgacgatccgaggattctccgactagtaataggcgtatcacgaggcc
CosAG14.14 tet LFP AvrII reverse	gcgagccagtacagcggagtcggccgcaagcgctgaggcgcctaggcgccggcttc
CosAG14.18 tet LFP AvrII reverse	tgcttcgcccacgggaacaagtgaaaatgggagaacaacgcctaggcgccggcttc
CosAG14.28 tet LFP SpeI forward	gtggtggtgatccaggtgagcaagggccgccggggcgcgtactagtaataggcgtatcacgaggcc
CosAG14.30 tet LFP AvrII reverse	tcgaggatgtccttcatggagacgacgggcatgccgccagcctaggcgccggcttc

2.8.2 Introduction of cosmids into *E. coli* BW25113/pIJ790 by electroporation

E. coli BW25113 is a λ-RED recombination-proficient host strain and contains pIJ790, carrying the antibiotic resistance marker *cat* (chloramphenicol resistance) and a temperature sensitive origin of replication (replicates at 30°C, not at 37°C). BW25113/pIJ790 was grown overnight at 30°C in 10 ml LB containing chloramphenicol (25 μg/ml). 100 μl of the overnight culture were inoculated into 10 ml SOB containing chloramphenicol (25 μg/ml) and grown for 3-4 h at 30°C with shaking at 200 rpm to an OD600 of ~ 0.6. Electrocompetent cells were generated from this culture and transformed with 100 ng cosmid DNA by electroporation (section 2.6.3). Transformants were selected by spreading onto L-agar containing carbenicillin (100 μg/ml), kanamycin (50 μg/ml) and chloramphenicol (25 μg/ml) and incubating overnight at 30°C.

2.8.3 PCR targeting of a cosmid

10 ml SOB (without $MgSO_4$) containing carbenicillin (100 μg/ml), kanamycin (50 μg/ml) and chloramphenicol (25 μg/ml) were 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 OD600 of ~ 0.6 and electrocompetent cells were prepared. 50 μl cell suspension was then mixed with ~ 100 ng (1-2 μl) of PCR product and electroporated (section 2.6.7). Selection for gene disruption in the cosmid of interest was carried out on LB agar containing carbenicillin (100 μg/ml), kanamycin (50 μg/ml) and apramycin (50 μg/ml)/ tetracycline (10 μg/ml) overnight at 37°C. Correct targeting of cosmids was confirmed by PCR and restriction analysis.

The targeting cassettes were subsequently removed using the FLP recombinase, detailed in section 2.6.9.

Gene disruptions were checked using a combination of PCR (section 2.7.1, 2.7.2 and 2.7.5) and restriction analysis (2.6.4).

Table 2.12: Primers used for the confirmation of gene disruptions.

Primer name	Primer sequence (5' → 3')
garA (con-3) forward	cgaaatcgccgggctattcg
garA (con-3) reverse	ctgtcgcgtaccgaggtg
garM (con-3) forward	gcgagctcaccttcgag
garM (con-3) reverse	cggtggtagccgtactg
garO (con-3) forward	gcatctgccacggcgac
garO (con-3) reverse	gggtctgtccagcgacgac
garR1 (con-3) forward	gtatccgaggtggtgctcac
garR1 (con-3) reverse	gtctgacccgctgacgac
garH (con-3) forward	gatgctgaccagtgcttcg
garH (con-3) reverse	gtgaggaactgcgcagtc
garT (con-3) forward	ccggtcgccagcatcac
garT (con-3) reverse	cacggtcccagcctaagcag
garK (con-3) forward	cgccgtgttcccgcaatcg
garK (con-3) reverse	gaccatcgcgaaaccggtgc
garR (con-3) forward	cactgttcaacggcgactgc
garR (con-3) reverse	cgcaccaccaccagtgtg
garPBP (con-3) forward	ctcaggcccttgacacg
garPBP (con-3) reverse	cgttgtcgttcgagttcacg
LigQ (con-3) forward	gggacatcggcgagac
LigQ (Con-3) reverse	caccttgagcgggctg
TcR Rev (con-1)	ccaagcctatgcctacagc
TcR Fwd (con-2)	gctcaagccttcgtcactg
10257 seq F	acgtccatgcgagtgtcc
10257 seq R	ccaaacggcattgagcgtc

Table 2.13: Primers used for the amplification of complementation constructs.

Primer name	Primer sequence (5' → 3')
garA NdeI *ermE* fwd	aaaaacatatgtctgctctcgccatcgag
garA HindIII *ermE* rev	aaaaaaagctttcagcaggcgcagatc
garM NdeI *ermE* fwd	aaaaacatatgtcaccggttccttc
garM HindIII *ermE* rev	aaaaaaagctttcaggtcagcggcg
garO NdeI *ermE* fwd	aaaaacatatgccggaagagatcgtgctc
garR1 NdeI *ermE* fwd	aaaaacatatgcgttcggcggctcg
garR1 HindIII *ermE* rev	aaaaaaagctttcaggacaggcagacgt
garK NdeI *ermE* fwd	aaaaacatatgctcaggatcgacgccc
garK HindIII *ermE* rev	aaaaaaagctttcatgcgcgcgctc
garR NdeI *ermE* fwd	aaaaacatatgacgcggccaccg
garR HindIII *ermE* rev	aaaaaaagctttcagccgacagcgc
garPBP NdeI *ermE* fwd	aaaaacatatgcgccggcgcg
garPBP HindIII *ermE* rev	aaaaaaagctttcagcggggtggc
ligQ NdeI *ermE* fwd	aaaaacatatgtctgagacggccggg
LigQ HindIII *ermE* rev	aaaaaaagctttcaggtcagcagttgacgg

2.8.4 Construction of integrative cosmids

The FLP mediated excision of the targeting cassette required the re-introduction of *oriT* to allow for the transfer of the cosmid into *Streptomyces*. The targeted cosmids were transferred back into *E. coli* BW25113/pIJ790.

E. coli BW25113/pIJ790 containing the SuperCos1 cosmid (containing the FLP mediated gene deletion) to be targeted was grown in 10ml L broth containing 50 µg/ml carbenicillin, 50 µg/ml kanamycin and 25 µg/ml chloramphenicol overnight at 30°C with shaking at 250 rpm. 10 ml SOB (minus Mg) containing carbenicillin, kanamycin, chloramphenicol and 10 mM L-arabinose was inoculated 1 in 20 with this overnight culture which was incubated with shaking at 30°C for 90 min. An additional 100 µl 1 M L-arabinose were added and incubation continued at 30°C for an additional 60 min (until OD_{600} ~0.6). Electrocompetent cells were generated from the induced culture as described in 2.6.7. 50µl of the cell suspension were mixed with ~100 ng gel purified 5247 bp *SspI* fragment derived from pIJ10702 (pMJCOS1). Electroporation was carried out as 2.6.7. Typically 300 µl aliquots of the transformation mix were plated onto two plates of L agar containing 50 µg/ml apramycin and 100 µg/ml carbenicillin and incubated overnight at 37°C.

2.8.5 Transfer of mutant cosmids into *Streptomyces*

Since *S. coelicolor* carries a methyl-specific restriction system it was necessary to passage the targeted cosmids containing an *oriT* through a non-methylating *E. coli* host. Constructs for all *Streptomyces* were transferred into *E. coli* ET12567 pUZ8002 so that they could be conjugated into a number of hosts. To achieve this, the cosmids were introduced by transformation into the non-methylating *E. coli* ET12567 containing the RP4 derivative pUZ8002. The cosmid was then transferred to *Streptomyces* by conjugation (section 2.6.11). Exconjugants were selected using the appropriate antibiotic.

2.8.6 Confirmation of PCR targeted mutants in *Streptomyces*

Confirmation was typically carried out using colony PCR (section 2.7.3) or PCR using genomic DNA (section 2.6.3 and 2.7.1).

2.9 Bioassay methods

2.9.1 Solid bioassay

Streptomyces and *Actinoplanes* strains were assessed for antimicrobial activity using *M. luteus* as test organism. 2.5 ml of molten agar were poured into each of the 25 sections of divided square Petri plates and allowed to set. Either 5 µl of a glycerol stock or 5 µl of water containing material scraped from a plate was used to inoculate each agar cube, and the plate incubated for 7 days at 30°C. *M. luteus* was grown as described in section 2.5.4. The overnight culture was diluted 1 in 25 into 50 ml L broth and grown at 30°C with shaking until an OD_{600} of 0.4-0.6 was reached. This culture was diluted 1 in 10 into molten L agar at 50°C and approximately 20 ml poured into 90 mm standard Petri dishes. The agar plugs with the strains to be assessed were subsequently added to the top of the plate and incubated at 30°C until a halo could be seen (typically overnight).

2.9.2 Liquid bioassays

For liquid bioassays, agar plugs taken from *Actinoplanes* agar cultures were typically placed in 50 ml conical flasks containing 8 ml of AAS seed medium with appropriate antibiotics (if required) and three or four glass beads, and incubated with shaking at 30°C for 3-4 days. 250 µl was taken and used to inoculate a 50 ml flask containing 8 ml of GM1 and incubated at 30°C for a further 3-4 days. *Streptomyces* cultures were similarly incubated but using TSB medium. 1 ml aliquots were taken and spun at 13000 rpm in a centrifuge for 10 minutes. The supernatant was either analysed by MALDI-ToF analysis or added to L agar plates pre-seeded with *M. luteus*. For the bioassays, *M. luteus* was grown as described in section 2.5.4. The overnight culture was diluted 1 in 25 into 50 ml L broth and

grown at 30°C with shaking until an OD_{600} of 0.4-0.6 was reached. This culture was diluted 1 in 10 into molten L agar at 50°C and approximately 20 ml poured into 90 mm standard Petri dishes. Holes were made in the agar using an unturned sterile P1000 tip. 50 µl of the spun supernatant was added to the hole and allowed to dry before incubation at 30°C until a halo could be seen (typically overnight).

2.10 MALDI-ToF analysis

Culture supernatants were diluted 1 in 5 with 5% formic acid. The diluted sample (ca. 0.8 µl) was spotted onto a PAC plate (Prespotted AnchorChipTM MALDI target plate, Bruker Daltonics, Bremen, Germany) and the spots washed briefly with 8 µl 5% formic acid according to the manufacturer. After drying, the samples were analysed by MALDI-ToF on a Bruker Ultraflex TOF/TOF. The instrument was calibrated using the prespotted standards (ca. 200 laser shots). Samples were analysed using a laser power of approx. 25% and spectra were summed from ca. 20 x 20 laser shots.

2.11 LC/MS analysis

High performance liquid chromatography – mass spectrometry (HPLC-MS or LC/MS) were performed on a Hewlett Packard 1050 series HPLC system linked to a Micromass Platform LC (operated with MassLynx version 3.5 software) with the following parameters:

Column:	Agilent Zorbax SB-C18 150 x 4.6 mm 5µ	
Mobile Phase A:	10% acetonitrile, 0.1% formic acid, in water	
Mobile Phase B:	90% acetonitrile, 0.1% formic acid, in water	
Flow rate:	1 ml/min	
Gradient:		
Time 0 min	100% A	0% B
Time 10min	0% A	100% B
Time 11min	0% A	100% B
Time 11.2min	100% A	0% B
Cycle time	15 min	
Wavelength:	200 – 400 nm	

Injection volume:	10 µl
Post column split:	1:10
Mass spectrometer:	Micromass Platform LC
Mode:	Electrospray positive ion
Cone voltage:	40V
Capillary voltage:	3.10V
Skimmer lens offset:	5V
Ion Energy:	1.4V

3 Computational analysis

3.1 Introduction

Bioinformatic analysis of DNA sequences has become a powerful technique that allows the identification and characterisation of genes and the proteins they encode through their homology to other characterised examples.

Comparison of putative gene products against a database of previously identified and characterised proteins can lead to the identification and classification of unidentified genes of interest based on amino acid sequence homology of the encoded protein (Chatterjee *et al.*, 2005). Using this approach and computational analysis, this study identified a number of lantibiotic biosynthetic genes likely to be required for actagardine and DAB biosynthesis.

The genes required for production of a particular lantibiotic are clustered together in the genome of the producing organism. The clusters contain genes that encode a prepropeptide, at least one modification enzyme and genes for transport and immunity. Most clusters also contain a regulatory gene or genes that control the expression of the biosynthetic genes (Chatterjee *et al.*, 2005, Schmitz *et al.*, 2006). Identification of lantibiotic biosynthesis genes can be accomplished because of the specialised functions that these proteins carry out, and the corresponding conservation of amino acid sequence across functional homologues.

The complement and organisation of genes is often well conserved between biosynthetic gene clusters that encode similar LanA prepropeptides, for example, nisin and subtilin produced by *Lactococcus lactis* and *Bacillus subtilis,* respectively (Chatterjee *et al.*, 2005). Genes found within lantibiotic gene cluster that might perform additional functions can sometimes be identified by their absence from similar lantibiotic gene clusters; for example, *mrsD* that encodes a flavoprotein responsible for the creation of the unusual S-[(Z)-2-aminovinyl]-methyl-D-cysteine in mersacidin (Blaesse *et al.*, 2003), and *cinR1* that encodes the SARP (*Streptomyces* antibiotic regulatory protein) in the cinnamycin gene cluster, the only example of the presence, in a lantibiotic gene cluster, of a member of this family of regulatory

genes (Blaesse *et al.*, 2003, Widdick *et al.*, 2003). Analysing the composition of the gene cluster may also give some indications of its evolutionary origin.

3.2 Computational analysis of actagardine and DAB biosynthesis

Actagardine biosynthesis is currently being studied by Novacta Biosystems. One of the main aims of their research is to produce a new clinically useful lantibiotic. To achieve this aim they attempted to clone the actagardine biosynthetic gene cluster into a heterologous host to both manipulate the lantibiotic that the cluster produced, and to enable high levels of lantibiotic production. To do this, a detailed description of all of the genes involved in actagardine biosynthesis was required. To this end Novacta created cosmid libraries of ~35-40 kb fragments of genomic DNA from *A. garbadinensis* and *A. liguriae*. The libraries were initially probed with DIG-labelled SBDIG-1 probe, designed based on the known amino acid sequence of actagardine, taking into account the codon usage of *Actinoplanes*. This probe was used in colony hybridisation experiments and successfully identified cosmid pLITAG01 derived from *A. garbadinensis* which was subsequently sequenced. The cosmid sequence was analysed and used to design a new probe designated SBDIG-2, ~2.3 kb in length, generated using pLITAG01 as a template. Both probes were then used to screen SuperCos1 genomic libraries of both *A. garbadinensis* and *A. liguriae*, identifying several cosmids that hybridised to both probes. Two cosmids, AG14 from *A. garbadinensis* and AL02 from *A. liguriae,* were subsequently fully sequenced. Heterologous expression of AG14 in *S. lividans* 1326 confirmed that the cosmid contained all of the genes required for the production of actagardine. Novacta Biosystems were unable to confirm that all of the genes required for DAB biosynthesis were present on AL02 since there was no production of DAB when the cosmid was conjugated into the same heterologous host.

An initial bioinformatic analysis carried out by Novacta Biosystems was completed several years ago. With the recent explosion in the amount of publicly available sequence information, a new analysis was performed to see if the original annotation could be improved. While gene annotations describing the potential function of each gene product have remained mostly the same, additional data added to the databases has provided further and more reliable gene homology comparisons and more detailed information about the encoded proteins compared

to the original Novacta study. In recent publications, Novacta has also updated its bioinformatic analyses of the biosynthetic gene clusters for both actagardine and DAB (Boakes *et al.*, 2010, Boakes *et al.*, 2009).

The cosmid sequences for AG14 and AL02 used in this study were kindly provided by Novacta Biosystems and visualised using the Artemis program (version 8) available from the Sanger Institute (Rutherford *et al.*, 2000). Open reading frames (ORFs) were first identified from the raw sequence using the graphic tool "GC frame plot" (Bibb *et al.*, 1984) available within the Artemis program. This algorithm visually displays ORFs based on the position-specific, non-random distribution of nucleotides within codons, and is particularly effective with sequences of high G + C content. Manual inspection of the "frame plot" allowed designation of start and stop codons to each individual ORF. Both cosmid sequences were also compared using webACT (web-based Artemis Comparison Tool, available from the Sanger Institute), a tool that visualises sequence homology between two or more selected sequences.

Likely protein coding sequences were extracted (in FASTA format) and submitted to the basic local alignment search tool database using BLASTP (http://blast.ncbi.nlm.nih.gov/Blast.cgi), using default parameters (Altschul *et al.*, 1990). Putative gene functions were then assigned to the ORFs that had database homologues with known or predicted function.

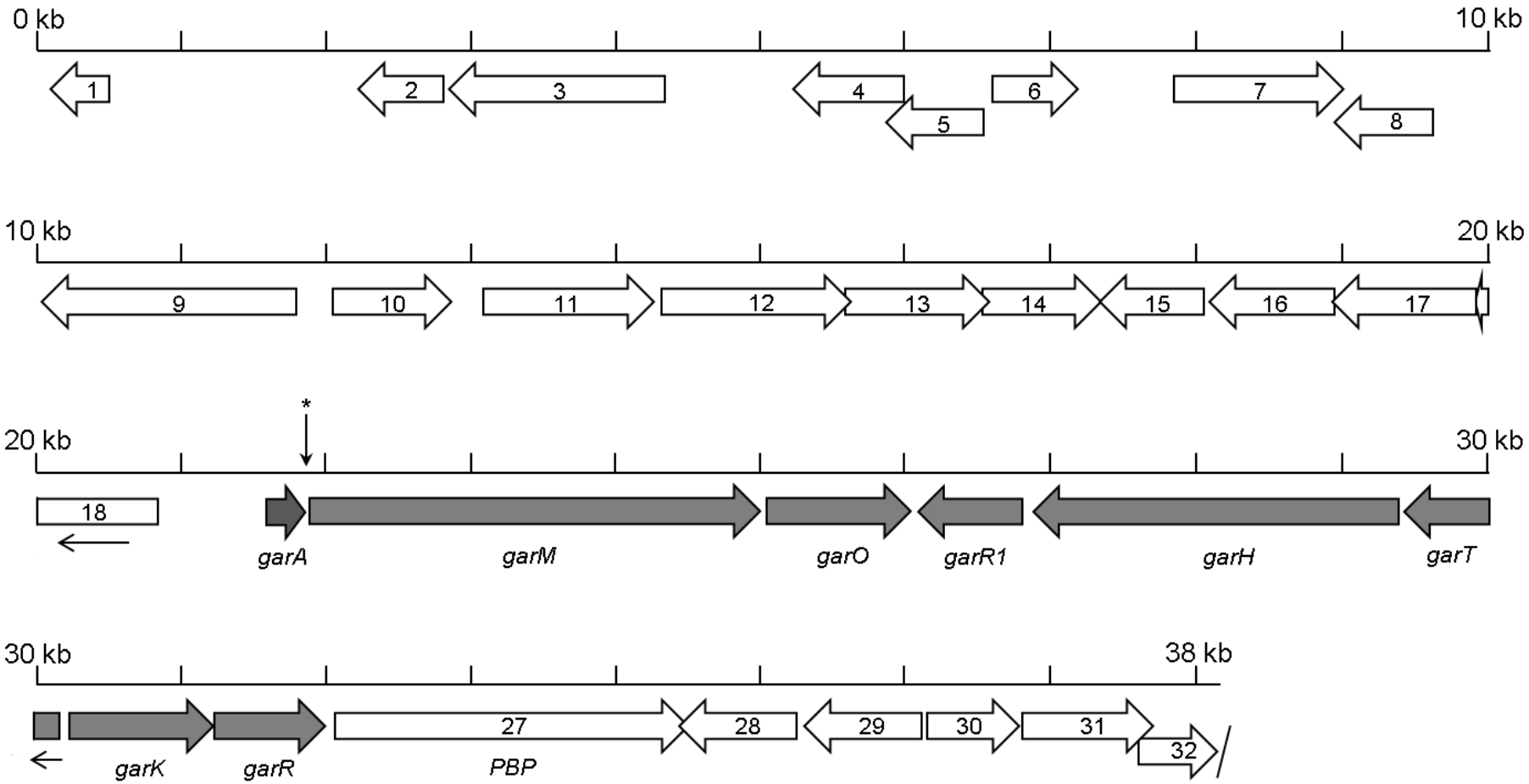
Bioinformatic analysis also provided data on the presence of DNA and RNA secondary structures. Using the computer program Clone Manager (version 6, Sci-Ed Software, http://www.scied.com/index.htm), both the AG14 and AL02 cosmids were analysed for the presence of stem-loop structures that might act as transcriptional terminators or attenuators for some of the genes within the biosynthetic clusters.

3.2.1 Overview of cosmids AG14 and AL02

Annotation of the sequenced cosmids identified 32 putative genes in AG14, and 31 genes in AL02 (figure 3.1). Table 3.1 contains the ORF number, as well as a brief description of the proposed function of each gene.

The genes highlighted in grey in figure 3.1 are hypothesised to be involved in actagardine and DAB biosynthesis. The actagardine gene cluster of AG14 is proposed to consist of eight genes, while the DAB gene cluster of AL02 appears to contain nine biosynthetic genes. The gene clusters are very similar; all the shared biosynthesis genes occupy the same relative position and orientation in both clusters. This high level of similarity between the two clusters is perhaps not surprising given that they produce very similar lantibiotics. This high level of similarity may indicate a recent transfer of the cluster from one *Actinoplanes* species to the other. However, it is also possible that both gene clusters have evolved from a yet unidentified common ancestor. Analysis using webACT failed to identify any significant sequence similarity between the two cosmids (see section 3.5) apart from the shared lantibiotic biosynthesis genes, with the possible exception of four genes to the right of each cluster (discussed later).

A


0 kb
10 kb
1
2
3
4
5
6
7
8
10 kb
20 kb
9
10
11
12
13
14
15
16
17
20 kb
30 kb
*
18
garA
garM
garO
garR1
garH
garT
30 kb
38 kb
garK
garR
27
PBP
28
29
30
31
32

B

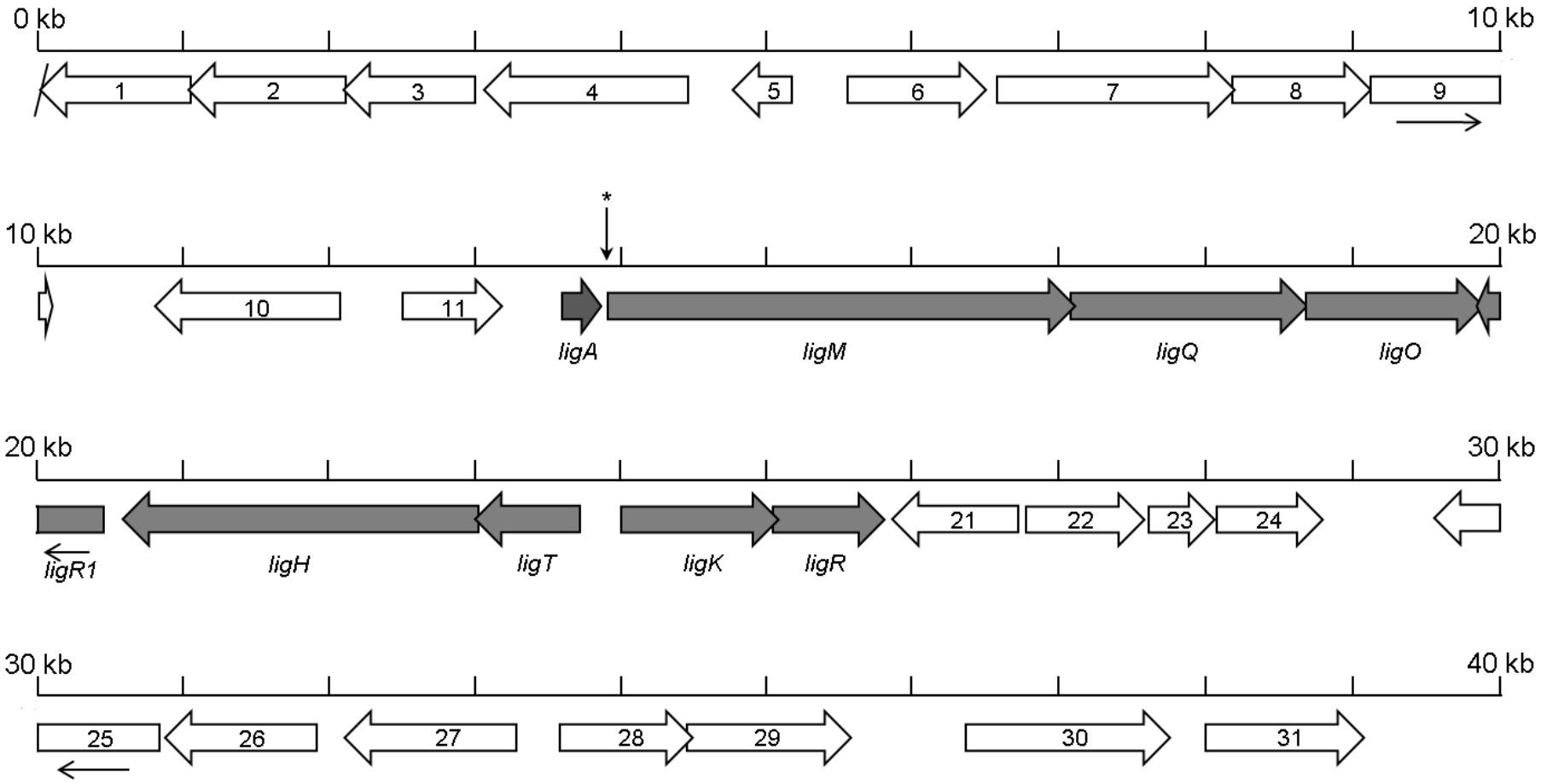

0 kb
10 kb
1
2
3
4
5
6
7
8
9
10 kb
20 kb
*
10
11
ligA
ligM
ligQ
ligO
20 kb
30 kb
ligR1
ligH
ligT
ligK
ligR
21
22
23
24
30 kb
40 kb
25
26
27
28
29
30
31

Figure 3.1: Schematic representation of the cloned DNA fragments present in cosmid AG14 from *A. garbadinensis* (A) and cosmid AL02 from *A. liguriae* (B). Open arrows indicate the orientation and position of each identified ORF. Genes present in the proposed biosynthetic gene clusters are in grey, with the respective *lanA* genes in dark grey. The asterisk highlights the approximate location of a stem-loop structure.

Cosmid AG14

ORF	Proposed function	Length
1	Unknown	146
2	Unknown	149
3	ATPase involved in cell division	518
4	Sugar hydrolase	203
5	Endoglucanase	219
6	Cytosine/adenine deaminase	161
7	Unknown	413
8	Unknown	227
9	Pyruvate oxidase	597
10	Hydrolase or acyltransferase	223
11	Aldose epimerase	396
12	ABC transport component	378
13	ABC sugar transport permease	352
14	ABC transport protein	253
15	Unknown	255
16	ABC transport permease	280
17	ABC transport permease	301
18	Substrate binding ABC transporter	445
19	GarA, actagardine prepropeptide	64
20	GarM, modification enzyme	1053
21	GarO, monooxygenase	336
22	GarR1, response regulator	231
23	GarH, transporter permease	812
24	GarT, ABC transporter	259
25	GarK, response regulator kinase	372
26	GarR, response regulator sensor	220
27	Penicillin binding protein	800
28	Methyltransferase	253
29	Hydrolase	261
30	Response regulator	222
31	Fructose bisphosphate aldolase	286
32	Hydrolase	192

Cosmid AL02

ORF	Proposed function	Length
1	Secretion system protein	375
2	Response regulator	389
3	Unknown	266
4.	Unknown	432
5.	Response regulator ATP binding	136
6.	ABC sugar transporter	363
7.	ABC sugar transporter ATP binding	501
8	ABC transport permease	319
9	ABC transport permease	320
10	Metallopeptidase	486
11	Transcriptional regulator	286
12	LigA, deoxyactagardine prepropeptide	64
13	LigM, modification enzyme	1046
14	LigQ, ABC transporter	575
15	LigO, monooxygenase	347
16	LigR1, response regulator	217
17	LigH, transporter permease	814
18	LigT, ABC transporter	240
19	LigK, response regulator kinase	362
20	LigR, response regulator sensor	218
21	Putative membrane protein	301
22	α/β hydrolase-like protein	290
23	Transcriptional regulator	145
24	Pyruvoyl-dependent arginine decarboxylase	175
25	Diaminopimelate decarboxylase	406
26	Kinase	309
27	Transcriptional regulator	367
28	Glycosyl transferase	317
29	Glycosyl transferase	369
30	Dihydrolipoamide dehydrogenase	459
31	Putative membrane protein	348

Table 3.1: ORFs found within cosmids AG14 and AL02 from *A. garbadinensis* and *A. liguriae* respectively. Listed are the proposed functions and encoded protein lengths (amino acids) of each ORF.

3.3 Biosynthetic gene analysis

Based on its predicted amino acid sequence, *garA* (ORF 19 of AG14) encodes the 64 amino acid actagardine prepropeptide. Similarly, *ligA* (ORF 12 of AL02) encodes the very similar 64 amino acid DAB prepropeptide. The prepropeptides are subsequently modified by their respective biosynthetic enzymes to produce mature actagardine and DAB, including cleavage of the 45 amino acid long leader sequence from each propeptide.

Both prepropeptides consist of two distinct parts, a leader sequence potentially involved in interacting with modification enzymes and maintaining the peptide in an inactive form, and the propeptide, the unmodified version of the mature lantibiotic (Chakicherla & Hansen, 1995, Patton *et al.*, 2008, van der Meer *et al.*, 1994). The prepropeptides share 78% amino acid identity (figure 3.2), with reduced similarity towards the N-terminus, perhaps reflecting less of a role for this part of the leader sequence in interacting with the modification machinery.

A BLASTP search using the GarA prepropeptide identified two other GarA homologs, ClvA from *C. michiganensis* and GnaA, a previously undiscovered actagardine-like prepropeptide encoded by a *lanA* gene found in *Ruminococcus gnavus* (with an identical *R. gnavus gnaA* also found in *Blautia hansenii*, see chapter 5 for more details). GarA shares a high degree of propeptide homology with LigA, ClvA, and some with GnaA, presumably reflecting the production of structurally similar lantibiotics. There is however little overall conservation of amino acid sequence between the leader sequences of these prepropeptides, and those of closely related lantibiotics (figure 3.3). A BLASTP hit with the mersacidin prepropeptide MrsA was also returned, and included a conserved PAG motif in the leader sequence (figure 3.3; see below).

The amino acid sequences of the five prepropeptides were aligned using ClustalW (http://www.ebi.ac.uk/Tools/clustalw2/index.html) (figure 3.3A). The alignment was adjusted manually and revealed a number of conserved residues found in all of the LanAs, particularly in the propeptide region (figure 3.3B). Mersacidin showed less similarity, reflecting its different structural characteristics (figure 3.3).

All the leader sequences with the exception of that for ClvA share a number of conserved residues found in a number of other leader sequences (Chatterjee *et al.*, 2005). In particular, the PAG motif (highlighted in red) is present in all five sequences (DAG in ClvA), although its function is unknown. Both MrsA and GnaA share a NPA motif (part of the highly conserved PAG motif) with cinnamycin (Widdick *et al.*, 2003, Chatterjee *et al.*, 2005). The leader sequence of GnaA is much shorter than the others and is missing many of the conserved residues along the length of the leader peptide. This may reflect the distant evolutionary relationship of GnaA to the other LanA prepropeptides. All five leader sequences lack amino sequence motifs, such as the FNLD motif, found in lantibiotic leader sequences from other rare *Actinomycetes* (Foulston & Bibb, 2010). However, interestingly, the GarA, LigA and MrsA propeptides are all preceded by an unusual AA motif (figure 3.3B) that is not observed in any other lantibiotic prepropeptide (Chatterjee *et al.*, 2005).

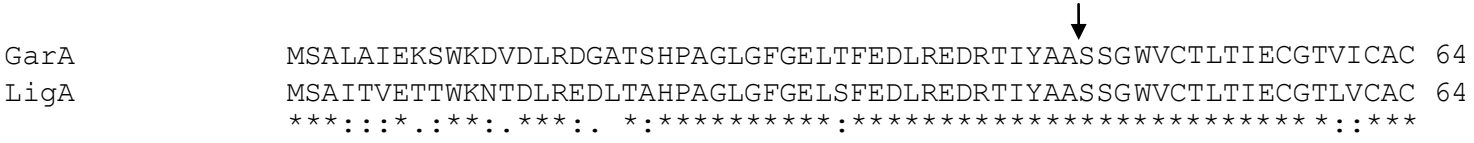


Figure 3.2: The prepropeptides of GarA and LigA. The arrow indicates the cleavage point of the leader sequence from the propeptide. These two peptides share 78% amino acid identity.

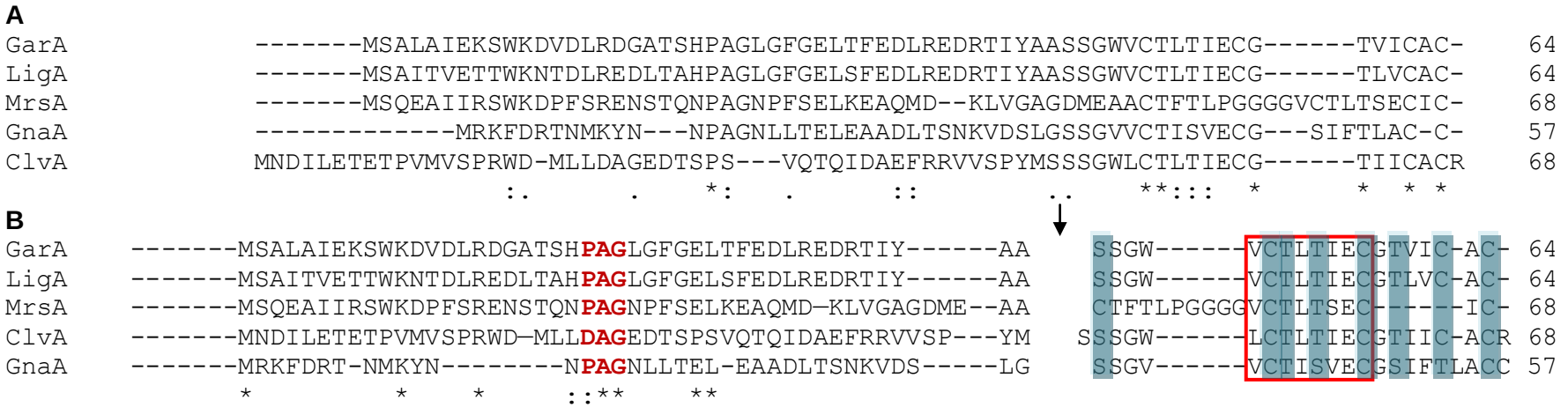


Figure 3.3: Alignment of the five LanAs that were returned in a BLASTP search using the GarA prepropeptide amino acid sequence as a query. The top alignment (A) was returned from ClustalW, the bottom (B) is a manually edited version of A. The PAG motif unique to these leader sequences is highlighted in red. The residues shaded in blue indicate the amino acids involved in lanthionine bridge formation (those for GnaA and ClvA are predicted). The red box highlights the motif likely to be involved in binding of these lantibiotics to lipid II; the arrow indicates the cleavage point.

All five propeptides also contain the lipid II binding motif (shown in the red box in figure 3.3B) that was first described and characterised in mersacidin (Brotz & Sahl, 2000). Presumably these eight residues also enable the other four lantibiotics to bind to lipid II.

Not surprisingly LigA, the DAB prepropeptide, shares homology with the same peptides identified in the GarA BLASTP search. However, a BLASTP search with LigA gave two additional hits to peptides (BSCJ1_09328 and BCSJ1_09323) from *Bacillus cereus SJ1*. These peptides are designated as hypothetical proteins and are found in a cluster of three possible *lanA*s. BCSJ1_09323 is identical to the third *lanA* and BCSJ1_09328 shares 76% amino acid identity with them, mainly in the leader sequence. These are situated next to two putative lantibiotic modification (*lanM*) and transport genes. The N-terminal regions of these peptides share striking similarity to the five lantibiotic leader sequences. Figure 3.4 shows an alignment of these two peptides with the GarA, LigA, MrsA and LicA (lichenicidin; (Dischinger *et al.*, 2009)) prepropeptides. LicA was included in this alignment since a BLASTP searches using BSCJ1_09328 and BCSJ1_09323 as query sequences identified the lichenicidin prepropeptide. The lichenicidin gene cluster also encodes multiple LanA peptides and LanM proteins (two of each), similar to the gene clusters containing BSCJ1_09328 and BCSJ1_09323. While BSCJ1_09328 and BCSJ1_09323 share homology with the leader sequences of other LanA prepropeptides, only BSCJ1_09328 contains a Cys residue in the propeptide region. Since the formation of lanthionine bridges requires Cys residues, BCSJ1_09323 and the other homologs presumably do not encode lantibiotics.

Using the Clone Manager program, stem-loop structures were identified between the *lanA* and *lanM* genes in both AG14 and AL02, and may represent rho-independent transcriptional terminators. The stem-loop in the AG14 cosmid has a ΔG = -157.8 kJ mol^{-1} and starts 11 bp downstream of *garA* and continues for 44 bp. In AL02, the stem-loop has a ΔG = -131.5 $kJmol^{-1}$, starting 17 bp downstream from the stop codon of *ligA* and continuing for 32 bases (Boakes *et al.*, 2010, Boakes *et al.*, 2009). It is likely that both *garAMO* and *ligAMQO* constitute operons. However, the GarA and LigA prepropeptides are required in higher amounts than the biosynthetic enzymes involved in their modification. Thus these stem-loop structures may serve to terminate many of the transcripts initiated upstream of the

two *lanA* genes, and may also act to stabilise the mRNA transcript of the prepropeptide, maintaining an appropriate stoichiometry between the LanA prepropeptide and the modification machinery. Similar stem-loop structures are observed in other lantibiotic gene clusters, and have been proposed to play similar roles (Altena *et al.*, 2000, Widdick *et al.*, 2003).

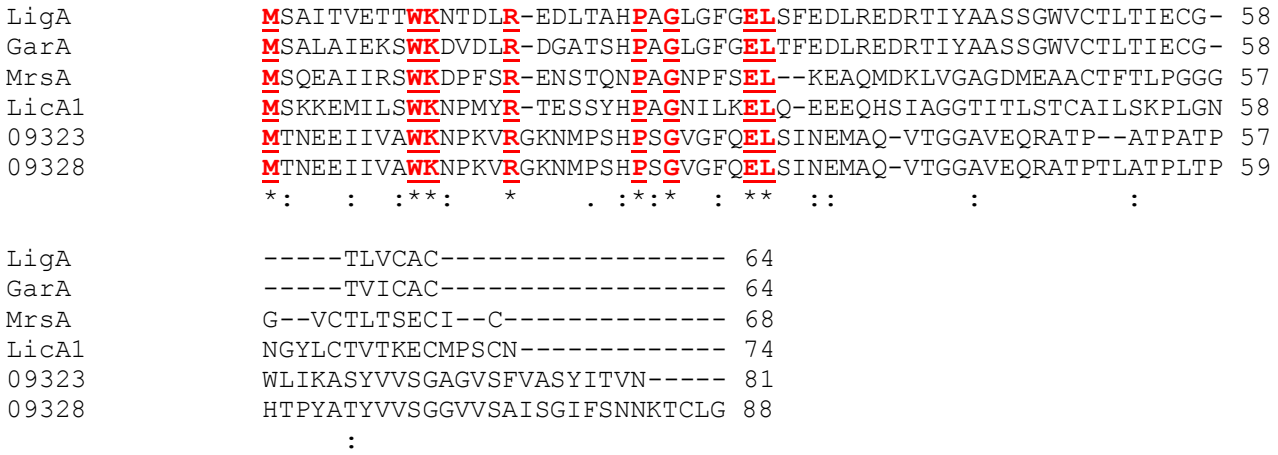


Figure 3.4: Alignment of four lantibiotic LanA peptides with two hypothetical peptides BSCJ1_09328 and BCSJ1_09323 from *B. cereus SJ1*. Residues that are conserved throughout all the putative leader sequences are highlighted in red.

Downstream of *garA* in the AG14 cosmid lies *garM*, which encodes a protein homologous to many characterised LanM enzymes. GarM is 1053 amino acids in length, slightly longer than LigM (1046 amino acids), its homologue in the AL02 cosmid. A BLASTP search using GarM as query revealed homology to MrsM (27 % amino acid identity), the modification enzyme responsible for lantibiotic ring formation in mersacidin biosynthesis; MrsM gave the highest level of identity after LigM. There were also a large number of hits to LanC cyclases from a range of bacteria.

LanM modification enzymes catalyse the selective dehydration of serine and threonine residues, and the formation of the lanthionine bridges via the conjugate addition of cysteine residues to the dehydrated residues. The N-terminal regions of GarM and LigM share all of the conserved residues found in other LanM synthetases that are thought to be required for the dehydration of the serine and threonine residues. Figure 3.5 is an alignment of seven LanM synthetases including those encoded by the actagardine, DAB, michiganin A and mersacidin gene clusters. Other LanMs returned from the BLASTP search from other clusters include GnaM (gnavucin), CinM (cinnamycin) and LctM (lacticin 481), one the most intensively studied LanMs. Dehydration involves one of two Asp residues, at positions 242 or 259 to accept a proton from the Ser/Thr substrate, allowing the deprotonated residue to attack the γ-phosphate of ATP, resulting in a phosphorylated peptide. Both GarM and LigM contain the Mg^{2+} ligands Asn247, Glu261 and Glu446 (highlighted blue, figure 3.5). They also contain Arg residues at positions 399 and 259, one of which is believed to carry out deprotonation of the α-carbon in conjunction with Lys159, resulting in elimination of the phosphate group from the serine and threonine residues (You & van der Donk, 2007). The N-terminal regions of GarM and LigM, like all other LanM synthetases, show no detectable homology to LanB enzymes.

The C-terminal regions of LanM enzymes share 20-27% homology to LanC enzymes, many of which are identified when using GarM as the query sequence in a BLASTP search. These homologous sequences are responsible for the cyclisation reaction required for the formation of lanthionine and methyl-lanthionine bridges (Chatterjee *et al.*, 2005). The most studied example of a LanC cyclase is

NisC, which catalyses the cyclisation of lanthionine rings in nisin (Li *et al.*, 2006). GarM and the other LanM synthetases returned from a BLASTP search using GarM as query contain a number of conserved residues found in both LanC and LanM enzymes. They share the conserved active site His212 residue (figure 3.6; red) and the ligands involved in binding zinc that are were identified in NisC and required for cyclisation activity (figure 3.6; blue) (Paul *et al.*, 2007, Li *et al.*, 2006).

```
GarM      LVYKPRP..GGVDMHMENVIA..PIDLET..EVRYLPRPTRRYSI..IESEKR
LigM      LVYKPRP..GAVDVHMENVIA..PIDLET..EVRYLPRPTRRYSM..IESEKR
ClvM      LVYKPRD..GASDLHLENVIA..PIDLET..SSRHIVRATRSYSL..VEAEER
CinM      FFYKPRS..GGSDLHFENVIV..VCDAET..TARFINWGTQIYAQ..AGREVA
MrsM      IVYKPRS..NAVDFHMQNLIA..LVDLES..TVRQILRGTSRYAN..VNSEKE
GnaM      ILYKPRC..EGGDIHFENLIA..IIDAET..EVRYLIRNTQQYVI..IEHEIE
LctM      LIYKPKS..NLTDLHFENVIS..IIDLET..TCRILFRNTMEYSV..IKSEIS
             *        * *  *        * *     *     *  *        *
           159      242           259     399   406         446(lctM)
```

Figure 3.5: Alignment of all LanM peptides that were returned from a BLASTP search using GarM as query sequence. Highlighted in red are the conserved Asp242/259 residues, thought to accept a proton from the Ser/Thr substrate, and the Mg^{2+} ligands Asn247, Glu261 and Glu446, all involved in dehydration of the prepropeptide.

Figure 3.6: Alignment of the C-terminal regions of LanM proteins with NisC. All members have the conserved His active site in red and metal ligands in blue.

ORF14 in the AL02 cosmid encodes a protein only found in the DAB cluster, and encodes a putative 575 amino acid ABC transporter designated ORF14 by Boakes *et al.* (2010) and LigQ in this study. LigQ contains all of the conserved domains expected for this class of transporter: the ATP binding site (#), Walker A/P-loop (blue), Walker B (red) and D loops (underlined, highlighted), Q-loop lid (green) and H-loop/switch (light blue) (figure 3.7) (Diederichs *et al.*, 2000, Marchler-Bauer *et al.*, 2009, Schneider & Hunke, 1998). ABC transporter systems are involved in the uptake of a wide variety of substrates against a concentration gradient in an ATP-dependent manner. Many of these systems are also involved in antibacterial drug export and hence resistance mechanisms (Schneider & Hunke, 1998). There were a number of BLASTP hits to putative multidrug export ABC transporters that hint to the possibility that LigQ may be part of an immunity mechanism. Although a LigQ homologue is not encoded in the actagardine cluster, a potential functional analogue is a penicillin-binding protein (GarPBP) that could conceivably confer resistance to actagardine. The biosynthetic gene cluster for michiganin A also encodes a likely immunity mechanism in the form of a LanFEG transport system. It also has a putative membrane protein with a PASTA domain similar to those found in penicillin binding proteins; this is discussed further in chapter 5. It is possible that three different gene clusters that produce structurally similar lantibiotics have evolved different mechanisms of self-resistance.

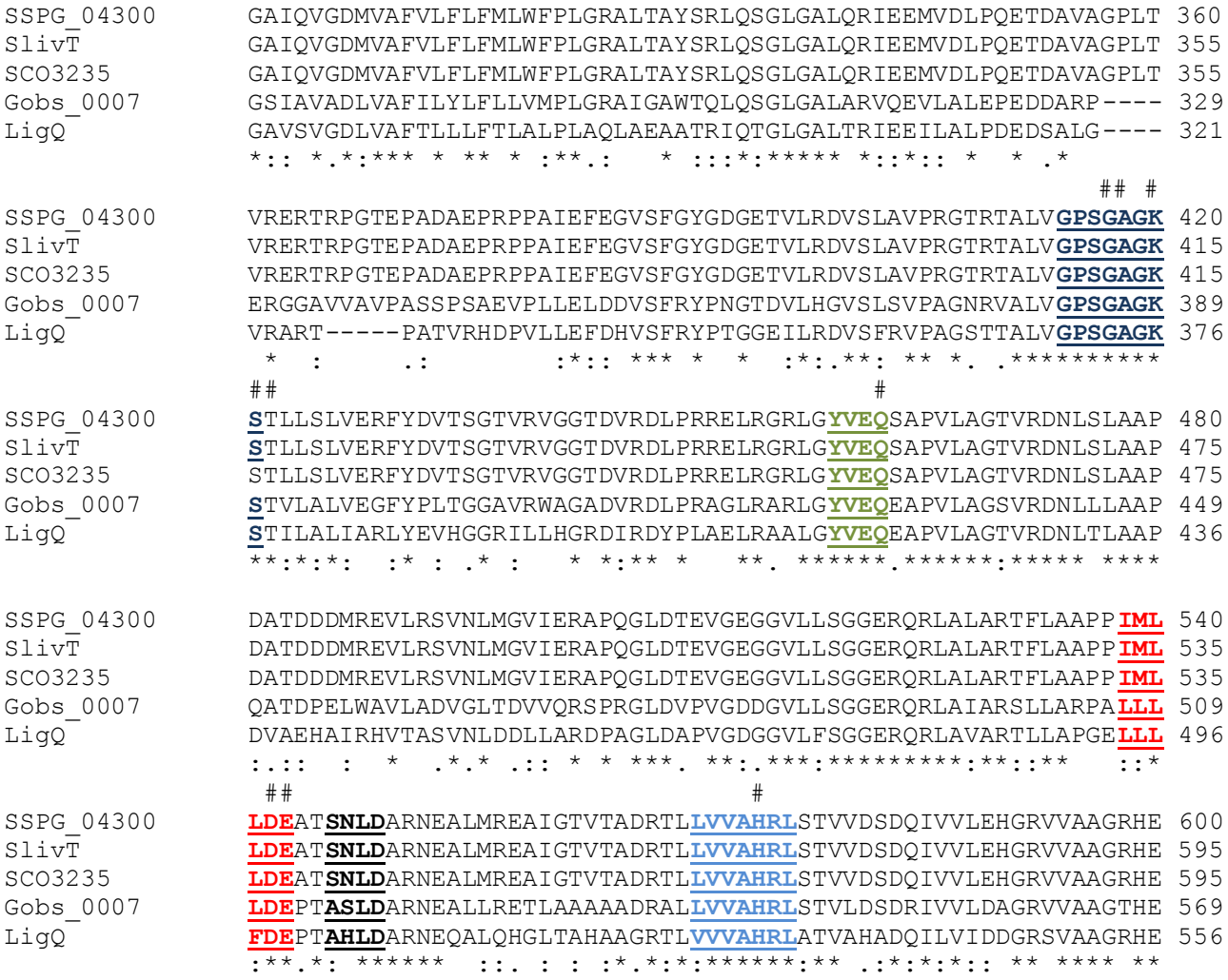
```
SSPG_04300      MSRSAMTSTQTRERPPARKTGAKNAPEPGPPAPPVRKVGLRSLFPYLQGYWPTLGIVAVL 60
SlivT           -----MTSTQTRERPPARKTGAKNAPEPGPPAPPVRKVGLRSLFPYLQGYWPTLGIVAVL 55
SCO3235         -----MTSTQTRERPPARKTGAKNAPEPGPPAPPVRKVGLRSLFPYLQGYWPTLGIVAVL 55
Gobs_0007       --------MSTSAETAARAT-------------------LRQLWPLVRPHRRPLTAAAVL 33
LigQ            -----------MSETAG-----------------------LLRRSLLDHRGKLAAVAGL 25
                             ....                        *   :  :   *  .* *

SSPG_04300      SLVVTLLTLSQPVLTRDVLADIEADRPVGRLVALLIGVLVVVAVLGGVRDYLLQRAAEGL 120
SlivT           SLVVTLLTLSQPVLTRDVLADIEADRPVGRLVALLIGVLVVVAVLGGVRDYLLQRAAEGL 115
SCO3235         SLVVTLLTLSQPVLTRDVLADIEADRPVGRLVALLIGVLVVVAVLGGVRDYLLQRAAEGL 115
Gobs_0007       SLVAAAGALAQPALVARVIEAVGAGRALLPVVGLLVVVLLASSALGAVQDYLLQRTAEGV 93
LigQ            AVAGVGCQLGQPFLIRRVLTAVQSAQPYRQLALAVLAVMVVGAALGAVQQFLLQRTGEAM 85
                ::. .   *.** *   *:  : : :.   :.  :: *::. :.**.*:::****:.*.:

SSPG_04300      VLTARRRLVARLLRLPITEYDQRRTGDMLSRVGADTTMLRAVVTSGLFDTVTNVVMVGGS 180
SlivT           VLTARRRLVARLLRLPITEYDQRRTGDMLSRVGADTTMLRAVVTSGLFDTVTNVVMVGGS 175
SCO3235         VLTARRRLVARLLRLPITEYDQRRTGDMLSRVGADTTMLRAVVTSGLFDTVTNVVMVGGS 175
Gobs_0007       VLSTRRTLADRLLRLPVAEYDRRRTGDLMSRVGADTTLLRATVTSGVVEVAGSAVVAVGA 153
LigQ            VFTVRRTLVAHLLRLPVAAYDERQSGDLVSRVGADTAQVRSVITSGVVDLAGGVLLVGGS 145
                *::.** *. :*****:: **.*::**::*******: :*:.:***:.: . ..::. *:

SSPG_04300      ALMMCLIDPTLFTATALGLGLGVLAVVGLSRRVRGASRDAQDRIGEMTSAVERAISAVRT 240
SlivT           ALMMCLIDPTLFTATALGLGLGVLAVVGLSRRVRGASRDAQDRIGEMTSAVERAISAVRT 235
SCO3235         ALMMCLIDPTLFAATALGLGLGVLAVVGLSRRVRGASRDAQDRIGEMTSAVERAISAVRT 235
Gobs_0007       VVAMALVDGWLLLVTVGSVAVGLAVVVSVARGVQRLSRQAQEEVGAMTAAVERALSAVRT 213
LigQ            IAGMIIIDPVLLGVSLAPVLCGAAGVRLVGRRLRPLSSAVQESIGALTASTTRALGAIRT 205
                   * ::*  *: .:   :  *   *  :.* ::  *  .*: :* :*::. **:.*:**

SSPG_04300      IRASGAEEREGKAVDGYAQQAYHAGMRIARLQAMINPITSTTIQVAFLVVLGLGGARVAS 300
SlivT           IRASGAEEREGKAVDGYAQQAYHAGMRIARLQAMINPITSTTIQVAFLVVLGLGGARVAS 295
SCO3235         IRASGAEEREGKAVDGYAQQAYHAGMRIARLQAMINPITSTTIQVAFLVVLGLGGARVAS 295
Gobs_0007       IRASGATAREVDVVAASAGRAYEAGVRVARLQALVSPAGSVAVQGAFLAVLGLGGYRVAS 273
LigQ            IRVAGATERETALIVAEADRARAAGVRLALVAAQAGPIVRLALQGAFVAVIGFGGYRVAN 265
                **.:**  **   : . * :*  **:*:* : *  .*    ::* **:.*:*:** ***.
```


SSPG_04300 GAIQVGDMVAFVLFLFMLWFPLGRALTAYSRLQSGLGALQRIEEMVDLPQETDAVAGPLT 360
SlivT GAIQVGDMVAFVLFLFMLWFPLGRALTAYSRLQSGLGALQRIEEMVDLPQETDAVAGPLT 355
SCO3235 GAIQVGDMVAFVLFLFMLWFPLGRALTAYSRLQSGLGALQRIEEMVDLPQETDAVAGPLT 355
Gobs_0007 GSIAVADLVAFILYLFLLVMPLGRAIGAWTQLQSGLGALARVQEVLALEPEDDARP---- 329
LigQ GAVSVGDLVAFTLLLFTLALPLAQLAEAATRIQTGLGALTRIEEILALPDEDSALG---- 321
SSPG_04300 VRERTRPGTEPADAEPRPPAIEFEGVSFGYGDGETVLRDVSLAVPRGTRTALVGPSGAGK 420
SlivT VRERTRPGTEPADAEPRPPAIEFEGVSFGYGDGETVLRDVSLAVPRGTRTALVGPSGAGK 415
SCO3235 VRERTRPGTEPADAEPRPPAIEFEGVSFGYGDGETVLRDVSLAVPRGTRTALVGPSGAGK 415
Gobs_0007 ERGGAVVAVPASSPSAEVPLLELDDVSFRYPNGTDVLHGVSLSVPAGNRVALVGPSGAGK 389
LigQ VRART-----PATVRHDPVLLEFDHVSFRYPTGGEILRDVSFRVPAGSTTALVGPSGAGK 376
SSPG_04300 STLLSLVERFYDVTSGTVRVGGTDVRDLPRRELRGRLGYVEQSAPVLAGTVRDNLSLAAP 480
SlivT STLLSLVERFYDVTSGTVRVGGTDVRDLPRRELRGRLGYVEQSAPVLAGTVRDNLSLAAP 475
SCO3235 STLLSLVERFYDVTSGTVRVGGTDVRDLPRRELRGRLGYVEQSAPVLAGTVRDNLSLAAP 475
Gobs_0007 STVLALVEGFYPLTGGAVRWAGADVRDLPRAGLRARLGYVEQEAPVLAGSVRDNLLLAAP 449
LigQ STILALIARLYEVHGGRILLHGRDIRDYPLAELRAALGYVEQEAPVLAGTVRDNLTLAAP 436
SSPG_04300 DATDDDMREVLRSVNLMGVIERAPQGLDTEVGEGGVLLSGGERQRLALARTFLAAPPIML 540
SlivT DATDDDMREVLRSVNLMGVIERAPQGLDTEVGEGGVLLSGGERQRLALARTFLAAPPIML 535
SCO3235 DATDDDMREVLRSVNLMGVIERAPQGLDTEVGEGGVLLSGGERQRLALARTFLAAPPIML 535
Gobs_0007 QATDPELWAVLADVGLTDVVQRSPRGLDVPVGDDGVLLSGGERQRLAIARSLLARPALLL 509
LigQ DVAEHAIRHVTASVNLDDLLARDPAGLDAPVGDGGVLFSGGERQRLAVARTLLAPGELLL 496
SSPG_04300 LDEATSNLDARNEALMREAIGTVTADRTLLVVAHRLSTVVDSDQIVVLEHGRVVAAGRHE 600
SlivT LDEATSNLDARNEALMREAIGTVTADRTLLVVAHRLSTVVDSDQIVVLEHGRVVAAGRHE 595
SCO3235 LDEATSNLDARNEALMREAIGTVTADRTLLVVAHRLSTVVDSDQIVVLEHGRVVAAGRHE 595
Gobs_0007 LDEPTASLDARNEALLRETLAAAAADRALLVVAHRLSTVLDSDRIVVLDAGRVVAAGTHE 569
LigQ FDEPTAHLDARNEQALQHGLTAHAAGRTLVVVAHRLATVAHADQILVIDDGRSVAAGRHE 556

```
SSPG_04300      ELTGTSPLYRELASHQLLIQ 620
SlivT           ELTGTSPLYRELASHQLLIQ 615
SCO3235         ELTGTSPLYRELASHQLLIQ 615
Gobs_0007       ELVETSPLYRELAAAQLLV- 588
LigQ            ELLVRDPTYREFATRQLLT- 575
                **   .* ***:*: ***
```

Figure 3.7: Full length alignment of the top four hits to a BLASTP search using LigQ as query sequence. Conserved regions are highlighted: ATP binding site (#), Walker A/P-loop (blue), Walker B (red) and D loops (underlined, highlighted), Q-loop lid (green) and H-loop/switch (light blue).

Transporters with the highest homology to LigQ are listed as:
SSPG_04300 (ABC transporter): *Streptomyces lividans* TK24
SlivT (ABC-type multidrug transport system, ATPase and permease components): *S. lividans* TK24
SCO3235 (Putative ABC transporter): *Streptomyces coelicolor* A3(2)
Gobs_007(ABC transporter related protein): *Geodermatophilus obscurus* DSM 43160
LigQ (Putative ABC transporter): *A. liguriae*

Following *garM* in the actagardine gene cluster there is an ORF (*garO*) that encodes a 336 amino acid protein with homology to luciferase type monooxygenases and coenzyme F420-dependent enzymes. This suggests that GarO is responsible for the addition of oxygen to the C-terminal methyl-lanthionine ring of actagardine (figure 3.8). There are no other published examples of enzymatic oxidation of lanthionine bridges, and apart from LigO (see below), GarO does not share homology with any previously described lantibiotic-associated biosynthetic enzyme.

Following *ligQ* in the *A. liguriae* biosynthetic gene cluster is *ligO*, which encodes a 347 amino acid putative monooxygenase with 72% amino acid identity to GarO. However, the oxidised version of DAB, actagardine B, is not apparently made by *A. liguriae* (Boakes *et al.*, 2010). There are several possible reasons why *ligO* is unable to effect the addition of an oxygen atom to the C-terminal methyl-lanthionine bridge. *ligR1*, a gene encoding a response regulator, runs in the opposite orientation to *ligO*, and overlaps the 3' end of *ligO* by 24 bp. This overlap may interfere with the production of a full length LigO monooxygenase. The C-terminus of LigO is conserved with other members of the luciferase-like monooxygenase protein family (figure 3.9), indicating that it is likely to be important for enzymatic function. The possible insertion of *ligQ* (ORF14) into the DAB gene cluster may also have interfered with transcription of *ligO*. Boakes *et al.* (2010) also proposed the presence of a stem-loop structure that might act as a transcription terminator, preventing *ligO* expression. The stem-loop structure spans a 30 nucleotide stretch of high GC sequence with a ΔG = -16.10 kcal mol^{-1}) (Boakes *et al.*, 2010). These authors proposed that this stem loop may lead to the production of a truncated and non-functional LigO enzyme.

However, the most likely reason for the inactive *ligO*, discovered in this study, is the presence of a likely frame-shift mutation via the insertion of a G nucleotide towards the C-terminus of the gene. By aligning GarO and LigO, it is possible to identify a likely frame-shift mutation in *ligO* (figure 3.9A, highlighted by the arrow), resulting in an overall amino acid sequence identity of 72%. The revised LigO amino acid sequence was subsequently aligned with GarO, markedly increasing the overall level of amino acid identity to 76% (figure 3.9B). It is possible that this putative frame-shift mutation has disrupted enzyme function. This region of all of the luciferase-like monooxygenases enzymes that show high levels of similarity to GarO

(figure 3.9) is highly conserved, suggesting that is it important for enzyme function. It is unlikely that this is simply an error in the cosmid sequence since fermentation of *A. liguriae* yields only DAB (Boakes *et al.*, 2010).

As stated previously, there are no monooxygenase gene homologues in any previously described lantibiotic biosynthetic gene cluster. This unique tailoring step in the production of actagardine appears to be non-essential for antibiotic activity since DAB appears to be just as active as actagardine. It is possible, however, that this modification is important in the natural habitat of *A. garbadinensis*. For example, this modification could reduce protease susceptibility and enhance solubility.

```
MSMEG_5029      3 LSVLDLVPVRTDQTTADALAATTHLAQTADRLGYHRFWVAEHHNMPAVAATSPPVLIAHLAAHTSQLRLGSGGVMLPNHA  82
garO            7 LSVLDQVPVFRDGSPAEAVRDAVALARSAEQYGYHRFWIAEHHGSAANACAAPEVVTAAVAAATSRIRVGSGGVLLPHYS  86
Rxyl_3089       5 LSVLDLSPVSAGSGSSRALRNTLELARLADRLGYARYWLAEHHNLPSVASPAPEVMIGHVANATARIRVGAGGIMLPNHA  84
ABO_0676       12 FSLLELASIREGDSVGQTLANSVAYARHAESLGFQRFWLAEHHNMQGISSAATSVLVGHIAGATESIRVGSGGVMLPNHP  91
RSc2462         9 LSILDQSPVIHGHSARDAIAATVDLAQMADALGYTRYWCAEHHGLRGVSNPAPEVMIARVASATRHLRVGSGGVMLPYYS  88
                         90       100       110       120       130       140       150       160
                  ....*....|....*....|....*....|....*....|....*....|....*....|....*....|....*....|
MSMEG_5029     83 PLAVAEQFALLEAAHPGRIDLGIGRAPGSDPVTSLALRGAAgrdDRDIEAFPQYLDDVVALM---GArGVRVPlprdllr 159
garO           87 PLKVAETFRVLAALYPGRIDLGFGRAPGGPPAMAELLNPYA---VRTDEAFLEQIGRLLGFL---GDtRT-VS------- 152
Rxyl_3089      85 PLKVVETFRLLEALHPGRIDLGIGRAPGTDPVTAAALRRGR---GAEADDFPELFGELLAFAgegFPeGHPFA------- 154
ABO_0676       92 PLVIAEQFGTLEILYPGRIDLGLGRAPGTDPMTSRALRRE----GLGAEQFPDDVAELQRLL---GPlDTS-Q------- 156
RSc2462        89 PFKLAEQFRLLEALFPNRIDLGVGRAPGGDMRTAQAVAMGD---YNRGDHFPQQVQELIWHLsgtLPpDHPTH------- 158
                        170       180       190       200       210       220       230       240
                  ....*....|....*....|....*....|....*....|....*....|....*....|....*....|....*....|
MSMEG_5029    160 dnyILKATPNATSVPQLWLLGSSMYSAHLAAAKGLPYVFAHHFSGHGTAEALAIYREEFQPSEAAPEPVTFLTVNAVVAE 239
garO          153 ---RVSVTPQVEEPPVPWMLGAGTGSARMAGMLGLPFCFAQFIATEECPEAIEAYRDAFRPSPWLERPQPMLALRVLCAD 229
Rxyl_3089     155 ---SVKAVPEDAELPPIWLLGSSGYSAKAAGEMGLGYAFASHFSPVDPAPPMLAYREHFRPSERFPEPHAILAASVVCAR 231
ABO_0676      157 ---GVNAIPGAGTQVPIWLLGSSLYSAKLAAQRGLPYAFAGHFAPRLYREALRIYRDQFQPSAQLSKPYAMLALPAVPAD 233
RSc2462       159 ---GVILQPEIDTRPELWMLGSSDFGGALAARLGIRFAFAHFINPHVGHIVAQQYRTDFAPGFE-PRPYSAAAVFVIAAD 234
                        250       260       270       280       290       300       310       320
                  ....*....|....*....|....*....|....*....|....*....|....*....|....*....|....*....|
MSMEG_5029    240 TTEEAEALLLPQLQMM---------ARLRTGQPLNALDLVEDAQtTVLTPQGDAVVADGRRRAVVGSPTEAAEQVRALAE 310
garO          230 SDAEAEELATCF--WMscttgwraqVQLTDDYRGGAPNLDDARR-YRLTAEDLALRESRPFLQISGTPAAVGKEIRRLQA 306
Rxyl_3089     232 TRERAEELASSMGLAW---------VRMRTGKPGPLPSPEEALS-YPYSPAERRLLESYRSMQVVGDPASVRKRLAEMAE 301
ABO_0676      234 SDEEAHYLTTTSYQRI---------LALFRGQPLWMKPPVQSMD-GLWNPQEEAGVKDFLGLQLLGTAASMQGQLDQLLD 303
RSc2462       235 TEAEARRCEAAVDLRR---------VQMALGVNGPIPTIEQAEA-HRHTERDQAIIARERPRSLIGTPESVAEQMLRLKD 304
                        330       340       350
                  ....*....|....*....|....*....|....*..
MSMEG_5029    311 EFDVDEVMINPVASAHRgadprtatARDTTLELLAKE 347
garO          307 VYGVSEVVLTTNCPGLP--------ARRRSYELLAGE 335
Rxyl_3089     302 RTAAEEVMVTTMVYDHE--------ARLRSYELLAEA 330
ABO_0676      304 SVEVDELMVTVDLYDPV--------KRRHALDILAsl 332
RSc2462       305 RFVADELIVLTVAPSYR--------ARMRSYELLAQA 333
```

Figure 3.8: Alignment of GarO with the closest members of the luciferase-like monooxygenase family (pfam00296) (Marchler-Bauer *et al.*, 2009).

MSMEG_5029 = alkanal monooxygenase alpha chain (*Mycobacterium smegmatis* str. MC2 155)
Rxyl_3089 = luciferase-like protein (*Rubrobacter xylanophilus* DSM 9941)
ABO_0676 = hypothetical protein ABO_0676 (*Alcanivorax borkumensis* SK2)
RSc2462 = hypothetical protein RSc2462 (*Ralstonia solanacearum* GMI1000)

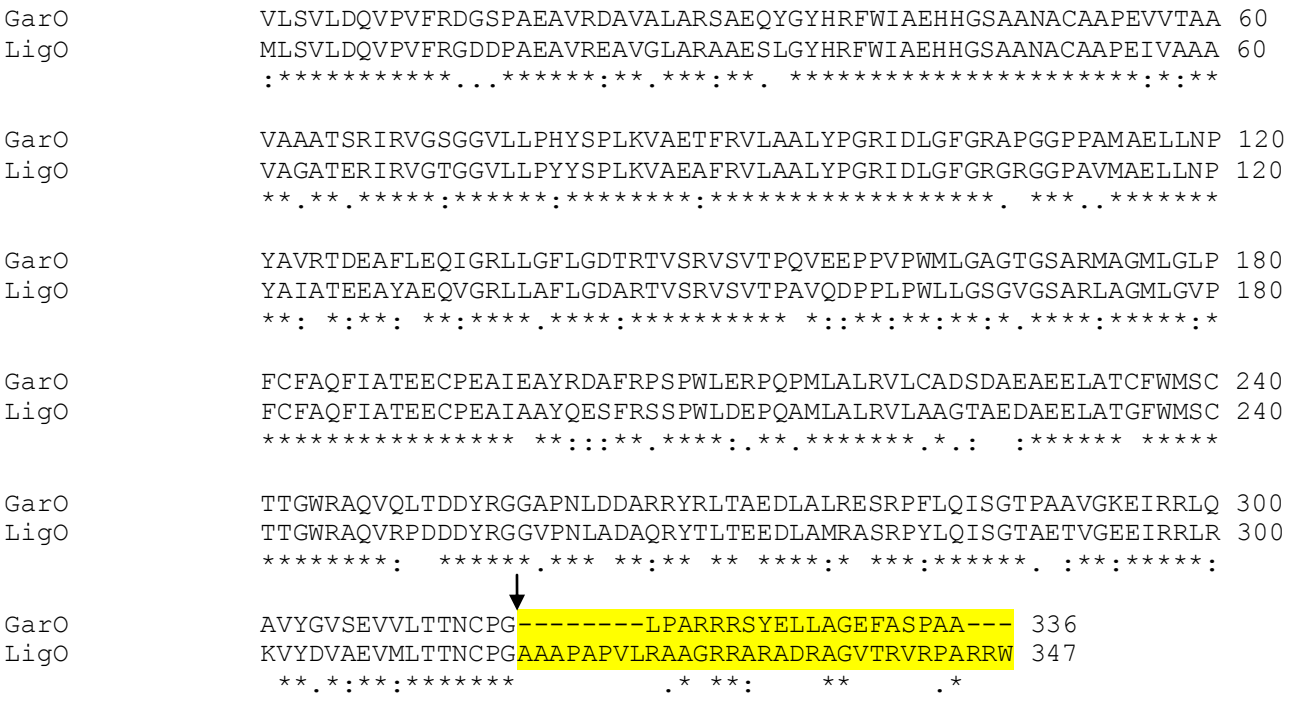


Figure 3.9A: Alignment of GarO and LigO, where a potential frame-shift mutation (arrow) has reduced the overall amino acid sequence identity to 72%

```
GarO            VLSVLDQVPVFRDGSPAEAVRDAVALARSAEQYGYHRFWIAEHHGSAANACAAPEVVTAA 60
LigO            MLSVLDQVPVFRGDDPAEAVREAVGLARAAESLGYHRFWIAEHHGSAANACAAPEIVAAA 60
                :***********...******:**.***:**. **********************:*:**

GarO            VAAATSRIRVGSGGVLLPHYSPLKVAETFRVLAALYPGRIDLGFGRAPGGPPAMAELLNP 120
LigO            VAGATERIRVGTGGVLLPYYSPLKVAEAFRVLAALYPGRIDLGFGRGRGGPAVMAELLNP 120
                **.**.*****:******:********:******************. ***..*******

GarO            YAVRTDEAFLEQIGRLLGFLGDTRTVSRVSVTPQVEEPPVPWMLGAGTGSARMAGMLGLP 180
LigO            YAIATEEAYAEQVGRLLAFLGDARTVSRVSVTPAVQDPPLPWLLGSGVGSARLAGMLGVP 180
                **: *:**: **:****.****:********** *::**:**:**:*.****:*****:*

GarO            FCFAQFIATEECPEAIEAYRDAFRPSPWLERPQPMLALRVLCADSDAEAEELATCFWMSC 240
LigO            FCFAQFIATEECPEAIAAYQESFRSSPWLDEPQAMLALRVLAAGTAEDAEELATGFWMSC 240
                *************** **:::**.****:.**.*******.*.:  :****** *****

GarO            TTGWRAQVQLTDDYRGGAPNLDDARRYRLTAEDLALRESRPFLQISGTPAAVGKEIRRLQ 300
LigO            TTGWRAQVRPDDDYRGGVPNLADAQRYTLTEEDLAMRASRPYLQISGTAETVGEEIRRLR 300
                ********:  ******.*** **:** ** ****:* ***:******. :**:*****:

GarO            AVYGVSEVVLTTNCPGLPARRRSYELLAGEFASPAA- 336
LigO            KVYDVAEVMLTTNCPGRPPRRRSYELLAAELGLTAPA 337
                 **.*:**:******* *.*********.*:. .*.
```

Figure 3.9B: Alignment of GarO and LigO after correction of the putative frame-shift mutation in *ligO*. The LigO amino acid sequence has been adjusted to remove the potential frame-shift mutation, increasing the overall amino acid identity between the two proteins to 76%. Note the much improved alignment (shown in yellow) at the C-terminus compared to Fig. 3.9A.

The product of *garR1* is a 231 amino acid protein with homology to response regulators from two-component signal transduction systems. The N-terminus of GarR1 is similar to known signal receiver domains such as those of CheY, OmpR, NtrC and PhoB (Marchler-Bauer *et al.*, 2009). Figure 3.10 shows the conserved residues present in this domain. GarR1 has the six residues (#) involved in coordination of the Mg^{2+} ion required for phosphorylation of the conserved aspartic acid residue (indicated by P) (Kern *et al.*, 1999). While the protein contains the conserved phosphorylated Asp residue (P), it is missing the Lys-Pro-Phe motif (black box) required for conformational change of the protein after phosphorylation. A similar situation is found in MrsR1 from the mersacidin cluster (Altena *et al.*, 2000, Sola *et al.*, 1999, Boakes *et al.*, 2009). The C-terminus of GarR1 aligns to DNA binding helix-turn-helix motifs and contains conserved residues for both dimerisation and DNA binding (Maris *et al.*, 2002, Boakes *et al.*, 2009). The GarR1 homolog in the *A. liguriae* cluster, LigR1, also lacks the Lys-Pro-Phe motif and intriguingly, a number of other conserved residues. It is unclear how these regulatory proteins are able to alter their configuration upon the presumed phosphorylation event; perhaps instead, the conformational change is induced by ligand binding. It is also unclear what their role is in regulating the production of actagardine and DAB. As mentioned previously, both GarR1 and LigR1 share homology with MrsR1, another orphan response regulator encoded by the mersacidin gene cluster. Guder *et al.* (2002) demonstrated that MrsR1 was required for mersacidin production (Guder *et al.*, 2002). Both the AG14 and the AL02 gene clusters have homologues (*garR/garK* and *ligR/ligK*) of the other regulatory genes identified in the mersacidin cluster, *mrsR2/K2*. It is therefore possible that the regulators in the actagardine and DAB gene clusters share the same functions as their homologues in the mersacidin gene cluster. Thus in the AG14 cluster, *garR1* may control expression of *garA*, and *garRK* may regulate expression of the potential immunity mechanisms, namely *garPBP* and/or the transport system encoded by *garTH* (an analogous situation may be true for the DAB gene cluster, with *ligQ* replacing *garPBP*) (Altena *et al.*, 2000, Guder *et al.*, 2002).

```
                    ##                                                                  P           #
ETR1          11 LVMDEN.[2].SRMVTKGLLV.[3].CEV     TTVS     SNEECLRVV.[3].HKVVFMDVCMP.[1].VENYQIA 69
GarR1         15 LVVDGE.[2].VSIGIKMILE.[3].GFA     VATT.[1].RENLRSAVE.[3].PAVVLLDGHSA.[1].SDGLEVL 74
LigR1          5 LVVAEE.[2].VSIGIKMILE.[3].GFS     VVAA.[1].RDSALAAVS.[3].PAVVLLDAHAT.[1].PESVPLL 64
PhoB           6 LVVEDE.[2].IREMVCFVLE.[3].FQP     VEAE.[1].YDSAVNQLN.[3].PDLILLDWMLP.[1].GSGIQFI 65
NtrC           7 WVVDDD.[2].IRWVLERALA.[3].LTC     TTFE     NGNEVLAAL.[4].PDVLLSXIRMP.[1].MDGLALL 66
FixJ           7 HIVDDE.[2].VRKSLAFMLT.[3].FAV     KMHQ     SAEAFLAFA.[4].NGVLVTDLRMP.[1].MSGVELL 66

                              #                       #   #
ETR1          70 LRIH.[9].RPLLVALSG.[3].KSTKEKC.[3].GLDGVLLKPVS.[ 1].DNIRDVLSDLL 127
GarR1         75 DQLR.[4].PPAIAMLTT.[4].ELVLDSL.[2].GACGFLLRDSQ.[ 1].EQLVAAVRALA 127
LigR1         65 TRLR.[4].GPALAVLAT.[4].STVLESL.[2].GACGFLLKDSQ.[ 1] EQLVAAVRALA 117
PhoB          66 KHLK.[6].DIPVVMLTA.[5].DRVRGLE.[1].GADDYITKPFS.[ 1].KELVARIKAVM 120
NrtC          67 KQIK.[4].MLPVIIMTA.[3].LDAAVSA.[3].GAFDYLPKPFD.[ 1].DEAVALVERAI 119
FixJ          67 RNLG.[4].NIPSIVITG.[3].VPMAVEA.[3].GAVDFIEKPFE.[ 1].TVIIEAIERAS 119
```

ETR1 = *Arabidopsis thaliana*
PhoB = *Escherichia coli*
NrtC = *Salmonella typhimurium*
FixJ = *Rhizobium meliloti*

Figure 3.10: Alignment of the N-termini of representatives of two component response regulator proteins. The # symbol represents residues involved in Mg^{2+} coordination. The black box represents the Lys-Pro-Phe motif missing from both GarR1 and LigR1, and normally involved in the conformational change of response regulators after phosphorylation of the Asp residue (labelled P).

Next to *garR1* lie *garH* and *garT*, which appear to encode transport proteins. *garH* encodes a putative 812 amino acid ABC transporter-associated permease, while *garT* encodes a putative 259 amino acid ATP binding cassette component. GarH has 71% sequence identity with LigH, the homolog present in the DAB cluster, and some amino acid identity to the C-terminal region of LigQ (~28%), and shares a number of the domains conserved in both LigQ and other ABC transporters, such as the ATP binding site, Walker A/P-loop, Walker B and D loops, Q-loop lid and H-loop/switch region (Diederichs *et al.*, 2000, Marchler-Bauer *et al.*, 2009). In addition, the protein may possibly include a Walker C motif involved in ATP hydrolysis (Neuwald *et al.*, 1999), differentiating it from LigQ and further highlighting its putative role as an ATP binding cassette of a two component transporter. Boakes *et al.* (2009) reported 57% identity of GarT to SCO4666 from *S. coelicolor*, a putative ABC-type antimicrobial peptide transporter. Both genes show similarity to a number of transport genes in *S. coelicolor, S. lividans*, and a number of other *Streptomyces*. GarT shows significant identity to CinT (27%) and MrsF (21%), but not to the other mersacidin transport components. As mentioned previously, the regulatory genes within the actagardine and DAB gene clusters are very similar to those found in the mersacidin cluster and have the potential to share similar functions in the regulation of the gene clusters. In mersacidin biosynthesis, MrsFEG confer immunity and are regulated by MrsR2/K2, while MrsT appears to be involved in mersacidin transport out of the cell. Thus while the transport-associated components of the *Actinoplanes* and *Bacillus* gene clusters may not be very similar, the two component regulatory systems encoded by both clusters may still regulate immunity (Altena *et al.*, 2000, Guder *et al.*, 2002).

There is no protease domain associated with GarH or GarT, so they are unlikely to cleave the leader sequence away from the mature lantibiotic. It is likely, as is the case with cinnamycin, microbisporicin and haloduracin, that an extracellular protease is responsible for leader peptide cleavage (Cooper *et al.*, 2008, Foulston & Bibb, 2010, Widdick *et al.*, 2003).

Adjacent to the *garTH* genes in the actagardine gene cluster are *garK* and *garR*, which encode proteins homologous to two-component sensory histidine kinases and response regulators, respectively. GarK shares 68% sequence identity with LigK and 39% sequence identity with ClvK from the michiganin A gene cluster. GarR

shares homology with lantibiotic two-component response regulators, including 73% sequence identity with LigR and 51% identity with ClvR. Much like other response regulators, including GarR1, GarR contains many of the conserved residues involved in signal reception and DNA binding. Interestingly, none of the actagardine response regulators, neither the orphan regulators nor the response regulators that are members of the two-component systems, have the otherwise conserved Lys-Pro-Phe motif; this is also true for the mersacidin response regulators (black box, figure 3.11) (Altena *et al.*, 2000, Marchler-Bauer *et al.*, 2009). In model systems studying the PhoB receiver domain of *E. coli*, Lys108 is likely to be influenced by phosphorylation of the conserved Asp65, changing the conformation of the protein and altering the complementatrity of the dimerisation surface; dimerisation is required for DNA binding and ultimately gene regulation (Sola *et al.*, 1999). As mentioned, GarR1 lacks the key Lys108 residue and therefore is potentially unable to activate its dimerisation surface. The other actagardine response regulators contain the conserved Lys activation residue, so they may be able to relay the activation signal via a conformational change, but would have to do so using different residues compared to those conserved in other regulators.

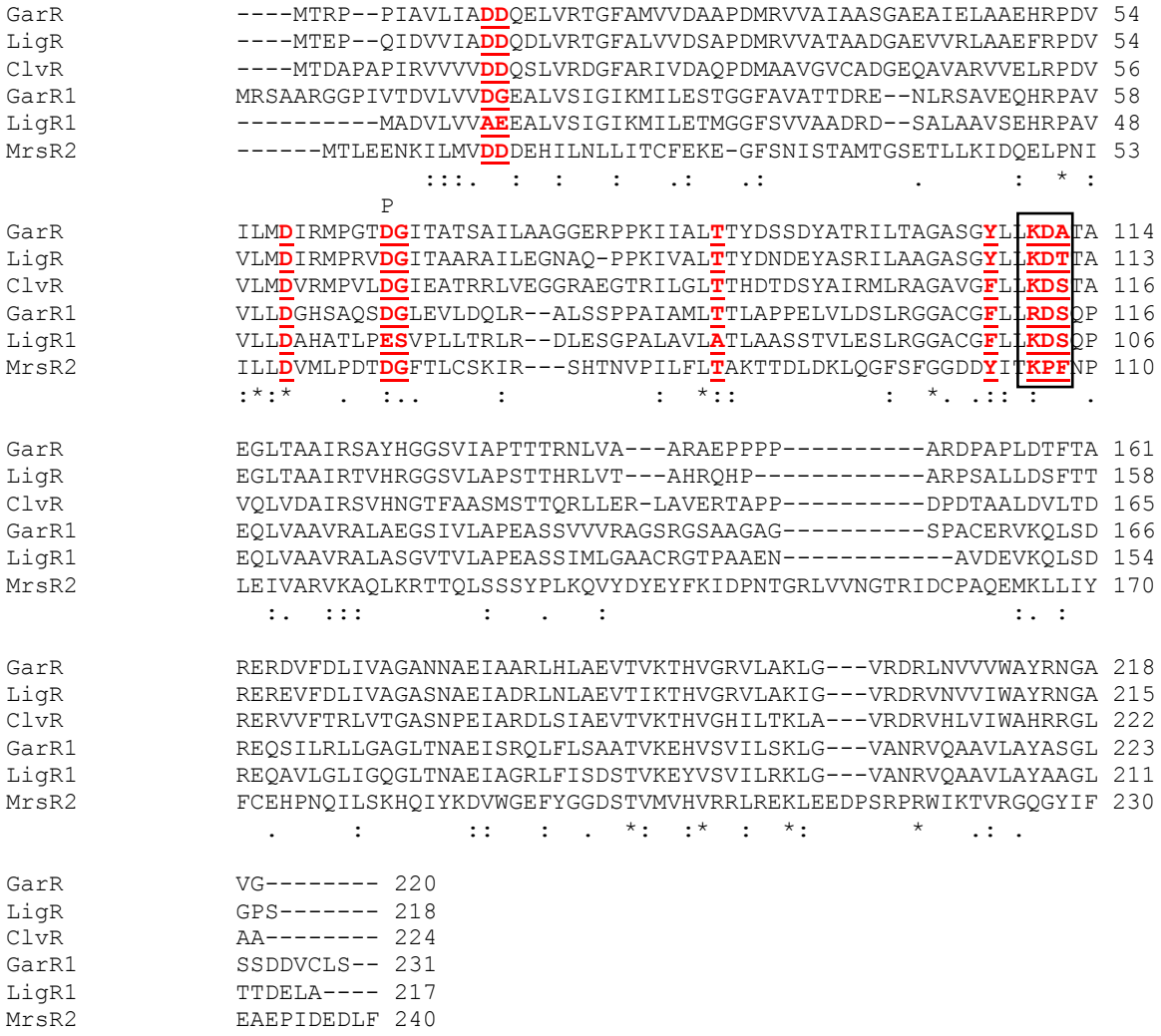


Figure 3.11: Alignment of the response regulator proteins (LanR) from the actagardine, DAB and michiganin A gene clusters, compared with the orphan LanR1 proteins from the actagardine and mersacidin gene clusters. Residues highlighted in red are discussed in the text.

The final potential lantibiotic biosynthetic gene within the actagardine cluster is *garPBP*. This gene is likely to encode a penicillin binding protein (PBP). GarPBP has three domains that are commonly associated with type I PBPs, namely a transglycosylase domain, a transpeptidase domain, and a serine/threonine kinase associated domain (PASTA domain) (Yeats *et al.*, 2002). The N-terminal glycosyltransferase domain catalyses the polymerisation of the glycan component of Lipid II (GlcNac-MurNAc-(pentapeptide) linked via a pyrophosphate bridge to undecaprenyl). The second domain has transpeptidase activity and catalyses the peptide cross-linking of adjacent glycan chains (Terrak *et al.*, 2008). This PBP also contains a PASTA domain at the C-terminus. PASTA domains are associated with the signal-binding sensing component of protein serine/threonine kinases (PSTKs) that are thought to act as sensors for the presence/concentration of unlinked peptidoglycan and activate, directly or indirectly, the expression of cell wall biosynthesis genes (Yeats *et al.*, 2002). It is believed that the PASTA domains on PBPs are required for localisation of the PBP to unlinked glycan chains often found near the site of cell division. They may also be part of an alternative sensor/signalling system, and are generally found in bacteria that do not have a PASTA-containing PSTK (Yeats *et al.*, 2002, Jones & Dyson, 2006).

The occurrence of *garPBP* adjacent to key genes for actagardine biosynthesis led to the hypothesis that the encoded PBP may be a resistance mechanism, functionally analogous to the transporter LigQ in the *A. liguriae* gene cluster. It is plausible that GarPBP might be able to circumvent actagardine-compromised lipid II incorporation into peptidoglycan, and confer immunity (Somma *et al.*, 1977). The presence of a gene encoding a PASTA domain-containing protein of unknown function downstream of the michiganin A gene cluster is intriguing, but this has yet to be implicated in immunity to michiganin A.

3.4 Comparison of actagardine and DAB gene clusters

Initial inspection of the genes within the putative actagardine and DAB biosynthetic gene clusters of *A. garbadinensis* and *A. liguriae* revealed a high degree of homology between the two. The proposed biosynthetic gene clusters consist of a collection of eight genes that are common to both cosmids and that stretch from *lanA* to *lanR*. The order, orientation and apparent operon structure of the biosynthetic genes is also conserved. There are two notable exceptions to this gene conservation, *ligQ* and *garPBP,* and these are discussed later. The proposed minimal gene sets for the two lantibiotics extend from their respective *lanA*s to *garPBP* in AG14 and to *ligR* in AL02. The genes outside of the proposed biosynthetic gene clusters differ in the two cosmid sequences, reflecting their different parental origins. Boakes *et al.* 2010 also noted a rapid reduction in sequence identity outside the proposed minimal gene sets.

Based on gene complement, it is difficult to predict the evolutionary relationship between the actagardine and DAB gene clusters, and whether one is derived from the other, or whether they are both derived from an ancestral cluster.

3.5 Discussion

Several computer programs were used to analyse the ORFs found in two sequenced cosmids, AG14 and AL02, and identified all the genes that are thought to be required for the production of each lantibiotic.

Computational analysis identified putative minimal gene sets for actagardine and DAB biosynthesis. Nine genes may be involved in the biosynthesis of actagardine and DAB (figure 3.12). This includes the genes that encode the prepropeptides of the lantibiotics, which also show similarity to a known actagardine homolog, michiganin A, and to a newly discovered putative lantibiotic, gnavucin. All actagardine family members share the lipid II binding motif that links this group to other well characterised lantibiotics such as mersacidin. This also suggests that the mode of action of this lantibiotic is the same as, or at least similar to, mersacidin, which binds to and sequesters lipid II, thus inhibiting cell wall biosynthesis (Somma *et al.*, 1977, Brotz *et al.*, 1998).

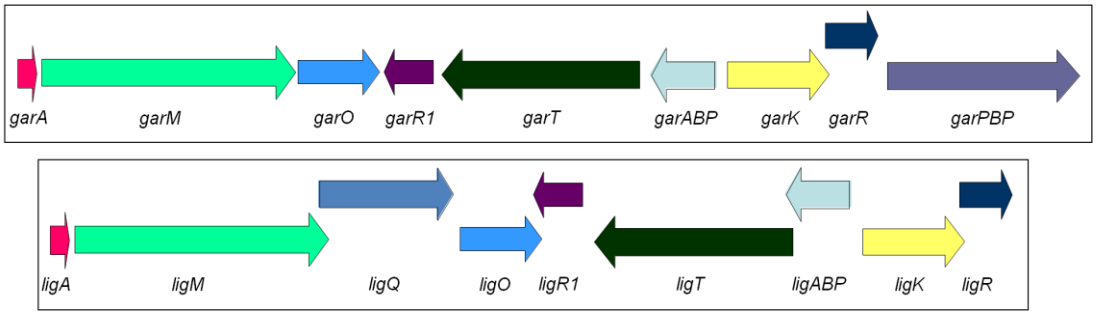


Figure 3.12: Schematic of the minimal gene sets for actagardine (top) and DAB biosynthesis (bottom).

The AG14 and AL02 cosmids were also found to encode typical type II modification enzymes encoded by *garM* and *ligM*, respectively. Both also encode two-component transporters (GarTH and LigTH) presumably responsible for export of the lantibiotic, and two-component regulatory systems (GarRK and LigRK).

The biosynthetic gene clusters in AG14 and AL02 also contain some unusual genes. They each contain an additional atypical response regulator gene (*garR1* and *ligR1*). These genes may be functionally analogous to *mrsR1* of the mersacidin gene cluster and directly regulate expression of their respective *lanA* genes (Guder *et al.*, 2002, Altena *et al.*, 2000, Boakes *et al.*, 2009). A corresponding gene is not present in the michiganin and gnavucin clusters.

Both gene clusters also contain a gene (*garO* and *ligO*), encoding a monooxygenase which has not been found in any other lantibiotic biosynthetic gene cluster. This monooxygenase is likely to be responsible for the addition of an oxygen moiety to the C-terminal methyl-lanthionine bridge of actagardine and appears to be non-functional in the AL02 cluster.

Neither of the two gene clusters contains a gene encoding a peptidase or peptidase domain that could cleave the leader sequence away from the mature peptide. The absence of such a function has been noted for other actinomycete lantibiotic biosynthetic gene clusters i.e. cinnamycin and microbisporicin (Cooper *et al.*, 2008, Foulston & Bibb, 2010, Widdick *et al.*, 2003).

By analogy to mersacidin biosynthesis, *garR1* may sense when it is appropriate for the cell to produce actagardine, activating transcription of the likely *garAMO* operon. The two-component regulatory system *garRK* is likely to regulate other genes in the cluster, such as *garTH*, resulting in transport of the lantibiotic out of the cell. Similar regulatory mechanisms are likely to exist in the DAB gene cluster. It is also possible that an immunity response is evoked to protect the cells, mediated by expressing *ligQ* or *garPBP*, potentially functionally equivalent genes.

While this bioinformatic analysis allows us to speculate on the function of each gene present in the two clusters, it does not provide direct evidence for their roles in lantibiotic biosynthesis. To accomplish this, a mutational study was carried out to determine the minimal gene set required for actagardine biosynthesis, and to determine the roles of each of the genes within it.

3.6 Summary

- All of the genes likely to be involved in actagardine and DAB biosynthesis were identified in cosmids AG14 and AL02, respectively, by bioinformatic analysis.

- Bioinformatic analysis of the nucleotide sequences of the cosmids identified gene clusters that contain a number of lantibiotic biosynthesis-associated proteins.

- Both gene clusters are highly homologous, with the exception of *garPBP* and *ligQ.*

- Both gene clusters contain a monooxygenase that could be responsible for the addition of an oxygen moiety on the C-terminal methyl-lanthionine bridge of actagardine, although LigO appears to be non-functional because of a frame-shift mutation.

- Computational analysis identified two other GarA homologs, ClvA and GnaA.

4 Heterologous production of actagardine and deoxyactagardine B (DAB)

4.1 Introduction

Bioinformatic identification of a number of putative lantibiotic biosynthetic genes in cosmids derived from the natural producers of actagardine and DAB was insufficient by itself to establish that these genes were indeed involved in, and necessary for, production of the two lantibiotics.

To establish that the two cosmids contained all of the genes required for actagardine and DAB biosynthesis, they were initially transferred to a heterologous *Streptomyces* host and assessed for their ability to confer actagardine and DAB biosynthesis. Previous studies had confirmed that the AG14 cosmid contains all of the genes required for the production of actagardine (Boakes *et al.*, 2009). However, DAB had not been produced when the AL02 cosmid was transferred into the same heterologous host (Novacta Biosystems, unpublished).

To determine which genes were required for the production of each lantibiotic, minimal gene sets (MGSs) were constructed based on earlier bioinformatic analysis. This would ultimately allow a mutational analysis of each gene within these minimal genes sets to confirm and understand its involvement in lantibiotic biosynthesis.

4.2 Rationale for the heterologous production of actagardine and deoxyactagardine B

Actagardine, formerly known as gardimycin, was isolated from two separate soil isolates, *Actinoplanes garbadinensis* and *Actinoplanes liguriae*. These organisms, and the lantibiotics they produce, were originally investigated in 1976 by Parenti *et al.*, and included studies on the growth and preservation of the natural producers, as well as purification of the two compounds and analysis of the biological properties of actagardine (Arioli *et al.*, 1976, Coronelli *et al.*, 1976, Parenti *et al.*, 1976). Novacta Biosystems subsequently isolated the actagardine analogue DAB from *A. liguriae*. DAB differs from actagardine by lacking the sulfoxide present in the C-terminal methylanthionine bridge and by the substitution of two amino acids in the C-terminal part of the mature peptide (V15L and I16V) (Boakes *et al.*, 2010).

A. garbadinensis and *A. liguriae* grew slowly, and did not appear to sporulate when grown on a range of agar media. Consequently, they had to be conserved as mycelial cultures, which provided variable inocula for liquid cultures. Furthermore, neither strain has been characterised genetically; only *A. garbadinensis* has been genetically manipulated (Boakes *et al.*, 2010, Boakes *et al.*, 2009), and with some difficulty. Consequently the preferred approach for characterising the lantibiotics that they produce was to express the respective gene clusters in a suitable heterologous host.

In this study, two species were chosen as heterologous hosts; *S. lividans* (strain TK24 and 1326) and *S. coelicolor* (strain M1146) (Gomez-Escribano & Bibb, 2010, Flinspach *et al.*, 2010). Heterologous expression of lantibiotic and modified peptide gene clusters in *Streptomyces* species has been used previously to study cinnamycin (Widdick *et al.*, 2003) and cypemycin (Claesen & Bibb, 2010) biosynthesis, and modified hosts have also been used to increase levels of secondary metabolite production (Flinspach *et al.*, 2010, Komatsu *et al.*, 2010, Widdick *et al.*, 2003, Gomez-Escribano & Bibb, 2010). The cosmid libraries of the natural producers led to the isolation of AG14 and AL02, which could be readily modified for transfer into a heterologous *Streptomyces* host (see below). Many tools have been developed for the genetic manipulation in these hosts. PCR-targeting (also known as REDIRECT technology) allows the efficient replacement of a gene of interest in a cosmid in *E. coli* with a selectable antibiotic resistance cassette, and if required, an in-frame deletion. The PCR reaction used to generate

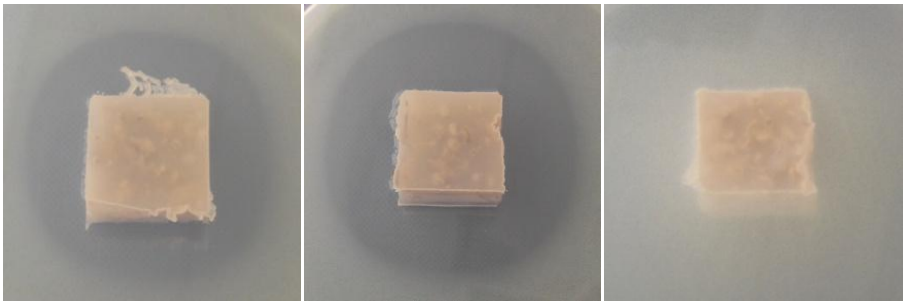
the targeting cassettes requires 39 nt homology extensions to the gene of interest (Gust *et al.*, 2003). The λ RED recombination system (*gam*, *bet*, and *exo*, carried on plasmid pIJ790) increases the rate of recombination when using linear DNA (see chapter 2). The mutant cosmid can then be transferred to a *Streptomyces* host to assess the affect of the mutation.

Novacta Biosystems produced actagardine from the AG14 cosmid in *S. lividans* 1326 (designated Nova263). *S. lividans* TK24 is a commonly used laboratory strain that has been used previously for the heterologous production of polyoxin (Chen *et al.*, 2009), novobiocin and clorobiocin (Eustaquio *et al.*, 2005). M1146 is a derivative of *S. coelicolor* M145 that has had four of its secondary metabolite gene clusters deleted in an effort to increase the level of production from heterologously expressed gene clusters by removing potential alternative sinks for carbon and nitrogen (Gomez-Escribano & Bibb, 2010). These strains are also useful for carrying out bioassays. Neither one produces antibiotics that are active against the test organism *Micrococcus luteus*, and therefore any antimicrobial activity that is observed is likely to result from the production of actagardine or DAB, or one of their derivatives. The construction of specific mutants, combined with such bioassays, would enable the identification of genes that are essential for actagardine or DAB production. A mutation that resulted in loss of antimicrobial activity can thus be readily attributed to loss of actagardine or DAB production, which could be subsequently confirmed by Matrix-Assisted Laser Desorption/Ionization – Time-of-Flight (MALDI-ToF) mass spectrum analysis (see chapter 2).

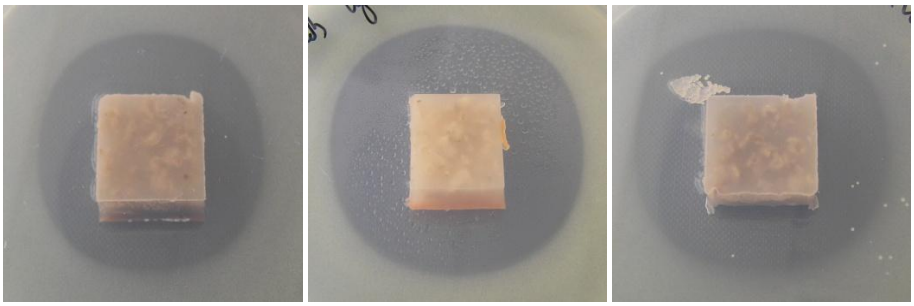
4.3 Transfer of the AG14 and AL02 cosmids to *Streptomyces lividans* TK24 and product analysis

AG14 and AL02 were supplied as SuperCos derivatives in *E. coli* DH10B by Novacta Biosystems. The cosmids were PCR-targeted with an *Ssp*I restriction fragment from pIJ10702 that contained the *Streptomyces* phage ΦC31 attachment site *att*P and integrase (*int*). The fragment was inserted into the neomycin resistance gene (*neo*) of AG14 and AL02 creating pIJ12206 and pIJ12207, respectively. Novacta Biosystems had previously made an equivalent construct to pIJ12206 named CosAG14HEapra in a similar way, and this was used in this study as a positive control. The cosmids were transferred to *S. lividans* TK24 by conjugation using the methylation deficient host *E. coli* ET12567 pUZ8002 (see chapter 2).

Several *S. lividans* TK24 exconjugants from each mating were tested for the successful integration of the respective cosmid by colony PCR. lanA-con3 primers (table 2.12, chapter 2) were used to amplify a ~1 kb PCR product that was sequenced, that included the entire *lanA* gene and short stretches of intergenic sequence up and downstream of the gene. Three *S. lividans* TK24 exconjugants confirmed to contain the AG14 (M1201) and AL02 (M1202) cosmids were analysed in a bioassay to test for lantibiotic biosynthesis. Cubes of OB-N agar inoculated with a clone containing one of the cosmids was incubated for seven days. The cubes were subsequently placed on plates of L agar pre-seeded with the sensitive test organism *M. luteus* and further incubated overnight. A halo of clearing in the *M. luteus* lawn around the cube of the *S. lividans* exconjugants (figure 4.1) indicated the production of antimicrobial activity. Comparison of the zones of inhibition after seven days of incubation of M1201 and M1202 suggested that they produce approximately the same amount of lantibiotic as the natural producers.



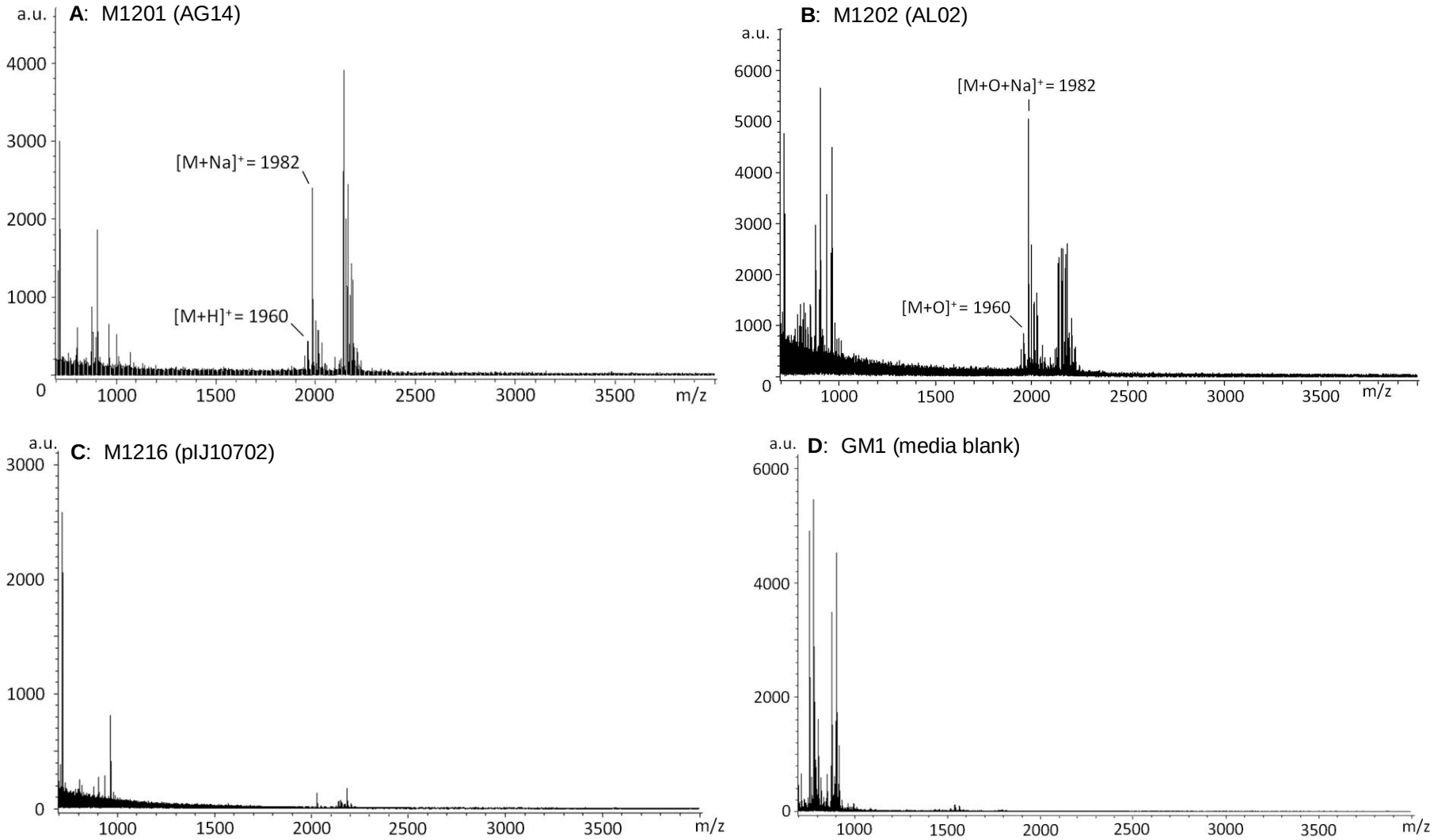
M1201 (AG14) M1202 (AL02) M1216 (pIJ10702)



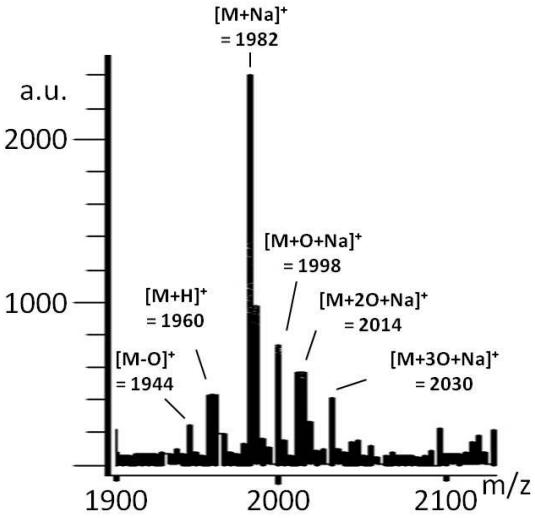
A. garbadinensis *A. liguriae* Nova263

Figure 4.1: Cubes of OB-N agar were inoculated with the natural producers, *S. lividans* TK24 containing the targeted AG14 or AL02 cosmids (strains M1201 and M1202, respectively), Nova263 (*S. lividans* 1326/CosAG14HEapra) from Novacta and *S. lividans* TK24 containing pIJ10702 (M1216) (empty vector control) and incubated for seven days. The cubes were placed on a plate of L agar seeded with *M. luteus*. Zones of inhibition, attributable to actagardine or DAB production, can be seen around the natural producers and the *S. lividans* derivatives containing the cloned gene clusters, but not around the empty vector control.

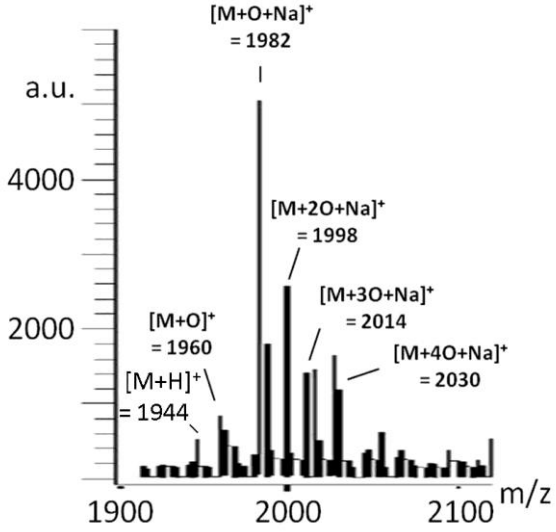
Supernatants of 7 day old liquid cultures of three M1201 and three M1202 clones grown in TSB were analysed to confirm that the antimicrobial activity was the result of actagardine and DAB production, respectively (see section 2.9.2, chapter 2). This was accomplished using MALDI-ToF mass spectrometry. All M1201 clones produced spectra with a mass peak of 1960 Da, the predicted mass of Ala(0)-actagardine $[M+H]^+$ (figure 4.2A). The M1202 clones yielded spectra with the same mass peak at 1960 Da, indicating production of Ala(0)-deoxyactagardine B with the addition of one oxygen adduct $[M+O]^+$ (figure 4.2B). No such peaks were observed in the three negative control clones of M1216 (*S. lividans* TK24 containing pIJ10702 (the empty vector), figure 4.2C) or media only control (figure 4.1D). In the expanded spectra of M1201 and M1202 (figure 4.1E and F), peaks corresponding to a number of actagardine and DAB adducts can be observed. Both spectra show that ions of actagardine with a sodium adduct and DAB with an additional oxygen moiety and a sodium adduct are the most abundant in the two spectra (1982 Da). Oxidation caused by the ionisation of the sample during analysis is probably responsible for the lack of a DAB peak in the M1202 spectra. This will be discussed in detail in the subsequent text. Further analysis of a Δ*garO* mutant using both LC-MS and MALDI-ToF confirmed the production of DAB by the genes present in the AL02 cosmid (see later).


A: M1201 (AG14)
a.u.
4000
3000
2000
1000
0
[M+Na]+ = 1982
[M+H]+ = 1960
1000
1500
2000
2500
3000
3500
m/z
B: M1202 (AL02)
a.u.
6000
5000
4000
3000
2000
1000
0
[M+O+Na]+ = 1982
[M+O]+ = 1960
1000
1500
2000
2500
3000
3500
m/z
C: M1216 (pIJ10702)
a.u.
3000
2000
1000
0
1000
1500
2000
2500
3000
3500
m/z
D: GM1 (media blank)
a.u.
6000
4000
2000
0
1000
1500
2000
2500
3000
3500
m/z

E: M1201 (AG14) expanded



F: M1202 (AL02) expanded

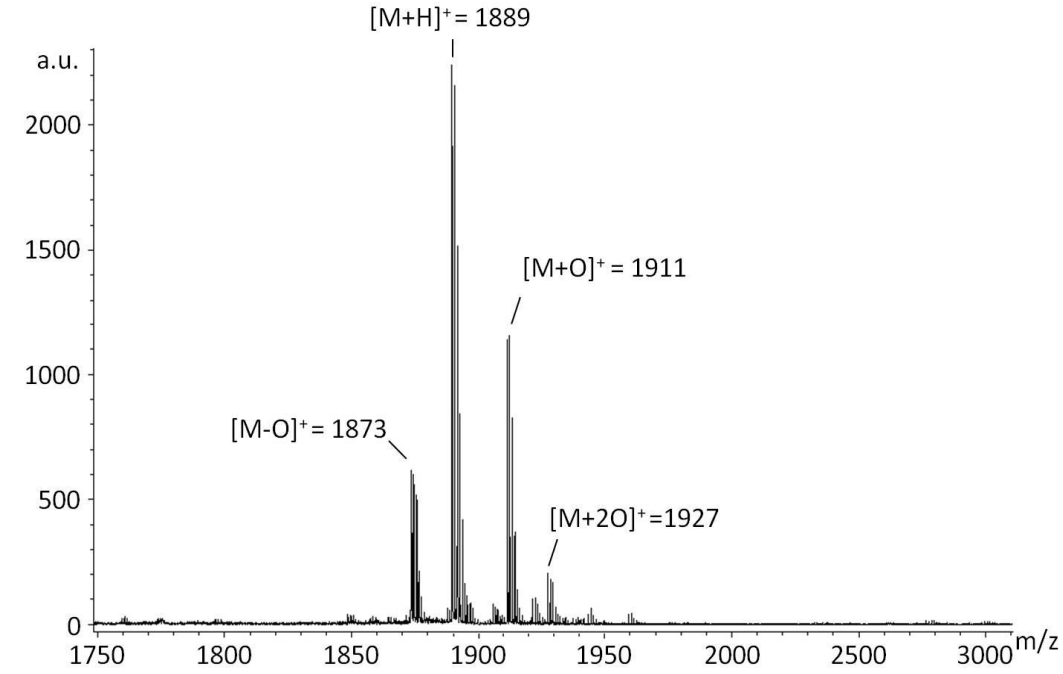


m/z	Description	
	Actagardine	Deoxyactagardine B
1944	Ala(0)-deoxyactagardine + H	Ala(0)-deoxyactagardine B + H
1960	Ala(0)-actagardine + H	Ala(0)-deoxyactagardine B + O
1982	Ala(0)-actagardine + Na	Ala(0)-deoxyactagardine B + O + Na
1998	Ala(0)-actagardine + O + Na	Ala(0)-deoxyactagardine B + 2O + Na
2014	Ala(0)-actagardine + 2O + Na	Ala(0)-deoxyactagardine B + 3O + Na
2030	Ala(0)-actagardine + 3O + Na	Ala(0)-deoxyactagardine + 4O + Na

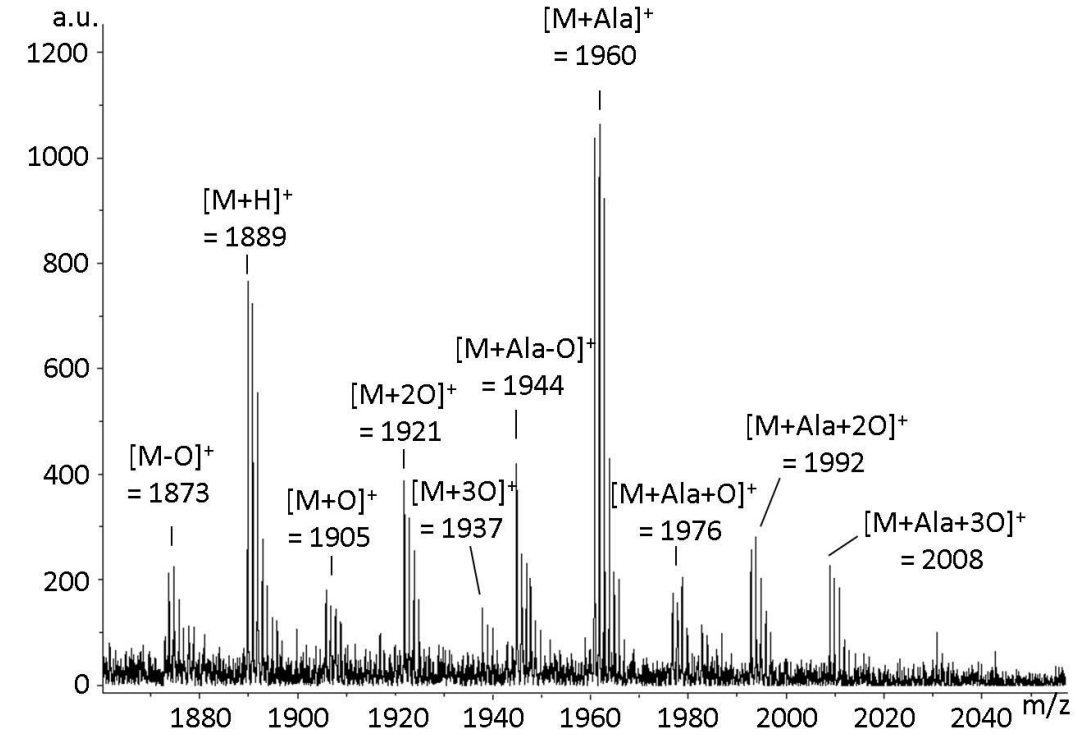
Figure 4.2: MALDI-ToF spectra of culture supernatants of M1201 (AG14 – A and E), M1202 (AL02 – B and F), M1216 (pIJ10702 (empty vector control – C)) and GM1 media control (D). The masses for the labelled peaks in E and F are shown in the table above. The mass peaks for actagardine and DAB $[M+H]^+$ are only present in the clones containing either the AG14 or AL02 cosmid, and are absent from the vector or media only controls.

MALDI-ToF analysis is a rapid and robust method of identifying peptides present in culture supernatants (Foulston & Bibb, 2010, Claesen & Bibb, 2010). MALDI-ToF analysis of supernatants containing actagardine or DAB produced many peaks, including those corresponding to Na adducts, and to the presence of oxidation products of the two lantibiotics, presumably reflecting oxidation of lanthionine bridges during ionisation (figure 4.2). Highly numerous oxidation products were not observed in the previous electro-spray mass spectrometric analyses of actagardine and DAB (chapter 5, and (Boakes *et al.*, 2010, Boakes *et al.*, 2009)). Such oxidation products were also observed upon MALDI-ToF analysis of microbisporicin (Foulston & Bibb, 2010). This is problematic since such oxidation could compromise the subsequent analysis of the Δ*garO* mutant (see chapter 5). MALDI-ToF analysis also identified a variant of actagardine where the alanine at position -1 (with respect to the cleavage site, but named 0 here and elsewhere) has not been cleaved from the mature peptide (figure 4.2). This phenomenon was described previously by Boakes *et al.* (2009), when the AG14 cosmid was expressed in *S. lividans* 1326. MALDI-ToF spectra of *A. garbadinensis* contain very small amounts of the ala(0)- derivative, and fewer oxidation events compared to a *S. lividans* 1326 derived actagardine standard, suggesting that nearly all of the prepropeptide is processed correctly in the natural producer (figure 4.3). The AG14 and AL02 cosmids do not encode a specific protease, and it appears that the specificity of the protease(s) present in *S. lividans* that are able to remove the leader sequence from the modified peptide is different to those in the natural producer. Interestingly, none of the thus-far sequenced actinomycete lantibiotic gene clusters encode a discernible protease or protease domain (Widdick *et al.*, 2003, Foulston & Bibb, 2010, Chatterjee *et al.*, 2005), and other host proteases are presumably responsible for cleavage.

A


[M+H]+ = 1889
a.u.
2000
1500
1000
500
0
[M+O]+ = 1911
[M-O]+ = 1873
[M+2O]+ =1927
1750
1800
1850
1900
1950
2000
2500
3000
m/z

B


a.u.
1200
1000
800
600
400
200
0
[M+Ala]+
= 1960
[M+H]+
= 1889
[M+Ala-O]+
= 1944
[M+2O]+
= 1921
[M-O]+
= 1873
[M+O]+
= 1905
[M+3O]+
= 1937
[M+Ala+2O]+
= 1992
[M+Ala+O]+
= 1976
[M+Ala+3O]+
= 2008
1880
1900
1920
1940
1960
1980
2000
2020
2040
m/z

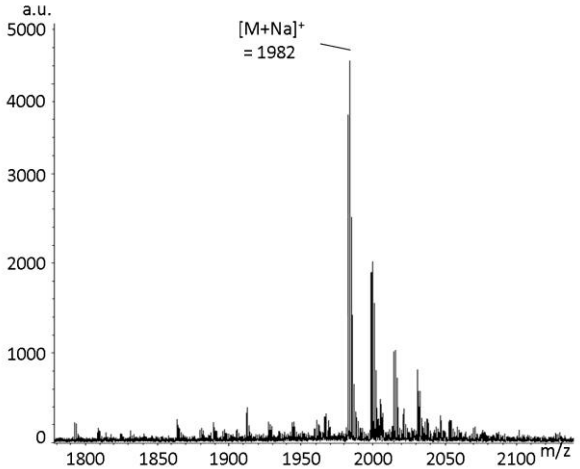
m/z	.	Description
1873.749	$[M-O]^+$	Deoxyactagardine
1889.844	$[M+H]^+$	Actagardine
1905.820	$[M+O]^+$	Actagardine + O
1921.821	$[M+2O]^+$	Actagardine + 2O
1937.815	$[M+3O]^+$	Actagardine + 3O
1944.777	$[M+Ala-O]^+$	Ala-(0)deoxyactagardine
1960.875	$[M+Ala]^+$	Ala(0)-actagardine
1976.825	$[M+Ala+O]^+$	Ala(0)-actagardine + O
1992.852	$[M+Ala+2O]^+$	Ala(0)-actagardine + 2O
2008.882	$[M+Ala+3O]^+$	Ala(0)-actagardine + 3O

Figure 4.3: MALDI-ToF spectra of culture supernatants of *A. garbadinensis* (A) and a *S. lividans* 1326-derived actagardine standard (B). Note the fewer oxidation species of actagardine from the natural producer, as well as the absence of Ala(0)-derivatives The table below the spectra lists the actagardine species that correspond to the highlighted masses.

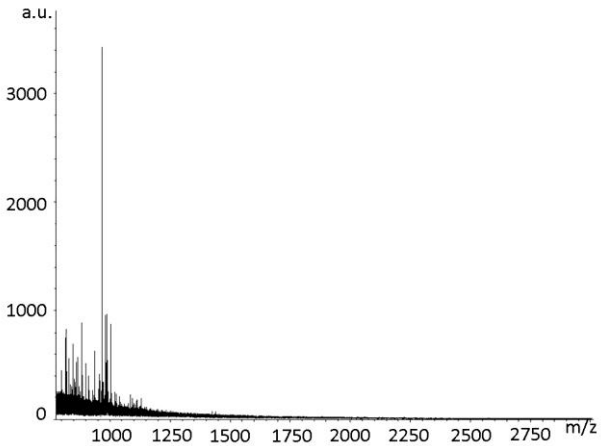
4.4 Transfer of the AG14 cosmid to *Streptomyces coelicolor* M1146 and product analysis

The AG14 cosmid was transferred to *S. coelicolor* M1146 and *S. lividans* 1326, in the latter case mirroring the experiment carried out by Novacta Biosystems (Boakes *et al.*, 2010, Boakes *et al.*, 2009). M1146 is a derivative of *S. coelicolor* M145, a fully sequenced strain of *S. coelicolor* A3(2) that lacks the plasmids SCP1 and SCP2. M1146 is deleted for four *S. coelicolor* A3(2) M145 secondary metabolic gene clusters (Gomez-Escribano & Bibb, 2010), those encoding actinorhodin, undecylprodigiosin, calcium-dependent antibiotic and Cpk production (CpK is a polyketide antibiotic (Pawlik *et al.*, 2007, Flinspach *et al.*, 2010)). All of these gene clusters were deleted in an attempt to increase the level of production from heterologously expressed gene clusters. As M1146 does not produce the more abundant secondary metabolites found in wild type *S. coelicolor*, there ought to be more primary metabolites available for the production of heterologous compounds. Another advantage of using M1146 is that it lacks antibiotic activity, thus facilitating subsequent bioassays on derivatives containing heterologous gene clusters. The AG14 cosmid, a Δ*garA* mutant derivative (see chapter 5) and an empty vector control were conjugated into M1146 via *E. coli* ET12567 pUZ8002, and the resulting strains grown in TSB liquid medium. Culture supernatants (50 µl) were applied to wells made in L agar pre-seeded with *M. luteus* (see chapter 2, section 2.9.2). Clear zones of growth inhibition of the *M. luteus* lawn were observed around the clones containing AG14, but not around the two negative controls (figure 4.4). Production of actagardine was confirmed by MALDI-ToF analysis (figure 4.5). The M1146 clones containing the AG14 cosmid produced more lantibiotic than *S. lividans* 1326, measured by the increased zone of inhibition.

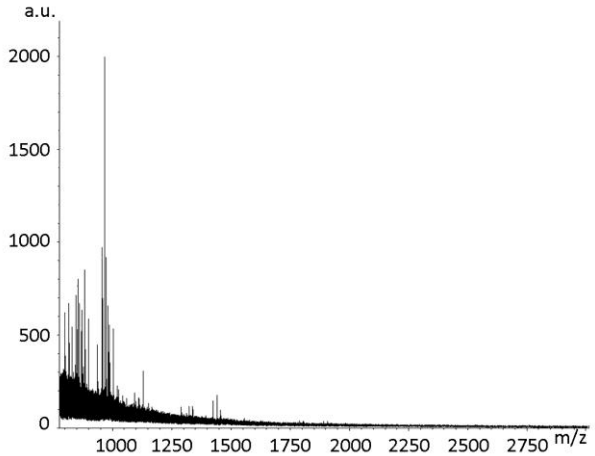
Figure 4.4: Bioassays of culture supernatants of *S. coelicolor* M1146 and *S. lividans* 1326 containing pIJ12206, pIJ12206Δ*garA*::scar and an empty vector control (pIJ10702).



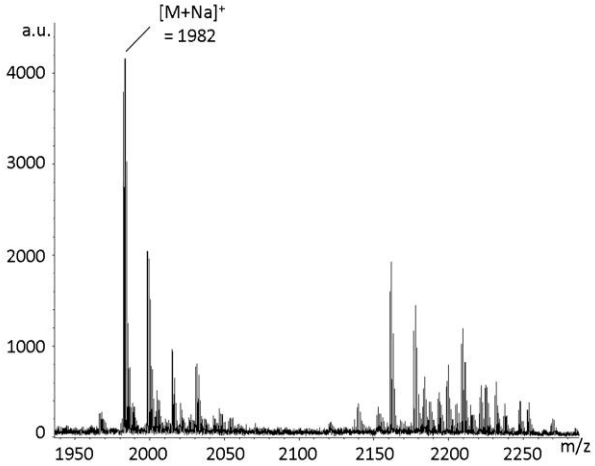
M1146 pIJ12206



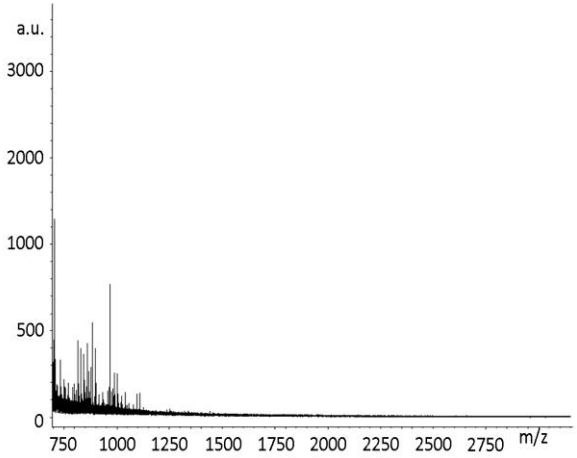
M1146 pIJ12206Δ*garA*::scar



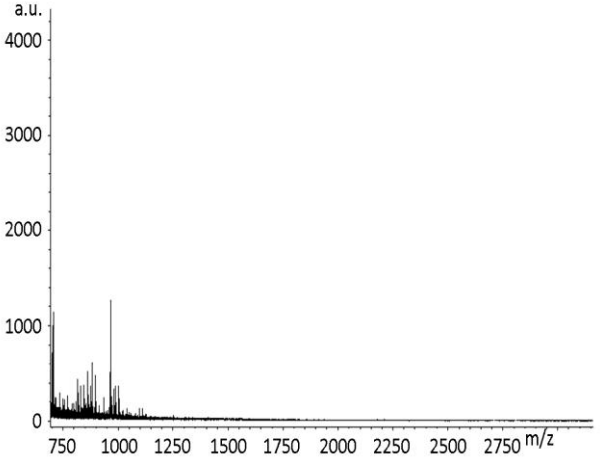
M1146 pIJ10702



S. lividans 1326 pIJ12206



S. lividans 1326 pIJ12206Δ*garA*::scar



S. lividans 1326 pIJ10702

Figure 4.5: MALDI-ToF spectra of culture supernatants of M1146 and *S. lividans* 1326 clones containing pIJ12206, pIJ12206Δ*garA*::scar or pIJ10702. The $[M+Na]^+$ ala(0)-actagardine peak is present in only the clones containing pIJ12206.

4.5 Determination of a minimal gene set for actagardine production

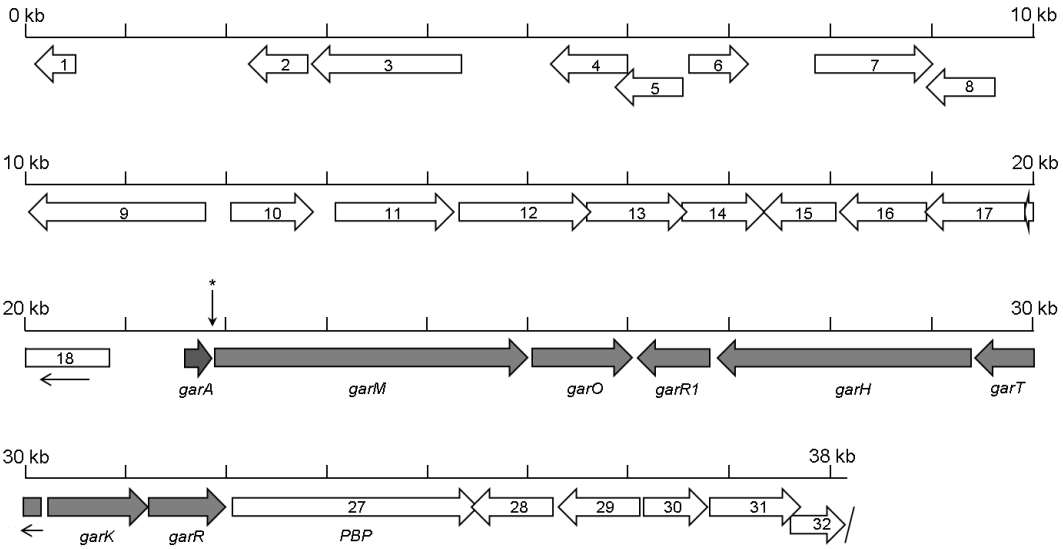
Computational analysis predicted nine putative biosynthetic genes involved in the biosynthesis of both actagardine and DAB. While computational analysis could identify genes in these clusters that were homologous to genes involved in the production of other lantibiotics, it could not discount the involvement of other genes present in both cosmids. Computational analysis was limited to the two sequenced cosmids, AG14 and AL02, thus there was also the potential for other genes present in the genomes of *A. garbadinensis* and *A. Liguriae* to have roles in actagardine and DAB biosynthesis, respectively. Resolution of these issues required the establishment of minimal gene sets (MGS) for the biosynthesis of both lantibiotics in a heterologous host.

Novacta Biosystems previously identified a number of cosmids from the natural producers that hybridised to the *garA* probe. Using primers that hybridise to the bacteriophage T3 and T7 promoters that are located at the ends of the inserts in the SuperCos vector, nine *A. garbadinensis* cosmids and ten *A. liguriae* cosmids were end-sequenced by Novacta. One cosmid from each natural producer was fully sequenced (Novacta Biosystems, unpublished).

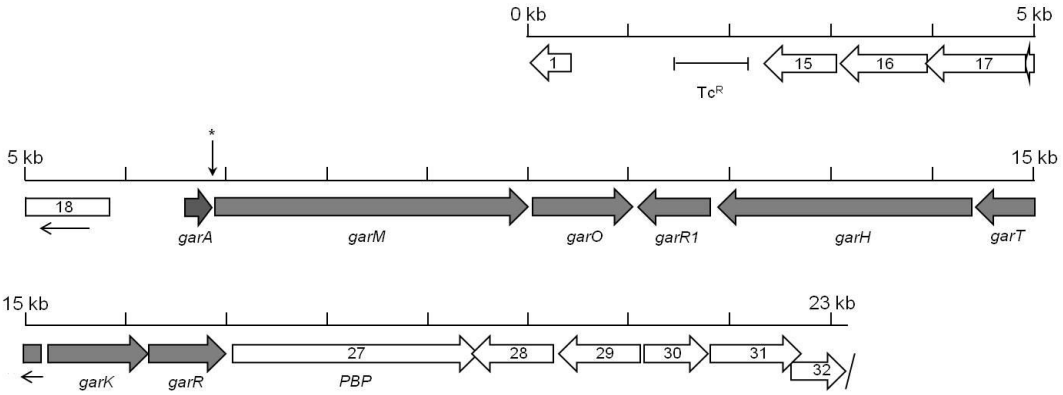
Computational analyses identified sets of conserved genes located downstream of *garA* and *ligA* in both clusters. As homology between the two putative lantibiotic gene clusters is high, it is likely that these genes would be essential for lantibiotic production. By comparing the two clusters it was immediately clear that the region of homology extended from the *lanA* genes to the *lanR* genes. The computational analysis also suggested a potential function for each identified ORF. The regions upstream of *lanA* in the AG14 and AL02 cosmids showed no significant sequence similarity (Boakes *et al.*, 2010), although both regions contain genes encoding ABC transporters. There is limited homology in the downstream region, where both clusters have genes encoding a putative membrane protein, a putative hydrolase and a response regulator in the subsequent four genes, one of those being *garPBP* in the AG14 cosmid. The relatively low level of similarity of these genes compared to those lying between the *lanA* and *lanR* genes suggested that they (with the possible exception of *garPBP*) were not likely to be part of the biosynthetic gene clusters, but this required experimental confirmation.

To identify the MGS for actagardine production, a tetracycline resistance cassette was used to replace segments of AG14 surrounding the proposed MGS. PCR-targeting was used to make these deletions. Essentially, 58/59 nt primers that corresponded to the DNA sequence adjacent to the genes to be deleted were used to amplify a targeting cassette containing a tetracycline resistance gene. Once this cassette was created using PCR, it was used to target cosmid DNA in *E. coli* BW25113 containing pIJ790 carrying the λ *red* genes that promote recombination. The three regions targeted are shown below (figure 4.6B, 4.6C and 4.6D). The first deletion (yielding cosmid AG14.2 – AG14.14) replaced all of the upstream genes except for the one located at the end of the cosmid insert and the four genes preceding *garA* with the tetracycline resistance cassette. The second deletion, yielding cosmid AG14.2-AG14.18, removed all upstream genes except the one located at the end of the cosmid insert. The third deletion (cosmid AG14.28-14.30) removed three of the genes downstream of *garPBP*. Each of the three deleted cosmids was transferred to *S. lividans* TK24. Figure 4.6E shows the subsequent bioassays. All deleted cosmids elicited antibiotic activity against *M. luteus* and were subsequently confirmed by MALDI-ToF to be producing actagardine (figure 4.6F). Thus the deleted genes were not required for actagardine production in *S. lividans* TK24, and do not appear to be part of the biosynthetic gene cluster.

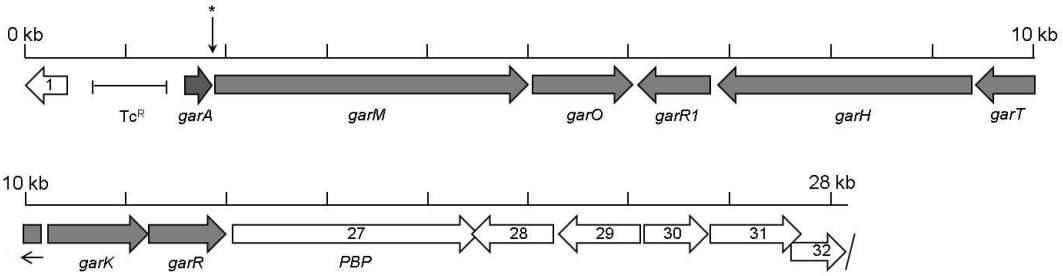
4.6A: Schematic of AG14 cosmid



4.6B: Schematic of cosmid AG14.2–14.14



4.6C: Schematic of cosmid AG14.2-14.18



4.6D: Schematic of cosmid AG14.28-30

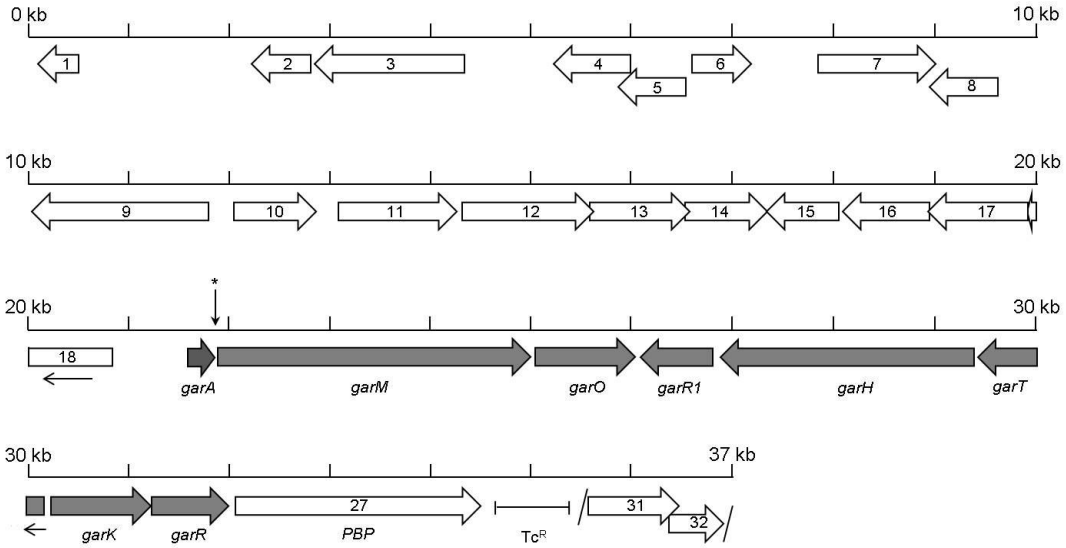
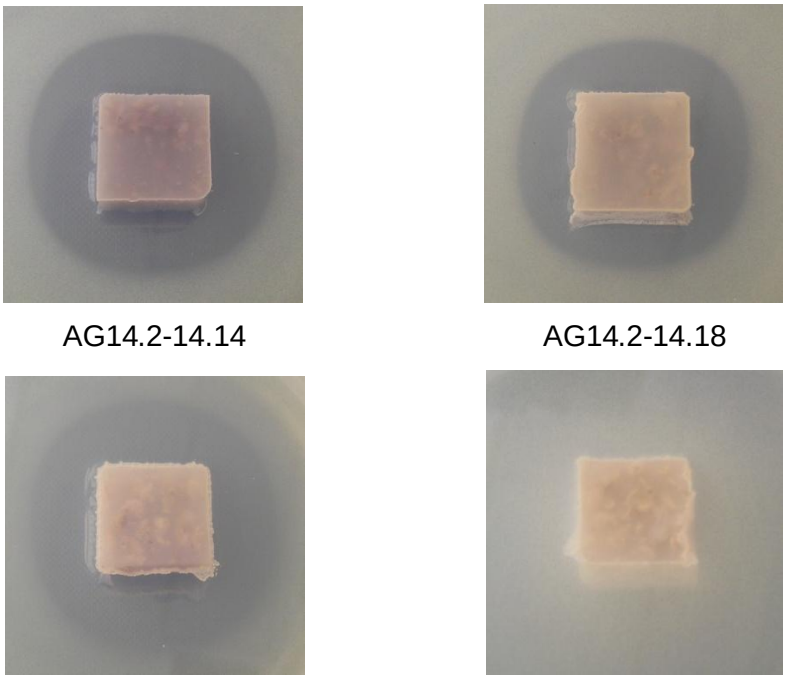


Figure 4.6: A-D. Schematics representing the replacement of genes with a tetracycline resistance targeting cassette. Using a PCR based targeting approach (chapter 2, section 2.8), the genes 2-14, 2-18 and 28-30 on the AG14 cosmid were replaced with the cassette, and the resulting clones tested for actagardine biosynthesis via bioassays (4.6E) and MALDI-ToF (4.6F).



AG14.2-14.14 AG14.2-14.18

AG14.28-14.30 pIJ10702

Figure 4.6E: Bioassays of *S. lividans* TK24 containing deletion derivatives of the AG14 cosmid: AG14.2-14.14, AG14.2-14.18 and AG14.28-14.30, and a vector only control (pIJ10702).

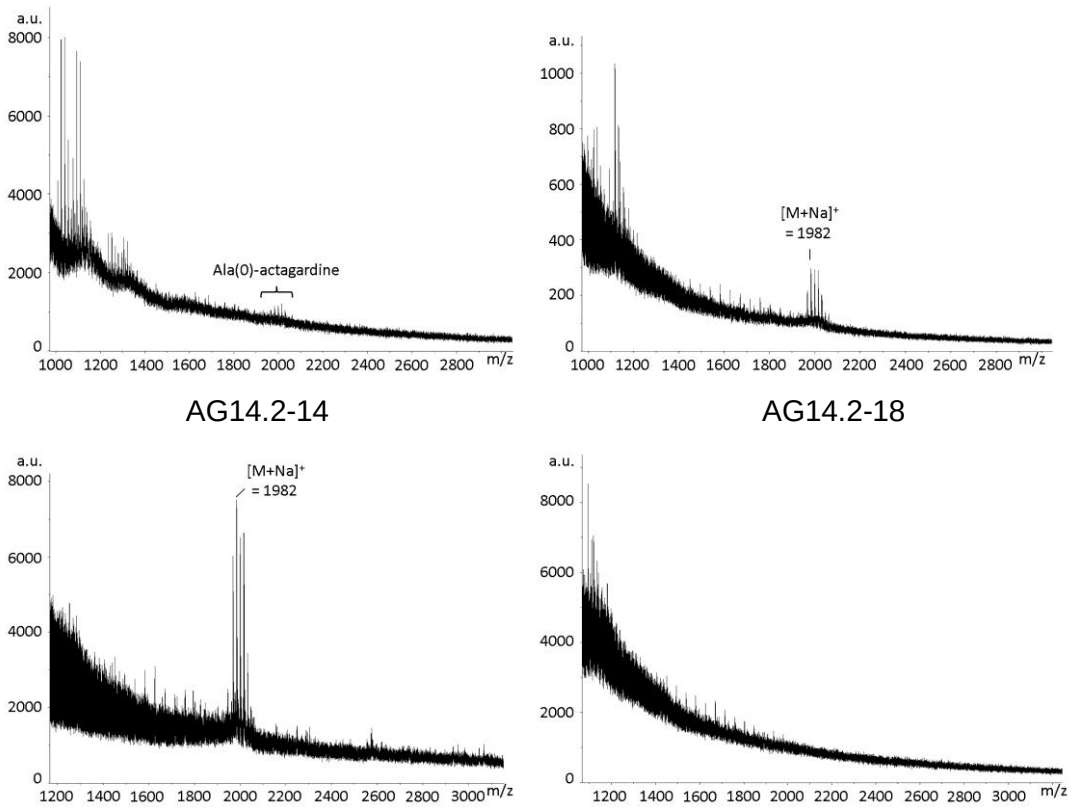


Figure 4.6F: MALDI-ToF spectra of *S. lividans* TK24 containing either the AG14.2-14, AG14.2-18, AG14.28-30 or pIJ10702 cosmids.

Since deletion of the genes upstream of *garA* in AG14.2-14.8 did not affect the production of actagardine, it appears that *garA* is the first essential gene in the actagardine biosynthetic gene cluster. Deletion of genes 28-30, potentially encoding proteins with functions analogous to genes found at a similar location in the DAB gene cluster, appear not to be involved in actagardine biosynthesis.

4.6 Deduction of a minimal gene set for deoxyactagardine B production

To deduce the minimal gene set for DAB production, cosmids from *A. liguriae* containing a different complement of genes to those found on the sequenced AL02 cosmid were assessed for their ability to produce DAB. Novacta Biosystems identified four cosmids (AL02, AL03, AL13 and AL17) by probing the *A. liguriae* genome library with a *garA* probe. All four cosmids were initially analysed by end-sequencing, using primers that anneal to the T3 and T7 bacteriophage promoters in SuperCos, and restriction enzyme analysis to identify a cosmid where the *ligA* gene was reasonably close the centre of the cosmid, which was then fully sequenced (Novacta Biosystems, unpublished). The three remaining end-sequenced cosmids all include some of the genes present within the sequenced and annotated AL02 cosmid. Figure 4.7A shows the location of the ends of the inserts present in AL03, AL13 and AL17 with respect to the insert present in AL02. The AL03 insert starts towards the end of ORF 9, and ends after ORF 31. The AL13 insert starts in ORF 3 and ends at ORF 31. The AL17 insert starts further upstream (to the left) of the AL02 cosmid insert and ends within ORF 29. AL03, AL13 and AL17 were supplied by Novacta Biosystems as SuperCos library clones and subsequently targeted using the *Ssp*I cassette (refer to chapter 2, section 2.8.4), transferred into *E. coli* ET12567 pUZ8002 and conjugated into *S. lividans* TK24. The subsequent clones were analysed by bioassay and MALDI-ToF. Figure 4.7B shows the results of the bioassay. All four cosmids elicited zones of inhibition in lawns of *M. luteus* and the production of DAB was confirmed by MALDI-ToF analysis (figure 4.7C). The most important cosmid to note is AL03, whose insert starts just two genes upstream of *ligA*, demonstrating that the genes upstream of ORF 10 are not required for DAB biosynthesis, analogous to the situation in cosmid AG14.2-14.18 cosmid. While this eliminates most of the upstream genes in the cosmid, this comparative analysis was not able to eliminate the involvement of genes downstream of *ligR* gene. However, as stated above, the lack of strong sequence homology downstream of *ligR* when compared to AG14 suggested that these genes were not involved in DAB biosynthesis (Boakes *et al.*, 2010).

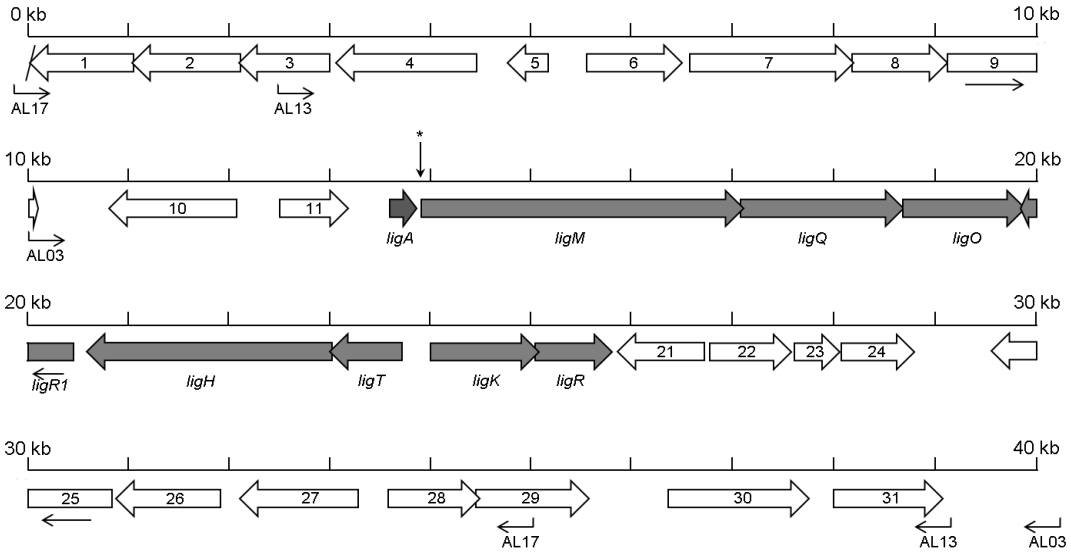
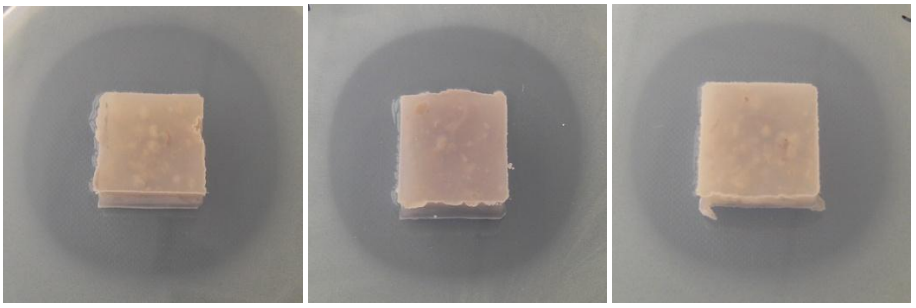
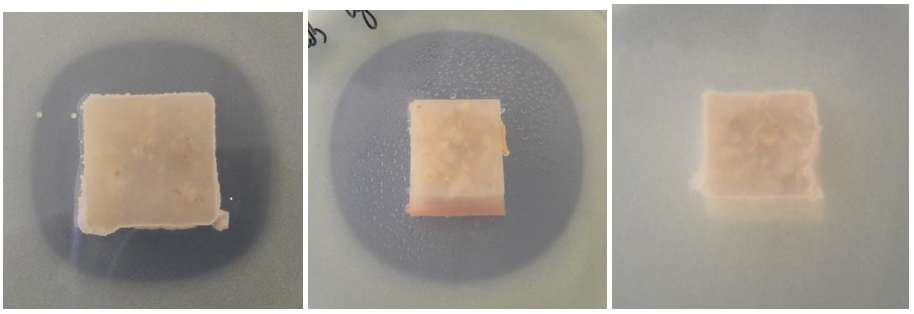


Figure 4.7A: Schematic of the AL02 cosmid showing the positions of the ends of the inserts cloned in AL03, AL13 and AL17.

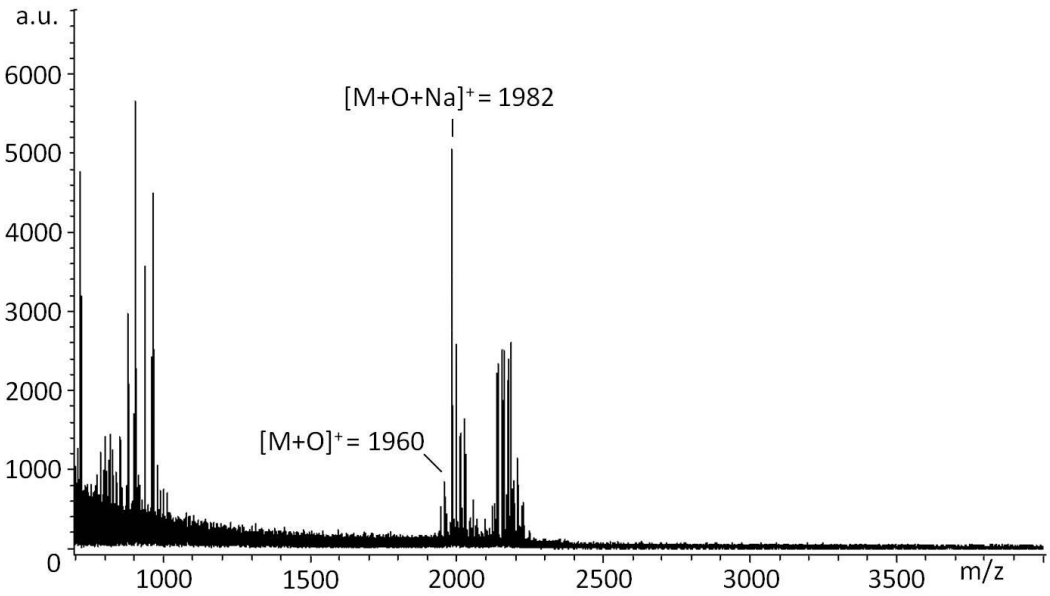


M1202 (AL02) M1203 (AL03) M1204 (AL13)

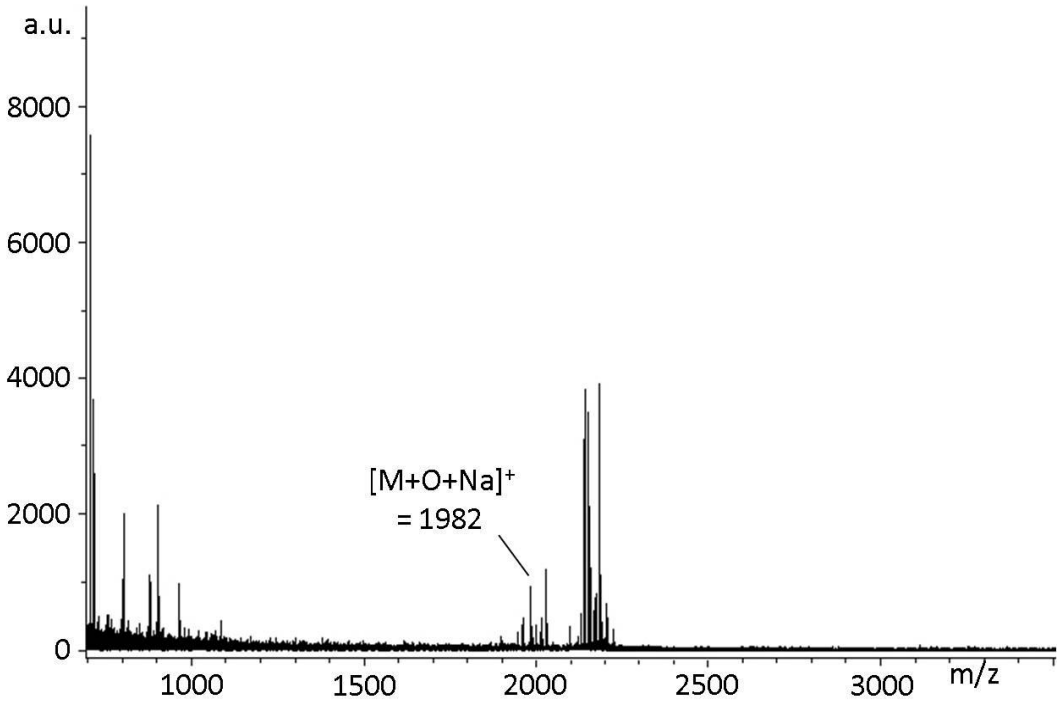


M1205 (AL17) *A. liguriae* M1216 (pIJ10702)

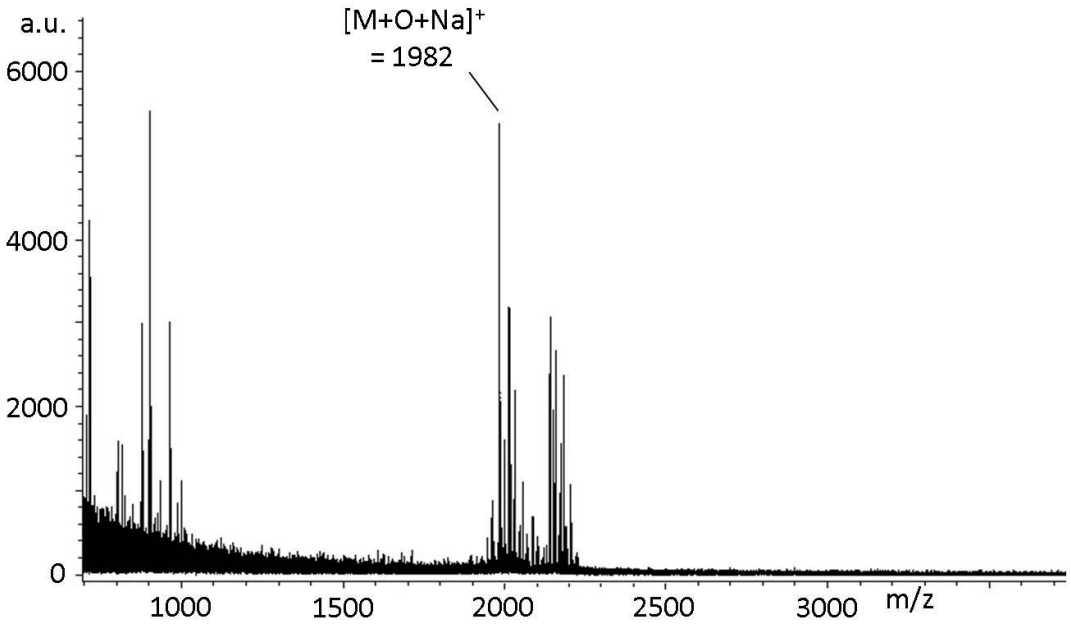
Figure 4.7B: Bioassay of *S. lividans* TK24 containing the different *A. liguriae* cosmids. Agar cubes inoculated with each clone were incubated for 7 days and then placed on a L agar plate pre-seeded with *M. luteus* and incubated for a further 24 hours. Zones of inhibition can be seen around all of the cosmid clones.



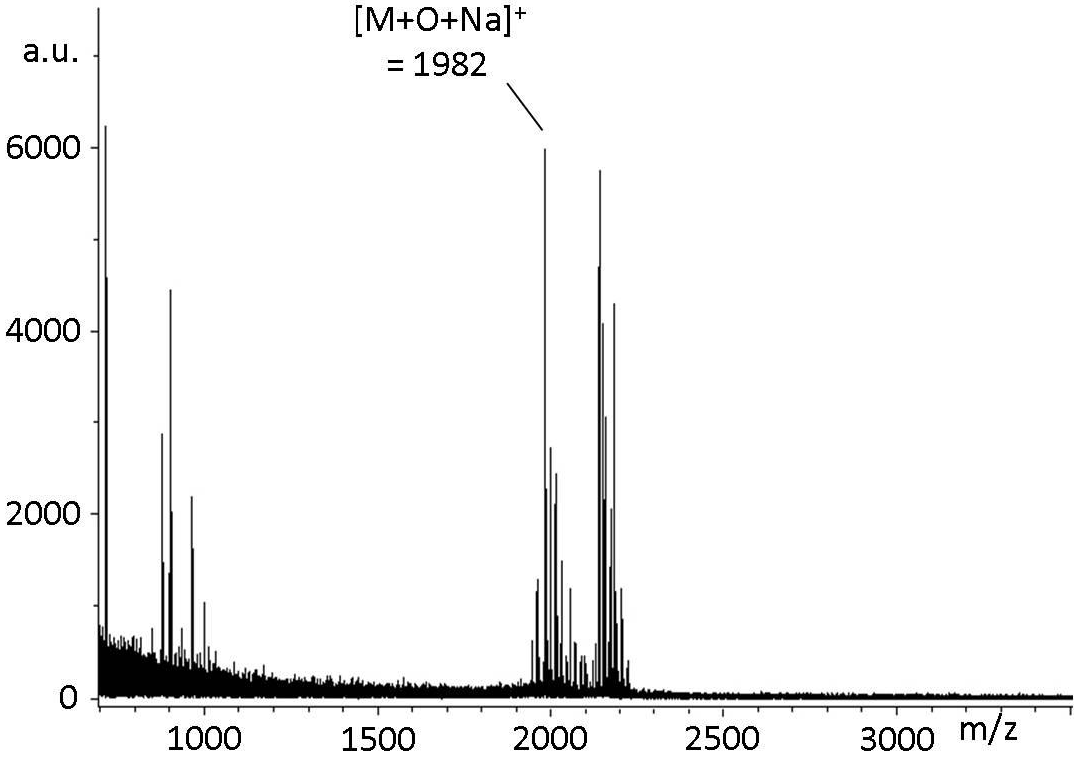
M1202 (AL02)



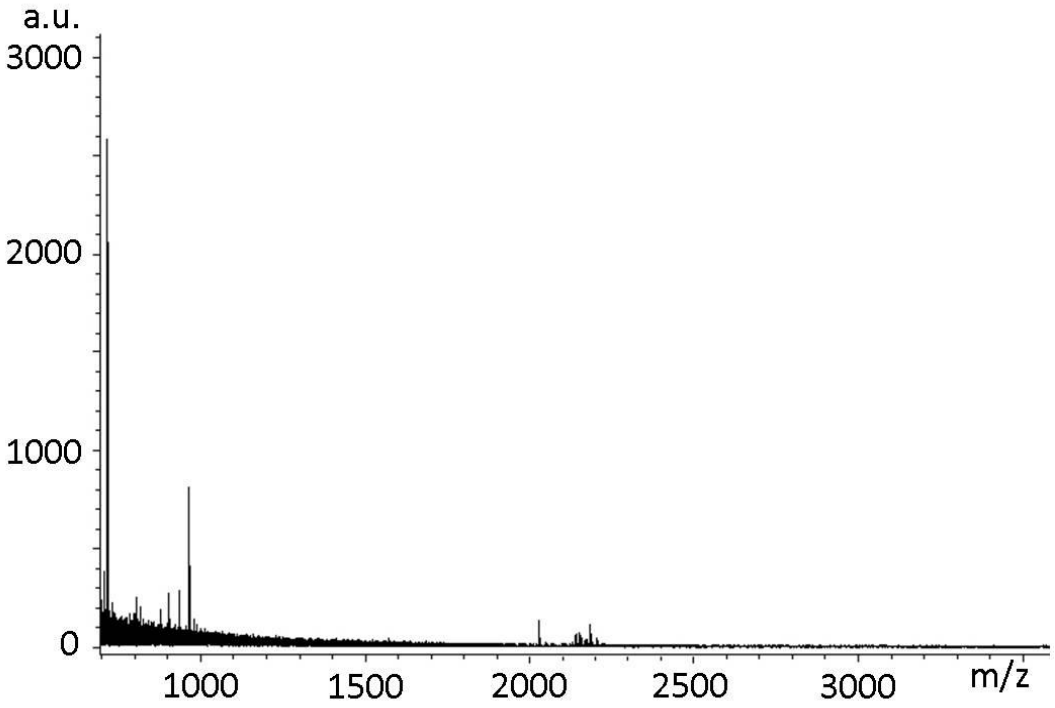
M1203 (AL03)



M1204 (AL13)



M1205 (AL17)



M1216 (pIJ10702)

Figure 4.7C: MALDI-ToF spectra of culture supernatants of five *S. lividans* TK24 clones. M1202 (AL02), M1203 (AL03), M1204 (AL13) and M1205 (AL17) produce Ala(0)-deoxyactagardine B ($[M+O+Na]^+$ = 1980), M1216 does not.

Figure 4.7: Schematic of the AL02 cosmid showing the positions of the ends of the inserts cloned in AL03, AL13 and AL17. Bioassay of *S. lividans* TK24 containing the different *A. liguriae* cosmids coupled with MALDI-ToF analysis concluded that all cosmids produced Ala(0)-DAB.

4.7 Discussion

In this chapter the limitations imposed by attempting to work with the genetically uncharacterised natural producers *A. garbadinensis* and *A. liguriae* were circumvented by expressing the actagardine and DAB biosynthetic gene clusters in two different *Streptomyces species*. Expression of the AG14 cosmid was successful, confirming that this cosmid contained all of the genes required for actagardine biosynthesis, as also shown by Boakes *et al.* (2009) (Boakes *et al.*, 2009). *S. lividans* TK24 was also used in this work to express the AL02 cosmid and for the first time demonstrated that this cosmid contained all of the genes required for DAB production. Like the natural producer, all of the *A. liguriae* cosmids conferred production of a deoxy variant of actagardine, potentially caused by a defective monooxygenase.

S. lividans 1326 is the heterologous host routinely used at Novacta Biosystems. It and the *S. lividans* TK24 and *S. coelicolor* M1146 strains analysed in this study produced ala(0) variants of actagardine and DAB, unlike the natural producers (Boakes *et al.*, 2010, Boakes *et al.*, 2009). The existence of an ala(0) variant could indicate that cleavage of the leader sequence occurs in a two step process. The first step would remove the leader sequence up to the ala(0) position, followed by a second cleavage that removed this residue to leave the mature peptide. The latter of these activities would be absent from the *Streptomyces* species used as heterologous hosts, resulting in production of the ala(0) variants. Alternatively, cleavage of the prepropeptide in the natural producers and heterologous hosts may be carried out a single protease with slightly different specificity. Given that neither gene cluster contains a gene encoding a protease or protease domain, cleavage of the leader sequences from the prepropeptides is presumably performed by protease(s) encoded elsewhere in the genomes of the producing species. This is not without precedent in lantibiotic biosynthesis (Foulston & Bibb, 2010, Cooper *et al.*, 2008).

Heterologous expression demonstrated that the transcription and translation machinery of *S. lividans* and *S. coelicolor* is able to recognise the *Actinoplanes* regulatory sequences required for actagardine and DAB biosynthesis. This had important consequences since it enabled the use of a genetically amenable

heterologous host for the subsequent mutational analysis of the actagardine biosynthetic gene cluster. While in principle this could have been attempted in *A. garbadinensis* (see Boakes *et al.* (2009)), the two *Actinoplanes* species used in this work grow more slowly than *S. lividans* and *S. coelicolor* and are genetically less tractable. Moreover, the use of *Streptomyces* species as heterologous hosts made available a range of vectors and techniques that could be used to characterise actagardine and DAB biosynthesis.

Heterologous expression of the DAB gene cluster also potentially allowed for a functional comparison of the two clusters to be made. For example, deletion of the *A. liguriae*-specific gene *ligQ* might allow us to understand its apparently unique role in DAB biosynthesis. Analysis of the apparently dysfunctional monooxygenase could also be undertaken.

A MGS was established for the production of actagardine by deleting many of the genes present on the AG14 cosmid. The first gene in the cluster is *garA* with the last being either *garR* or *garPBP.* All other genes on the AG14 cosmid appeared to be non-essential for actagardine biosynthesis. The high level of homology shared by many of the genes present on both the AG14 and AL02 cosmids also helps to define a likely MGS for DAB biosynthesis. Comparison of the genes present in the different *A. liguriae* cosmids provided by Novacta Biosystems allowed us to eliminate a role for many of those carried on AL02 in DAB biosynthesis. Further experimental analysis to more precisely define the MGS for DAB was not carried out since it appeared likely that only genes with a high level of conservation to *gar* genes were likely to be required. Consistent with levels of amino acid sequence similarity, Boakes *et al.* (2010) noted that there is a distinct drop in nucleotide sequence homology between the two cosmids in the regions preceding the *lanA* and following the *lanR* genes. It falls from 65% between *lanA* and *lanR* (rising to 76% if *ligQ* is discounted) to 30-42% in the flanking regions (Boakes *et al.*, 2010).

Expression of AG14 in *S. lividans* TK24, and the establishment of a minimal gene set, allowed the use of PCR-targeting to create deletions in each of the proposed

biosynthetic genes for actagardine biosynthesis. These studies are reported in chapter 5.

4.8 Summary

- Actagardine was produced heterologously when cosmid AG14 was introduced into *S. lividans* TK24 and *S. coelicolor* M1146.

- *S. coelicolor* M1146, from which four secondary metabolite gene clusters had been deleted, appeared to produce increased levels of actagardine compared with *S. lividans* TK24.

- For the first time, this study demonstrated that all of the genes required for DAB biosynthesis were contained in cosmid AL02.

- Deletion analysis of the AG14 cosmid was used to define a likely MGS for actagardine biosynthesis.

- Comparison of the genes contained in three other *A. liguriae* cosmids that conferred DAB production when introduced into *S. lividans* TK24 identified a reduced gene set for DAB biosynthesis.

- The experimentally defined and deduced MGSs for actagardine and DAB biosynthesis correspond to regions of AG14 and AL02 with high levels of nucleotide sequence identity that drop off significantly outside of the proposed MGSs.

5 Mutational analysis of actagardine biosynthesis and analysis of novel actagardine homologs

5.1 Introduction

In chapter 4, heterologous expression of both the AG14 and AL02 cosmids in *Streptomyces* hosts confirmed that they contained all of the genes required for production of actagardine and DAB, respectively. Computational analysis carried out in chapter 3 identified a number of putative lantibiotic biosynthesis genes, based on their homology to known examples. Chapter 4 also identified a MGS for the AG14 cosmid and confirmed that eight or nine genes encode all of the necessary components required for actagardine biosynthesis, including the prepropeptide, and enzymes involved in modification, transport, regulation and potential immunity to actagardine. Nine genes were proposed to be involved in DAB biosynthesis based on homology to characterised lantibiotic biosynthetic genes and the high degree of homology shared between these genes to those in the AG14 MGS.

Computational analysis identified genes that were unique to each cluster, and that had not been found in previously characterised lantibiotic gene clusters. *garPBP* and *ligQ* have no homologs in their opposing biosynthetic gene cluster, although in principle they could perform a similar function in providing producer immunity against the corresponding lantibiotic. The LanO monooxygenases are thought to be capable, in principle, of adding an oxygen moiety to the C-terminal methlylanthionine bridge (although *ligO* appears non-functional), a modification that has not been seen in any other lantibiotic. Deletion of these genes would determine whether they were essential for biosynthesis or for the bioactivity of their respective lantibiotics.

While computational analysis is able to assign a potential function to each of the ORFs, it is unable to conclusively demonstrate that the gene is essential for lantibiotic biosynthesis. The role of each of these genes in the biosynthesis of actagardine and DAB required experimental confirmation.

Many lantibiotic gene clusters such as those for nisin, the most studied lantibiotic, and microbisporicin, a lantibiotic produced by an actinomycete, have undergone numerous gene deletion experiments to provide increased understanding of the role each gene plays in the biosynthesis of its respective lantibiotic (van den Berg van Saparoea *et al.*, 2008, Foulston & Bibb, 2010). To determine the role of each of the genes proposed to be involved in actagardine and DAB biosynthesis, attempts were made to delete each individually from the corresponding cosmid. The effects of these gene deletions were assessed using bioassays and MALDI-ToF mass spectroscopy.

5.2 Construction, confirmation and analysis of gene deletions within the biosynthetic gene clusters of actagardine and DAB

As mentioned previously, deletion of each putative biosynthetic gene will determine whether the protein it encodes is essential for lantibiotic biosynthesis. The computational analysis carried out in chapter 3 demonstrated that nearly all of the actagardine and DAB biosynthesis genes have a homolog in the other cluster and that the homology between them is very high (see figure 5.1). This observation indicates that homologous genes found in both clusters are likely to perform the same function within the biosynthetic pathways of these lantibiotics.

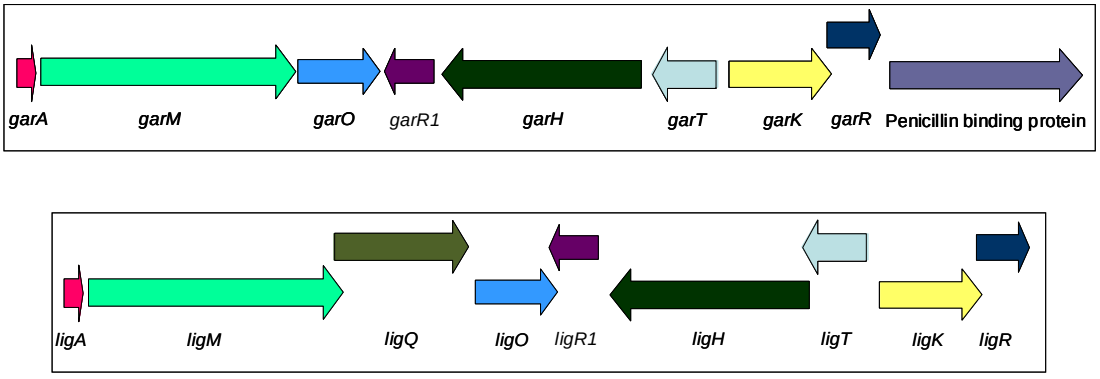


Figure 5.1: A diagram of the biosynthetic gene clusters from both *A. garbadinensis* (AG14, top) and *A. liguriae* (AL02, bottom). Genes that are predicted to perform the same function in both clusters are represented in the same colour. Genes that are found in only one cluster are uniquely coloured.

With this in mind, the results from a deletion series of all the biosynthetic genes from the actagardine MGS on the AG14 cosmid are highly likely to apply to the homologous genes present on the AL02 cosmid. This principle, however, cannot be applied to genes that only appear in either one of the clusters, these being *garPBP* and *ligQ.* While it is possible that these two genes play equivalent roles in host immunity, they encode two quite different proteins, one a transporter, the other a penicillin binding protein. In this study, *ligQ* is the only gene within the AL02 cluster that was deleted; *garPBP* was deleted in conjunction with the other putative actagardine biosynthetic genes.

Many of the putative biosynthetic genes in the actagardine MGS have homologues in the gene clusters of other Class II lantibiotics, for example *garA, garM*, and *garTH* (Chatterjee *et al.*, 2005). These were thus assumed to be essential for actagardine biosynthesis, and deletion of these genes would likely result in the loss of actagardine biosynthesis. Computational analysis carried out in chapter 3 also identified genes that are only found in the two *Actinoplanes* clusters, and are not found in other Class II gene clusters. The *lanO* genes, encoding putative monooxygenases, are unique to these clusters and have the potential to perform a novel lantibiotic modification; enzymatic addition of an oxygen moiety to the C-terminal methylanthionine bridge of the lantibiotics. However, the existence of DAB suggests that the LigO monooxygenase is either dysfunctional, or unable to recognise its peptide substrate, which differs from that of actagardine.

The biosynthetic genes present in the AG14 and AL02 gene clusters appear to be organised into several operons, as outlined in chapter 3 (and represented again in figure 5.1). These consist of *garA* to *garO,* which are likely to be translationally coupled, *garT* to *garR1,* and *garK* to *garR*. Any gene deletions that are made within the cluster must not interfere with the expression of downstream genes or the results would not reflect a single gene deletion. To accomplish this, in-frame deletions were made that replaced essentially all of the coding sequence of the targeted gene with an 81 bp scar sequence. These scar sequences retain the reading frame in the deleted gene and lack stop codons, and so should not interfere with the expression of downstream genes (figure 5.2). Promoters and ribosome binding sites are unaffected by the 81 bp insertion. The only potential issue is that the deletions might affect the stability of the resulting multi-cistronic mRNA, leading to phenotypic effects not directly attributable to deletion of the targeted gene.

The deletion mutants were produced using a PCR-based targeting method (also known as REDIRECT) that first replaces the target gene with a selectable marker that is subsequently converted into an 81 bp scar (Gust *et al.*, 2003). Using the REDIRECT protocol, primers that match the chromosomal DNA sequence adjacent to the gene that is to be deleted were used to amplify a targeting cassette from pIJ799 (*aac(3)IV* ($Apra^R$) + *oriT*) containing the apramycin resistance gene and the origin of transfer, both of which are flanked by FRT sites (Gust *et al.*, 2003). The FRT sites are small segments of DNA that are recognised by the yeast FLP recombinase encoded by the BT340 plasmid (Datsenko & Wanner, 2000). Incubation at 42^0C activates expression of the FLP recombinase that selectively recombines the two flanking FRT sites, leaving the 81 bp in-frame non-polar deletion in place of the targeting cassette (Cherepanov & Wackernagel, 1995). This strategy also allows the subsequent use of the apramycin resistance gene as a marker on the targeted cosmid, allowing for apramycin selection when conjugating into *Streptomyces.* This was accomplished by targeting the SuperCos backbone of the scarred cosmids with an *Ssp*I restriction fragment from pIJ10702 (also known as pMJCos[1]) that contained the apramycin resistance gene, *oriT*, ΦC31 integrase gene and *attP* attachment site.

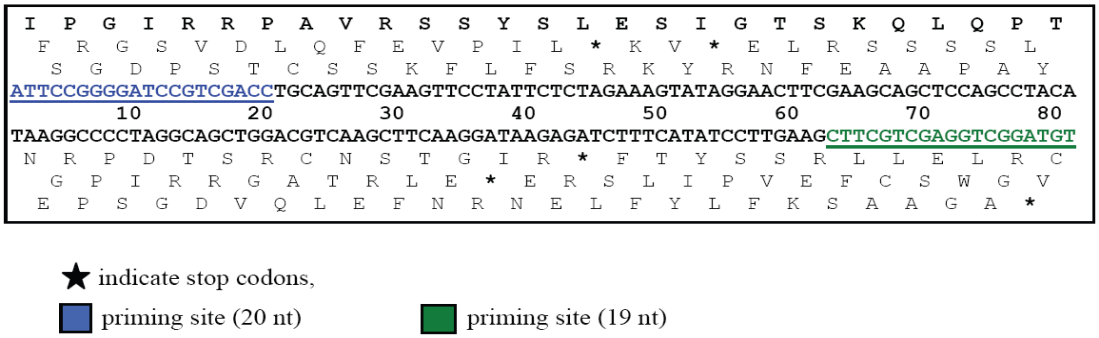


Figure 5.2: Sequence of the 81 bp scar sequence remaining after FLP-mediated excision of the apramycin cassette. Translation of the targeted reading frame is printed in bold. The 20 and 19 nt priming sites used in PCR amplification are underlined and printed in colour (adapted from Recombineering in *Streptomyces*, Gust *et al.* (2006)).

All scar mutants were made in the AG14 SuperCos (pIJ12201) cosmid and confirmed in *E. coli* using both restriction digests and PCR analysis (figure 5.3). Confirmed clones were subsequently targeted with the *Ssp*I fragment, introduced into *S. lividans* TK24 by conjugation and validated again by colony PCR (chapter 2, table 2.1). Bioassays were performed by incubating the *S. lividans* TK24 exconjugants on OB-N agar plugs for seven days. An agar plug was subsequently

transferred onto a pre-inoculated L agar plate and incubated for a further 24 hours. All phenotypes were confirmed using MALDI-ToF mass spectrum analysis of supernatants obtained from liquid cultures.

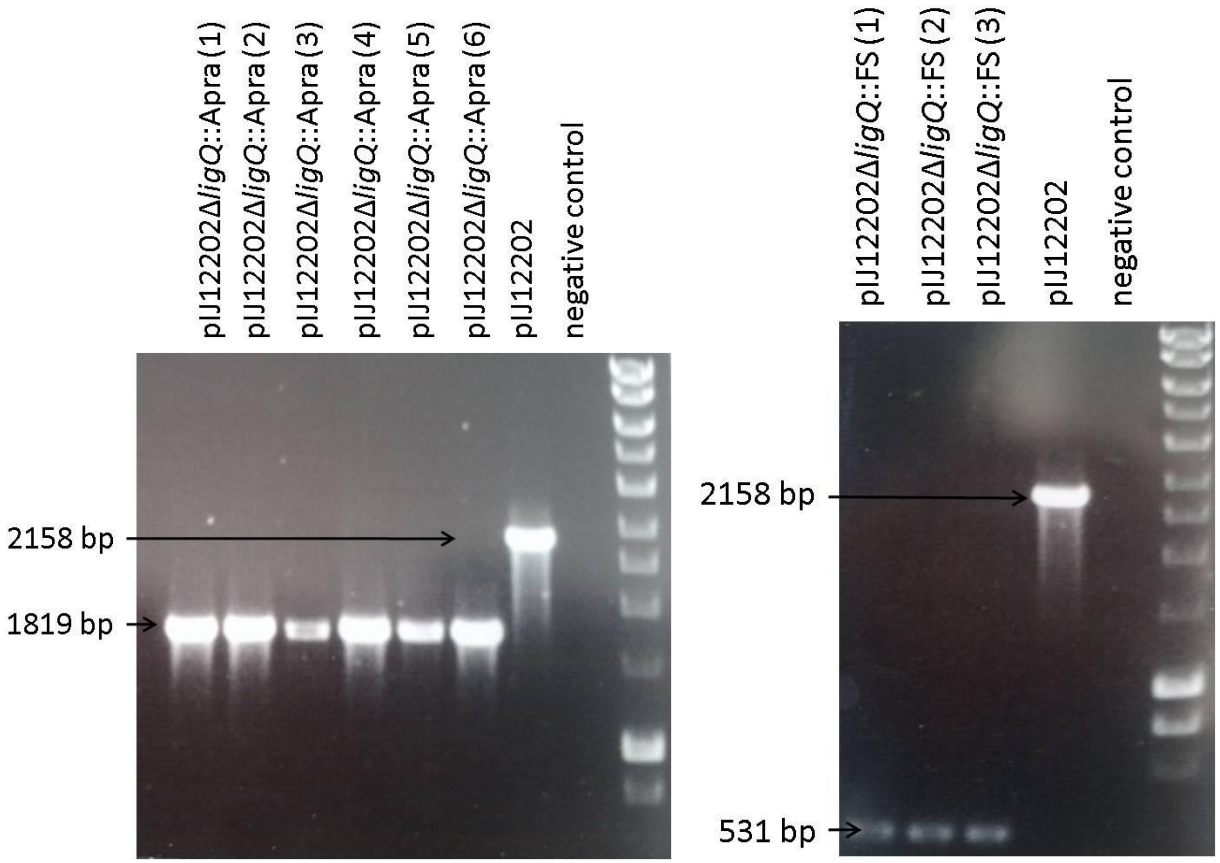


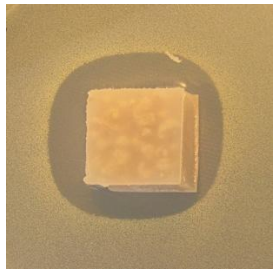
Figure 5.3: Representative PCR confirmation of the replacement of *ligQ* with the apramycin resistance cassette (left), followed by the replacement of the cassette with the 81 bp scar (right) using the lan-con3 primers in *E. coli* DH5α clones. On the left, the pIJ12202 clone amplified a 2158 bp fragment encompassing the wild type gene. The smaller 1819 bp fragment demonstrated that six clones have successfully had the gene replaced by the cassette. On the right, three clones have had the apramycin cassette replaced with the 81 bp scar sequence (513 bp).

Streptomyces clones that contained a deletion that resulted in the loss of production of actagardine (or DAB) were subsequently subjected to attempted complementation with an in-*trans* chromosomally integrated (ΦBT1 *attB* site) copy of the deleted gene using the *Streptomyces* constitutive *ermE** promoter on pIJ10257 (Hong *et al.*, 2005). This was done in an attempt to confirm that the gene deletion was indeed responsible for loss of production, and that it did not reflect an effect on downstream genes resulting from, for example, altered transcript stability, or a spontaneous mutation elsewhere in the genome. The complementation constructs were produced by PCR amplification. Primers with *Nde*I and *Hind*III restriction sites incorporated into the 3‘ ends of the forward and reverse primer sequences were used to amplify the gene to be complemented. The gene was amplified from the ATG start codon onwards since pIJ10257 contains both promoter and RBS elements. After amplification, the PCR products were cut with *Nde*I and *Hind*III and subsequently cloned into the pGEM T-easy vector (Promega) and introduced into *E. coli* DH5α (Invitrogen) by transformation. The resulting constructs were digested with *Nde*I and *Hind*III, the amplified gene subjected to gel purification, and ligated with the similarly cut expression vector pIJ10257. All constructs were sequenced to verify that the genes had been cloned without errors. The constructs were subsequently introduced into *E. coli* S17-1 by transformation, conjugated into *S. lividans* TK24 and subjected to bioassay and MALDI-ToF mass spectroscopy analysis.

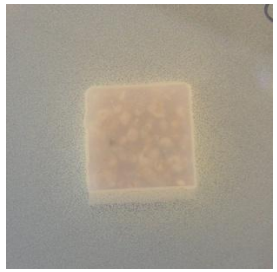
Figure 5.4A contains images of all the *S. lividans* TK24 clones that contain a deletion in a single putative lantibiotic biosynthetic gene in the AG14 MGS, and in *ligQ* from the AL02 cosmid. Lack of lantibiotic biosynthesis is indicated by loss of a zone of inhibition around the agar plug. Further information on each specific gene deletion is detailed in the subsequent text.

Figure 5.4B contains images of the bioassay results of *S. lividans* TK24 clones containing the complementation constructs of single gene deletion mutants that resulted in a loss of production phenotype. The *garO* mutant was included even though the clone produced a zone of inhibition. This complemented clone should revert to producing actagardine, which is detectable using MALDI-ToF mass spectrum analysis. However, in the majority of cases, attempts to complement the deleted gene failed to restore actagardine biosynthesis. MALDI-ToF mass spectroscopy of culture supernatants also failed to reveal restoration of actagardine

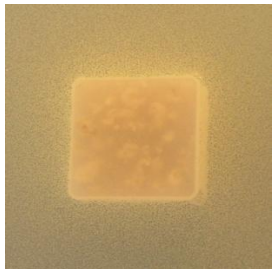
production, even when a zone of inhibition was present (i.e. the complemented *garR* mutant).



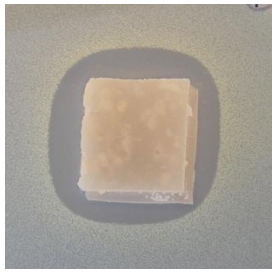
AG14



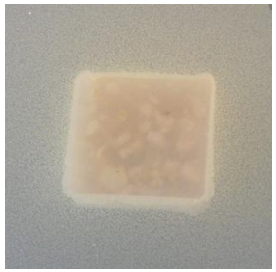
AG14Δ*garA*



AG14Δ*garM*

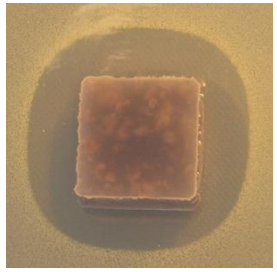


AG14Δ*garO*

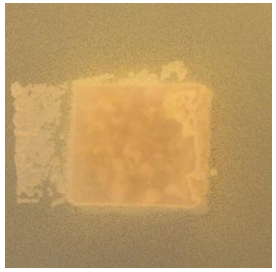


AG14Δ*garR1*

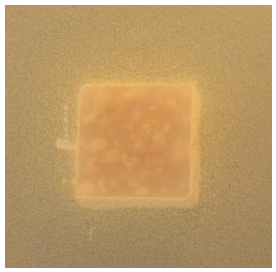
AG14Δ*garH*



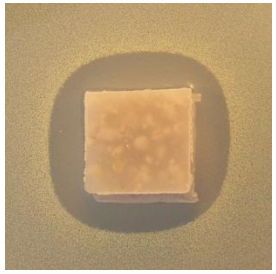
AG14Δ*garT*



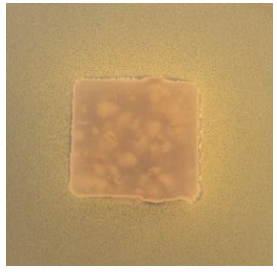
AG14Δ*garK*



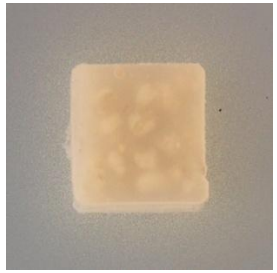
AG14Δ*garR*



AG14Δ*garPBP*

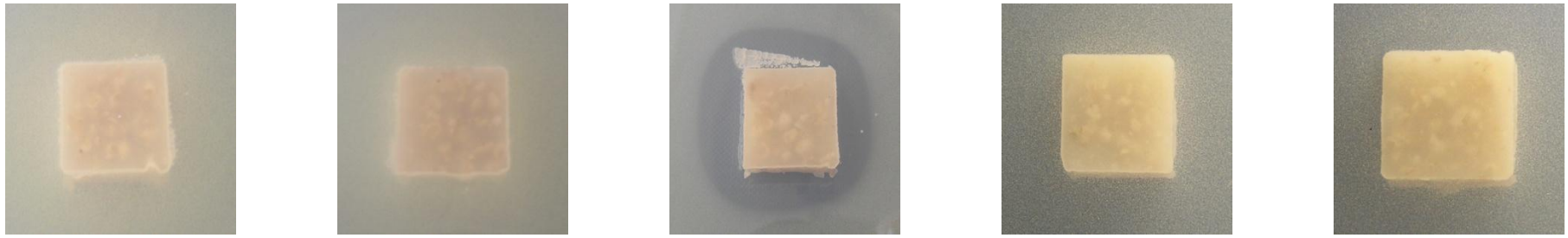


AL02Δ*ligQ*

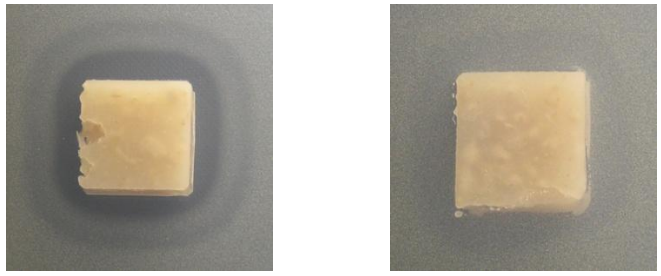


pMJCos^{-1}

Figure 5.4A: Images of *S. lividans* TK24 clones that contain a single deletion of a putative lantibiotic biosynthetic gene within the AG14 and AL02 cosmids. Clones were incubated for seven days on OB-N agar cubes and subsequently placed on L agar plates pre-seeded with the sensitive test organism *M. luteus* and incubated for a further 24 hours. Zones of inhibition can be seen around clones where the gene deletion has not abolished biosynthesis of the lantibiotic. Clones where *M. luteus* has grown close to the agar plug demonstrate that the *S. lividans* derivative is no longer producing a lantibiotic.



garA *garM* *garO* *garR1* *garK*



garR *ligQ*

Figure 5.4B: Images of *S. lividans* TK24 clones that contain a single deletion of a putative lantibiotic biosynthetic gene within the AG14 and AL02 cosmids and that have been subjected to attempted complementation with an *in-trans* integrated copy of the gene under control of the *ermE** promoter. Clones were incubated for seven days on OB-N agar cubes and subsequently placed on L agar plates pre-seeded with the sensitive test organism *M. luteus* and incubated for a further 24 hours. A zone of inhibition around the *garR* mutant indicated successful complementation (the zone has suffered from re-growth of *M. luteus* leading to a double ring). Clones where *M. luteus* has grow close to the agar plug indicate that the mutant has not been complemented and that biosynthesis has not been restored.

The first gene in the actagardine biosynthetic gene cluster is *garA* that encodes the prepropeptide. GarA is the precursor of mature actagardine, and is modified and exported by other enzymes and proteins encoded within the MGS. Figure 5.4A contains an image of the *S. lividans* TK24 clone containing the cosmid pIJ12206Δ*garA*. This clone notably lacks a halo around the agar plug after seven days growth on OB-N agar, in contrast to the AG14 cosmid (positive control) that does have a zone of inhibition, suggesting that the AG14Δ*garA* clone is no longer able to produce actagardine. MALDI-ToF mass spectrum analysis of supernatants from liquid cultures of the strain failed to detect peaks corresponding in size to the known masses of actagardine, providing further evidence that deletion of *garA,* as expected, completely abolished actagardine biosynthesis.

Actagardine, like all lantibiotics, is ribosomally synthesised. The sequence of the *lanA* structural gene correlates directly with the peptide sequence of the mature lantibiotic, with the exception of the amino acids that are modified to form (methyl-) lanthionine bridges. The actagardine MGS is capable of producing a number of variant actagardine molecules (Boakes *et al.*, 2009). Some lantibiotic biosynthesis pathways can tolerate significant amino acid substitutions and truncations within the prepropeptide; for example, the nisin biosynthetic machinery can produce hexapeptides and non-lantibiotic peptides when they are fused to the nisin leader sequence (Kluskens *et al.*, 2005, Kuipers *et al.*, 2008, Rink *et al.*, 2007). The nisin biosynthesis machinery can even tolerate lantibiotic prepropeptides from different lantibiotic classes, including a two-component lantibiotic from *Streptococcus pneumoniae* (Majchrzykiewicz *et al.*, 2010). Loss of actagardine biosynthesis, and the failure to detect any other modified peptides in culture supernatants of the *garA* mutant, indicates that either *garA* is the only *lanA* gene in the heterologous host that encodes a substrate for the actagardine biosynthetic enzymes, or that deletion of *garA* has compromised the function of other *gar* genes in the cluster.

The *garA* prepropeptide has also been disrupted in the natural producer *A. garbadinensis*. Boakes *et al.* (2009) disrupted *garA* using an apramycin resistance gene (from pIJ773) and observed a similar loss of actagardine production. Although not attempted, deletion of *ligA* from either the natural producer or from the AL02 cosmid would be similarly expected to result in the loss of DAB production.

An attempt was subsequently made to complement the Δ*garA* mutant with pIJ10257-*garA* (figure 5.4B). Several repeats of the bioassay and MALDI-ToF

analysis of culture supernatants failed to reveal complementation. The reasons for this are unknown, particularly since in other studies a complementation construct was used to generate a number of actagardine variants in the natural producer and to restore limited production of actagardine in an *A. garbadinensis* Δ*garA* mutant (Boakes *et al.*, 2009). pIJ10257 containing the *ermE** promoter has been used previously to successfully complement mutants and to over-express various genes in a number of *Streptomyces* strains (McArthur & Bibb, 2008, Claesen & Bibb, 2010, Hong *et al.*, 2005). PCR and DNA sequence analysis confirmed that the complementation construct used in this experiment was assembled correctly.

There are several possible explanations for the lack of complementation. Boakes *et al.* (2009) commented on the low levels of actagardine production in an *A. garbadinensis* Δ*garA* mutant that had been complemented with *garA* using its own promoter and inserted at the ΦC31 integration site. Reduced levels of production of secondary metabolites when their corresponding gene clusters have been inserted at the ΦC31 integration site have been reported (Baltz, 1998, Rodriguez *et al.*, 2003). Boakes *et al.* (2009) suggested that the low levels of complementation were associated with *in trans* expression and likely to result from instability and loss of integrity of mRNA transcripts. It is therefore possible that integration of *garA* gene into the ΦBT1 attachment site used in these experiments has an even more severe effect, resulting in lack of restoration of detectable actagardine biosynthesis. The chromosomal context of *garA* in the complementation clone may also have an adverse affect on *garA* expression. For example, local “chromatin” structure may play a role. There may also be regulatory elements in the region of DNA directly upstream of *garA* that are crucial for its expression and that are lacking in the complementation construct. Like that of many secondary metabolites, the production of actagardine is regulated and occurs late in growth in liquid culture, and is probably coordinated with morphological differentiation in agar-grown cultures. Attempting to express *garA in trans* might have disrupted such regulatory mechanisms, resulting in lack of complementation.

The constitutive expression of *garA* from the strong *ermE** promoter might also interfere with regulatory mechanisms. It is conceivable that the stoichiometry between GarA and the biosynthetic machinery is important. The presence of excess actagardine prepropeptide might down-regulate expression of the rest of the cluster, or disrupt its function.

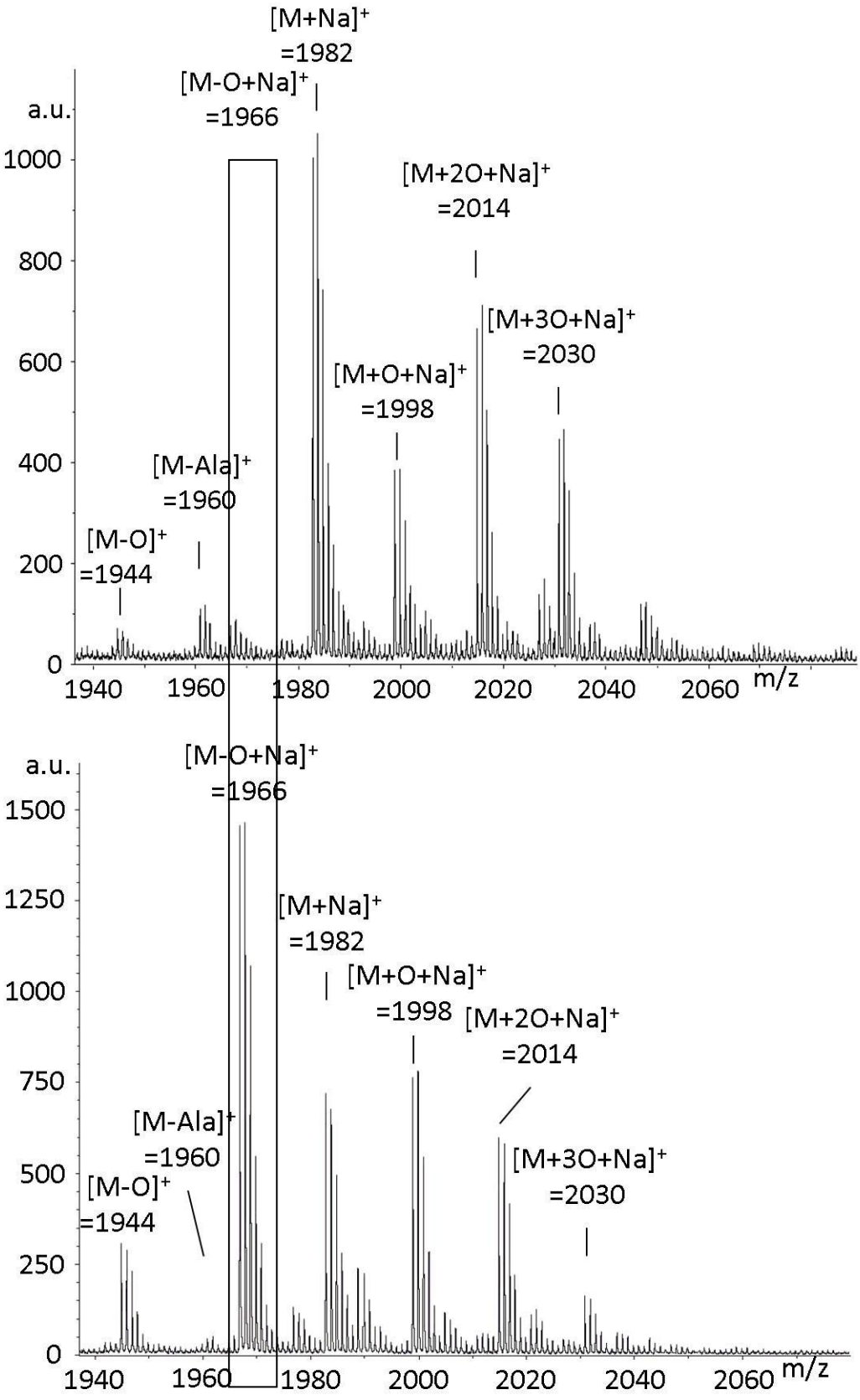
To try to identify the cause of the lack of complementation, Reverse Transcription-PCR (RT-PCR) was performed on RNA isolated from *S. lividans* TK24 clones containing the wild-type cosmid and a number of deletion mutants. RNA isolated after 3 days growth failed to yield RT-PCR products, perhaps reflecting a temporal delay in *gar* gene expression in the heterologous host. RNA extracted from cultures harvested at later time points were of poor quality, and again no RT-PCR products were detected. These experiments failed to provide any useful insights into the lack of genetic complementation.

garM was identified by bioinformatic analysis to encode a biosynthetic protein responsible for the selective dehydration of serine and threonine residues to Dha and Dhb, respectively. It is also responsible for cyclisation of the modified peptide by the addition of cysteine thiols to the dehydrated amino acid residues (You & van der Donk, 2007, Paul *et al.*, 2007). *garM* lies directly downstream of *garA* apparently in the same operon, and the two genes are separated by a stem-loop structure (Boakes *et al.*, 2009). Replacement of *garM* in the AG14 cosmid with the 81 bp scar sequence abolished actagardine production, as assessed by bioassay and MALDI-ToF mass spectrum analysis of culture supernatants. This suggests that in the heterologous host GarM is the only enzyme able to modify the actagardine prepropeptide.

Attempts to complement the *garM* deletion mutant did not restore actagardine production when the complementation clone was analysed using both bioassay and MALDI-ToF mass spectrum analysis of culture supernatants. In principle, reintroduction of the synthetase gene should have been sufficient to restore the biosynthesis of actagardine by the heterologous host. As for the *garA* mutant, there are a number of possible reasons why complementation failed.

The likely last gene in the *garA/garM* operon is *garO.* *garO* is thought to encode a novel enzyme for lantibiotic biosynthesis. The GarO monooxygenase is thought to be responsible for the oxidation of the C-terminal methylanthionine bridge found in actagardine. The homolog found in the *A. liguriae* cluster, *ligO,* is thought to be non-functional as both the natural producer and heterologous hosts containing the AL02 cosmid all produce DAB which lacks the oxidised methylanthionine (Boakes *et al.*, 2010). Deletion of *garO* did not abolish the ability of the heterologous host to produce a compound active against *M. luteus*, as demonstrated by the zone of inhibition surrounding the pIJ12206Δ*garO* clone (figure 5.4A).

Both the actagardine and DAB biosynthetic gene clusters contain a gene, *lanO,* encoding a putative monooxygenase that has not been observed in any other lantibiotic biosynthetic gene cluster. As previously discussed in chapter 3, the *ligO* gene carried on the AL02 cosmid appears to have undergone a frame shift mutation towards the end of the gene that may have rendered the LigO monooxygenase non-functional. As also previously discussed in chapter 4, MALDI-ToF analysis identified Ala(0)-DAA, and a number of other oxidised actagardine derivatives, in culture supernatants of *S. lividans* TK24 clones containing the AG14 cosmid, most of which are likely incomplete oxidation products and oxidation artefacts caused by the MALDI-ToF analysis. Along with Ala(0)-DAA, peaks were observed that correspond to Ala(0)-actagardine peptides with up to three additional oxygen atoms, as well as their corresponding sodium adducts. The additional oxygen atoms are likely to be present on all of the lanthionine bridges in the lantibiotic. While complicating analysis of the Δ*garO* mutant, MALDI-ToF analysis can still be used to assess the affect of deleting the monooxygenase gene. Figure 5.5 shows a comparison of two MALDI-ToF spectra of culture supernatants of clones containing the wild type AG14 cosmid and the Δ*garO* mutant. The intensity of the peaks corresponding to Ala(0)-DAA (1944) and its sodium adduct (1966) are markedly elevated in the Δ*garO* clone. Note also that the distribution of peaks associated with MALDI-ToF associated oxidation of actagardine is shifted to the left in the *garO* mutant. These results provide strong evidence to suggest that the GarO monooxygenase is indeed responsible for the addition of the oxygen moiety to the actagardine peptide.


[M+Na]+
=1982
[M-O+Na]+
=1966
a.u.
1000
800
600
400
200
0
[M+2O+Na]+
=2014
[M+3O+Na]+
=2030
[M+O+Na]+
=1998
[M-Ala]+
=1960
[M-O]+
=1944
1940
1960
1980
2000
2020
2040
2060
m/z
[M-O+Na]+
=1966
a.u.
1500
1250
1000
750
500
250
0
[M+Na]+
=1982
[M+O+Na]+
=1998
[M+2O+Na]+
=2014
[M-Ala]+
=1960
[M+3O+Na]+
=2030
[M-O]+
=1944
1940
1960
1980
2000
2020
2040
2060
m/z

m/z	Description
1944	Ala(0)-deoxyactagardine +H
1960	Ala(0)-actagardine + H
1966	Ala(0)-deoxyactagardine + Na
1982	Ala(0)-actagardine + Na
1998	Ala(0)-actagardine + O + Na
2014	Ala(0)-actagardine + 2O + Na
2030	Ala(0)-actagardine + 3O + Na

Figure 5.5: MALDI-ToF spectra of culture supernatants of two heterologously expressed AG14 cosmids. The top spectrum is from the wild type AG14 cosmid. The bottom spectrum is from a pIJ12206Δ*garO*::scar mutant. The major masses are noted on the spectra, and summarised in the table above. Note that the peak that corresponds to the Ala(0)-DAA sodium adduct (in the black box) has a higher relative intensity in the Δ*garO* clone compared to AG14, suggesting that this gene is indeed able to add the oxygen moiety to the C-terminal methylanthionine bridge of actagardine.

To confirm that the above interpretation of the MALDI-ToF spectra was correct, liquid chromatography/mass spectroscopy (LC-MS) analysis was conducted during a visit to Novacta Biosystems on a Δ*garO*::tet mutant, where the *garO* gene has been replaced with a tetracycline resistance marker. Three clones of the Δ*garO*::tet mutant were grown in liquid culture for seven days, and the supernatants subjected to LC-MS analysis and compared to the supernatants of clones containing the wild type AG14 cosmid. Figure 5.6 shows the results from this experiment.

m/z	Description	
	AG14	AG14Δ*garO*
1) 973		Ala(0)-deoxyactagardine
2) 981	Ala(0)-actagardine	Ala(0)-deoxyactagardine + O
3) 989	Ala(0)-actagardine + O	Ala(0)-deoxyactagardine + 2O
4) 1001	Ala(0)-actagardine + O + Na	

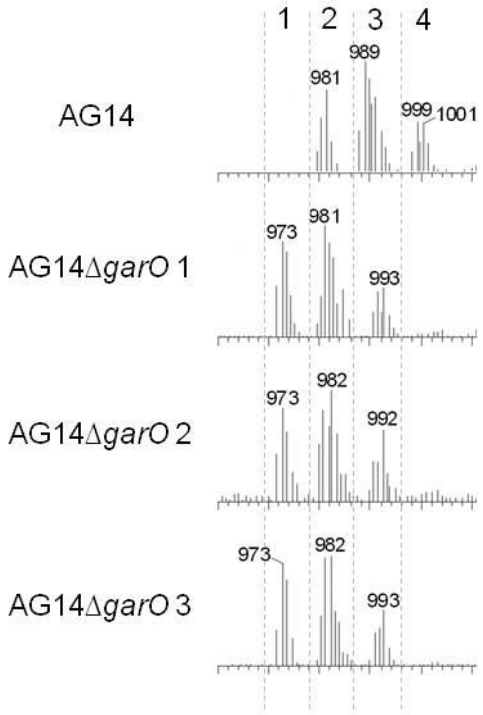


Figure 5.6: LC-MS comparison of three AG14Δ*garO*::TcR clones and the wild type AG14 cosmid in *S. lividans* TK24. The table contains a description of the major masses identified in the spectra. A new peak of 973 Da is observed only in the Δ*garO* mutants, confirming that the GarO monooxygenase is responsible for enzymatic oxidation of actagardine.

LC-MS analysis of culture supernatants from *S. lividans* TK24 containing the wild type AG14 cosmid produced a peak of 981 Da corresponding to Ala(0)-actagardine, and two others representing a single oxidation event (989 Da) and a single oxidation event with an additional sodium adduct (1001 Da). The observed mass peaks in the LC-MS specra are doubly charged leading to an observed mass that is half of that seen in MALDI-ToF specra. In the supernatants of the cultures containing the Δ*garO* deletion mutant, a new peak (973 Da) was observed, corresponding to Ala(0)-DAA. Whether the singly oxidised species observed upon LC-MS analysis of both the wild type and mutant cultures (compared to the numerous oxidations seen in MALDI-ToF spectra (Chapter 4 figure 4.2)) reflects the presence of oxidised actagardine in the culture supernatant, or whether it arose as a consequence of the analysis is not known. Later LC-MS analysis of a pIJ12206Δ*garO*::scar mutant yielded the same results as seen with the tetracycline resistance marked mutant.

As for *garA* and *garM*, attempts to complement the Δ*garO* deletion mutant did not restore wild type levels of Ala(0)-actagardine production; the complementation clone exhibited the same profile as the *garO* mutant upon MALDI-ToF analysis (data not shown).

Bioassay analysis of the pIJ12206Δ*garO*::scar mutant failed to reveal any apparent decrease in bioactivity against *M. luteus*, although this observation may have been compromised by the potential presence of the spontaneously oxidised derivatives. Boakes *et al.* (2009) subsequently confirmed that deletion of *garO* in the natural producer had the same effect, using HPLC-MS to identify Ala(0)-DAA in the culture supernatants of the mutant. This study also demonstrated that the antimicrobial activity of Ala(0)-DAA against *M. luteus* was slightly lower than that of actagardine, but not for all of the Gram-positive pathogens tested (figure 5.7).

	Actagardine	Deoxyactagardine	Vancomycin
Enterococcus faecalis 29212	8, 8	8, 8	2, 2
E. faecium 7131121 (VRE)	64, 64	64, 64	> 64, > 64
E. faecium 19579 (VSE)	16, 16	32, 32	0.5, 1
Micrococcus luteus 4698	1, 1	2, 4	1, 1
Staphylococcus aureus R33 (EMRSA)	8, 16	8, 16	1, 1
S. aureus SH1000 (MSSA)	8, 8	16, 32	2, 2
Streptococcus epidermidis 11047 (MSSE)	32, 32	32, 32	1, 2
Streptococcus pneumoniae R6	16, 32	16, 32	0.5, 0.5

VRE, vancomycin-resistant enterococcus; VSE, vancomycin-sensitive enterococcus; EMRSA, epidemic methicillin-resistant *S. aureus*; MSSA, methicillin-sensitive *S. aureus*; MSSE, methicillin-sensitive *S. epidermidis*.

Figure 5.7: The antimicrobial activity of actagardine, deoxyactagardine and vancomycin against a range of Gram-positive microorganisms shown as minimum inhibitory concentrations (μg ml^{-1}), (Boakes *et al.*, 2009).

This observation was not unexpected as *S. lividans* containing pIJ12207 and producing Ala(0)-DAB also shows clear bioactivity. Thus it appears that biosynthetic oxidation of actagardine has a marginal effect on its antimicrobial activity. It is possible that the oxidised methylanthionine bridge serves another biological purpose, possibly influencing solubility or increasing resistance to proteases.

As mentioned previously, the LigO monooxygenase appears to have a frame-shift mutation towards its C-terminus, and no actagardine B or Ala(0)-actagardine B (Ala(0)-AB) has been identified in culture supernatants from either the natural producer or *S. lividans* derivatives expressing the cloned *lig* gene cluster. Consistent with lack of functionality of LigO, replacement of *garA* with *ligA in A. garbadinensis* led to the production of actagardine B (Boakes *et al.*, 2010). This demonstrates that the minor amino acid differences between actagardine and DAB are unlikely to be responsible for lack of oxidation of the C-terminal methylanthionine bridge of DAB in *A. liguriae*.

.

Deletion of *garR1*, encoding an orphan response regulator, eliminated the production of actagardine in *S. lividans* TK24. Actagardine biosynthesis was not detectable either by bioassay or by MALDI-ToF mass spectrum analysis. As described previously in chapter 3, the *garR1* response regulator and the *garRK* two component regulatory system have corresponding homologues in the mersacidin gene cluster (Guder *et al.*, 2002, Boakes *et al.*, 2009). In the mersacidin cluster, MrsR1 regulates the expression of *mrsA* and *mrsRK* regulate the expression of the immunity transporter MrsFEG. It is therefore possible that the *garR1* response regulator is directly responsible for regulation of *garA* expression. Its absence from the gene cluster would prevent synthesis of the prepropeptide and the subsequent expression of the other biosynthetic genes through the *garRK* signal transduction system, and hence the production of mature actagardine (Guder *et al.*, 2002, Boakes *et al.*, 2009).

The loss of actagardine production in the *garR1* deletion mutant was not restored when the gene was introduced back into the chromosome under the control of the constitutive *Streptomyces ermE** promoter. The possible reasons for this are as for the other mutants where complementation failed.

The two transporter proteins, GarH and GarT, appear not to be essential for biosynthesis and export of actagardine in *S. lividans* TK24, since both pIJ12206Δ*garH*::scar and pIJ12206Δ*garT*::scar yielded zones of inhibition comparable to the wild type cosmid (figure 5.4A), and production of Ala(0)-actagardine was confirmed by MALDI-ToF mass spectrometry. Due to the wealth of secondary metabolites produced by *S. lividans*, it is probable that other transporter systems present within the heterologous host are able to export actagardine from the cell. There is evidence for some lantibiotics that a number of lantibiotic biosynthetic enzymes may form a complex associated with the cell membrane (Chatterjee *et al.*, 2005). The transporter complex could act as an anchor at the cell membrane for the biosynthetic machinery and then immediately export the modified propeptide from the cell. If such a complex occurred in actagardine biosynthesis, deletion of *garTH* would be expected to reduce the level of actagardine production. Since this was not observed, such a complex, if it occurs in *A. garbadinensis*, is not essential for actagardine production.

Heterologous expression of cosmids pIJ1206Δ*garR*::scar and pIJ12206Δ*garK*::scar produced no ala(0)-actagardine that was detectable by either bioassay or by MALDI-ToF mass spectrum analysis. Thus the response regular and sensor kinase appear to be essential for actagardine biosynthesis.

Unlike the actagardine cluster, deletion of the *mrsR2/K2* regulatory genes had no effect on the biosynthesis of mersacidin (Guder *et al.*, 2002). This observation may indicate a role for *garRK* that is different to that of *mrsR2/K2* and may suggest that *garRK* regulate processes in the biosynthetic pathway other than the (unconfirmed) immunity mechanism.

Attempts to complement the *garK* mutant again failed to restore actagardine biosynthesis. Conversely, expression of *garR* from the *ermE** promoter in the *garR* deletion mutant did appear to restore actagardine production in agar-grown cultures (figure 5.4B), although this could not be confirmed by MALDI-ToF mass spectrometry of culture supernatants from liquid-grown cultures. As mentioned previously, this may reflect relatively low levels of of actagardine production in liquid cultures of *S. lividans* TK24 containing the *gar* gene cluster.

The gene encoding the penicillin-binding protein located at one end of the biosynthetic gene cluster has the potential to be involved in producer immunity to actagardine. As mentioned previously, there is a similarly located potential analog in the gene cluster for michiganin A biosynthesis, annotated as an unknown protein with a PASTA domain (figure 5.10) (Holtsmark *et al.*, 2006). As discussed in Chapter 3, PBP and serine/threonine kinase associated domains (PASTA domains) are required for localisation of PBPs to unlinked glycan chains often found near the site of cell division (Yeats *et al.*, 2002, Jones & Dyson, 2006). PBPs have been associated with the enhanced resistance of *Listeria monocytogenes* to nisin. It is believed that the PBP is able to shield lipid II or possibly reduce the effective extracellular concentration of lipid II (Gravesen *et al.*, 2004).

Deletion of *garPBP* did not affect the production of actagardine in the heterologous host, and thus the gene appears to be outside the minimal gene set required for biosynthesis. However, it is possible that the same mutation made in the natural producer could influence actagardine production. As for the transport proteins, it is conceivable that *S. lividans* TK24 may be able to compensate for the lack of GarPBP by providing an alternative immunity mechanism to actagardine (*S. lividans* 1326 (the strain from which TK24 was derived) is naturally resistant to high concentrations of actagardine (up to 100 µg/ml); Novacta Biosystems, personal communication).

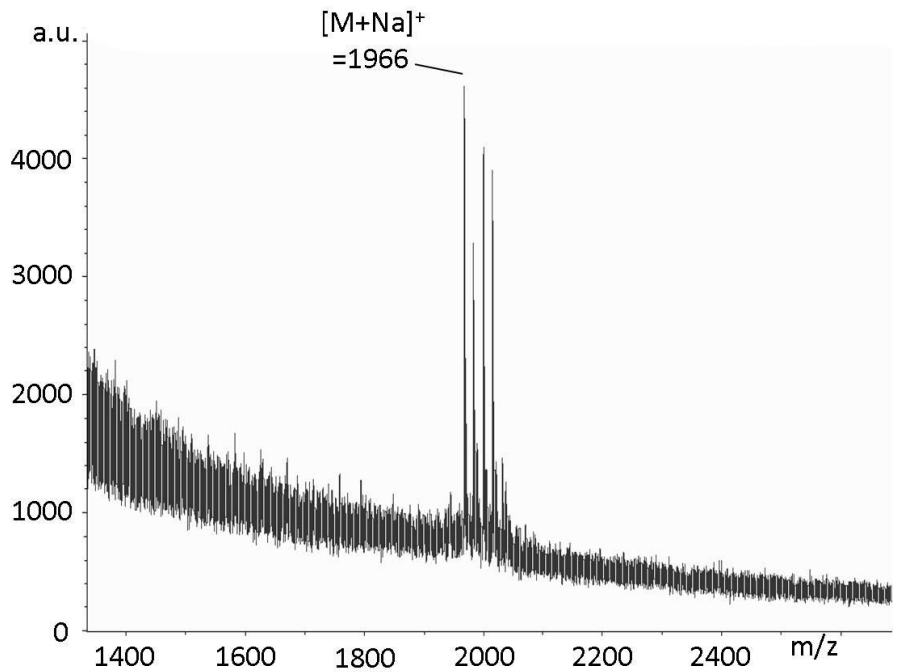
Nonetheless, this result suggests that GarPBP (and most likely the protein with a PASTA domain associated with the michiganin A gene cluster) is not involved in lantibiotic immunity. Even though *S. lividans* TK24 is intrinsically resistant to actagardine, given the likely interplay between immunity and biosynthesis, deletion of the PBP gene, if part of the gar gene cluster, would have been expected to have had some effect of actagardine biosynthesis, but none was apparent.

The *A. garbadinensis* and *A. liguriae* gene clusters are highly homologous, sharing eight out of a total of nine genes potentially required for biosynthesis of actagardine and DAB. The DAB cluster has the full complement of genes found in the actagardine cluster, and the function of these genes is likely to be identical to those found in the actagardine cluster with one exception. The DAB gene cluster encodes an additional ATP binding cassette transporter (LigQ) that has 51% amino acid sequence identity to a putative ABC transporter from *S. coelicolor*, SCO3235 (Boakes *et al.*, 2010). Deletion of *ligQ* appeared to almost abolish activity against

M. luteus, (figure 5.4A), although some DAB was detectable upon MALDI-ToF analysis of culture supernatants (figure 5.8). Expression of *ligQ* using the *ermE** promoter only marginally increased levels of Ala(0)-DAB production, but was not able to restore wild type levels, based on the size of the zone of inhibition, and could not be confirmed by MALDI-ToF analysis (as for *garR* previously) (figure 5.4B). Thus while not absolutely essential, the LigQ ABC transporter appears to play an important role in determining the level of Ala(0)-DAB synthesis and/or export in *S. lividans* TK24.

It is possible that LigQ is performing a similar role to LanFEG immunity transporters found in the gene clusters for the related lantibiotics michiganin A and mersacidin (Guder *et al.*, 2002, Holtsmark *et al.*, 2006). In the absence of LigQ-mediated immunity, Ala(0)-DAB synthesis could be prevented as a fail-safe mechanism to prevent the organism killing itself. If LigQ does encode an immunity mechanism for DAB production, it is surprising that the actagardine gene cluster does not possess an equivalent mechanism for self-resistance.

A



B

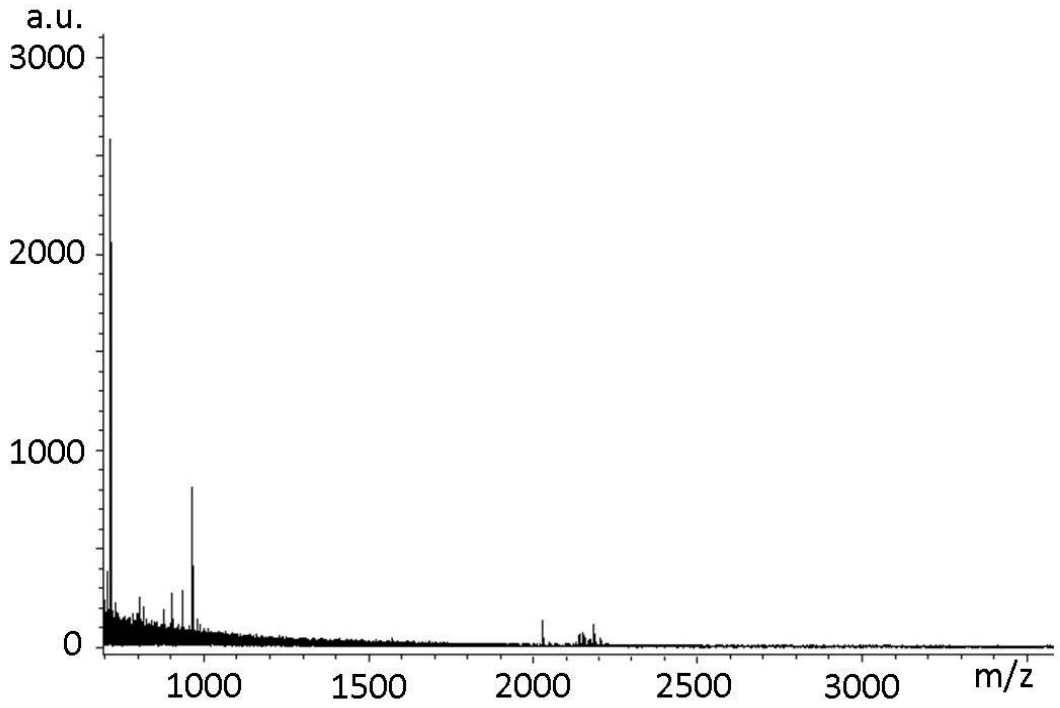


Figure 5.8: MALDI-ToF spectra of an *S. lividans* clone AL02Δ*ligQ*::scar (A) and vector only control M1216 (B). The ala(0)-deoxyactagardine B ($[M+Na]^+$ = 1966) peak is present in the AL02Δ*ligQ*::scar clone and not the negative control (M1216).

5.3 Analysis of actagardine homologs

5.3.1 Introduction

The gene clusters that produce actagardine and DAB, found in *A. garbadinensis* and *A. liguriae*, respectively, are very similar and produce lantibiotics that differ by only two amino acids and one lanthionine bridge modification. To date, two other lantibiotics (one characterised, one putative) have been identified that share similar peptide sequences to these two lantibiotics. Michiganin A is an actagardine homolog that is produced by *Clavibacter michiganensis* (Holtsmark *et al.*, 2006). *Clavibacter michiganensis* subsp. *michiganensis* is described as a plant-pathogenic actinomycete, infecting tomato and causing bacterial wilt and canker (Gartemann *et al.*, 2003). The gene cluster for michiganin A is however markedly different from those encoding actagardine and DAB despite producing a lantibiotic that is very similar to actagardine (figure 5.9 and 5.10). As discussed in chapter 3, bioinformatic analysis identified another homolog of actagardine named in this study as gnavucin. The gene encoding the prepropeptide for this putative lantibiotic, *gnaA*, was identified in the genomes of two bacteria, *Ruminococcus gnavus* ATCC 29149 and *Blautia hansenii* DSM 20583. *R. gnavus* is described as a Gram positive, non spore-forming obligate anaerobe found normally in the human alimentary tract (Hata & Smith, 2004). The *R. gnavus* E1 strain is responsible for producing the lantibiotic ruminococcin A, which differs significantly from gnavucin, when grown in the presence of trypsin (Dabard *et al.*, 2001, Gomez *et al.*, 2002, Marcille *et al.*, 2002). The structure of gnavucin was predicted from the amino acid sequence of the encoded peptide and modelled on the structure of actagardine. Only one other putative lantibiotic biosynthetic gene was identified in close proximity to the *gnaA* gene, a *lanM* type synthetase gene. It is thus possible that these two genes represent a remnant of an earlier lantibiotic gene cluster, and it seems unlikely that the predicted lantibiotic is produced by its host strain. In contrast, in addition to the *gnaA* homologue, *B. hansenii* contains an intact *lanM* gene.

All three gene clusters were analysed using the same tools and methods outlined in chapter 3. The biosynthetic gene cluster of michiganin A has been submitted to the National Centre for Biotechnology by Holtsmark *et al.* (2006). The biosynthetic genes for gnavucin were identified by BLAST searching using the amino acid sequence of GarA as query sequence. The frame-plot algorithm could not be used

to identify other putative lantibiotic biosynthetic genes close to *gnaA* due to the low G-C content of the *R. gnavus* genome. Instead, ORFs were identified manually by locating gaps in the series of stop codons using Artemis, and then searching by BLAST for homologues in the public database.

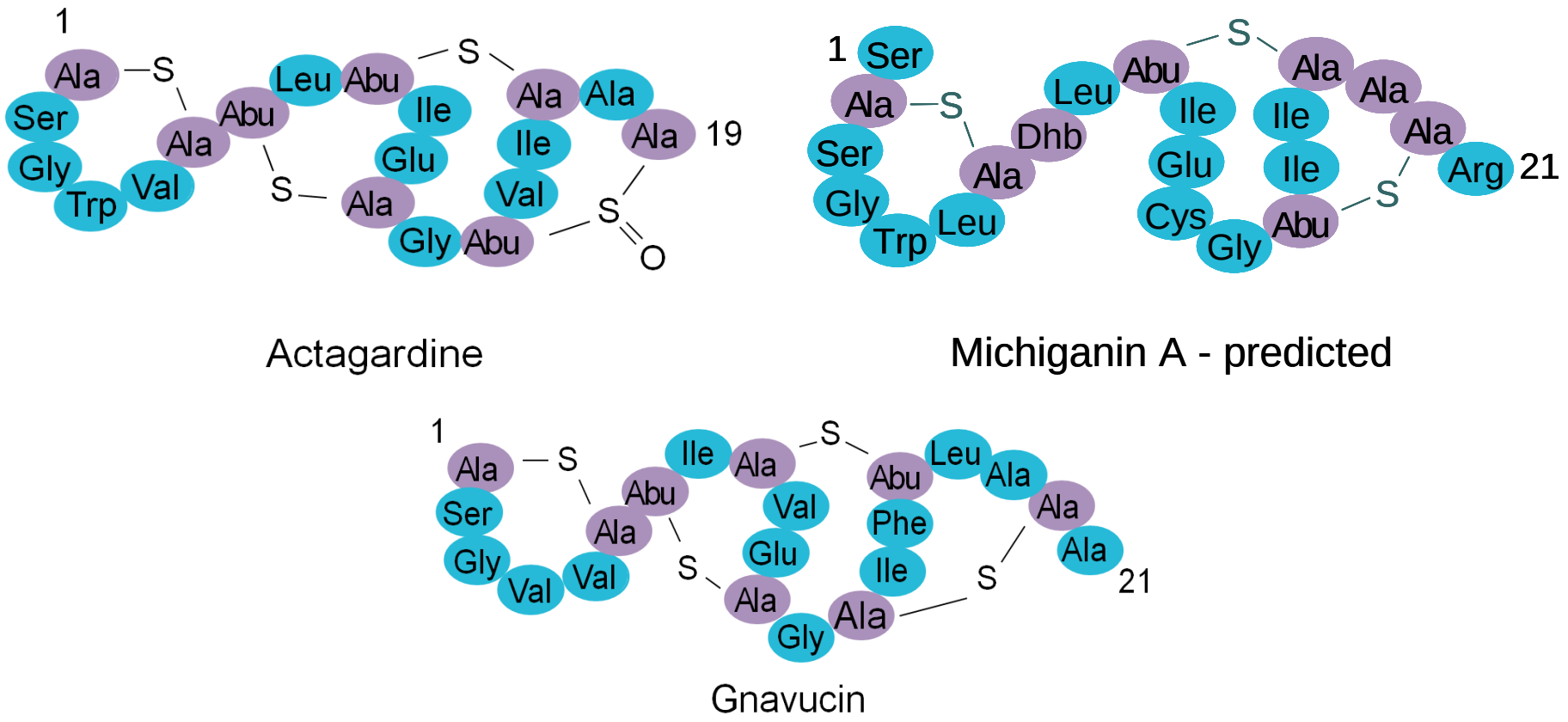


Figure 5.9: The actual and predicted structures of actagardine and two homologs, michiganin A and gnavucin. The structure of michiganin A displayed here is taken from Holksmark *et al*. (2006). It is possible that a lanthionine bridge exists between Dhb9 and Cys14, which would result in the lantibiotic closely resembling actagardine in structure; it is unclear why the authors chose not to include it in their predicted structure. The structure of gnavucin was predicted from its primary amino acid sequence using the actagardine structure as a template.

5.4 Computational analysis of the *gnaA* and *clvA* gene clusters

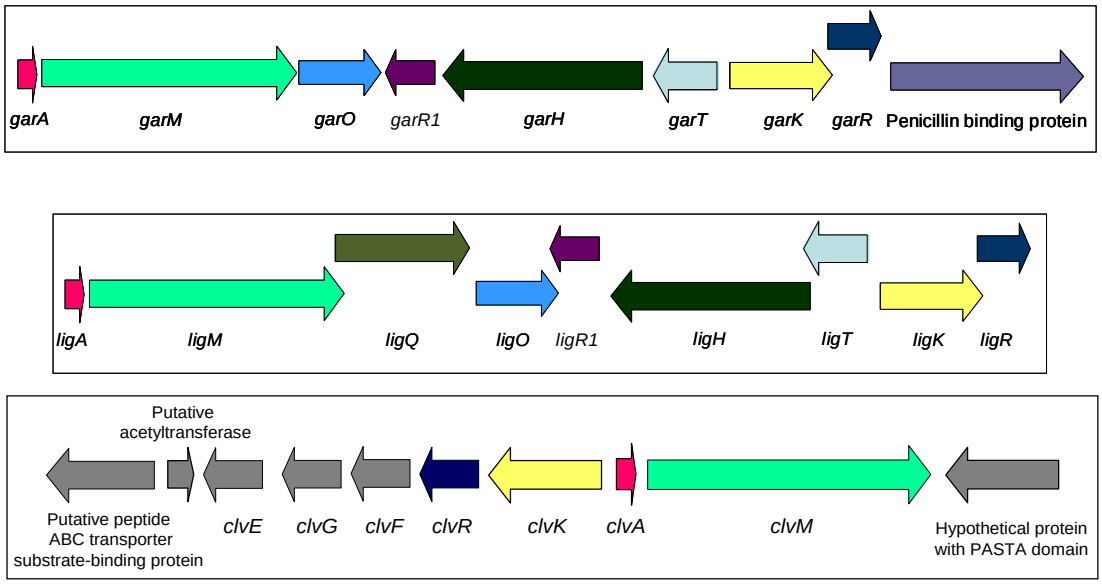


Figure 5.10: The proposed biosynthetic gene clusters of actagardine (top), DAB (middle) and michiganin A (bottom). Genes highlighted in the same colour are predicted to be functionally analogous.

As seen in figure 5.10, while the gene clusters for actagardine and DAB biosynthesis are very similar, that for michiganin A is quite different. The michiganin A gene cluster only shares similarity with respect to the relative position of the *lanA* and *lanM* genes in the cluster, and lacks *lanO* and *lanTH* transporter genes. Instead, it encodes a ClvFEG transporter to presumably export the peptide and to provide immunity to the lantibiotic (Holtsmark *et al.*, 2006). As mentioned previously in chapter 3 section 3.1.2, the ClvA prepropeptide has significant homology to GarA, sharing 29 % amino acid identity, which increases to 89 % when comparing the propeptide amino acid sequences. The leader sequence of ClvA lacks the PAG motif found in the leader sequences of all of the other actagardine homologs and mersacidin (chapter 3, figure 3.4). ClvM shares some homology with GarM, with 29 % amino acid identity over the length of the entire protein. ClvK has higher homology to LigK than GarK, sharing 38 % amino acid sequence identity. ClvR also shares higher homology with LigR than GarR, sharing 52 % amino acid identity. ClvFEG is likely to form an ABC transporter composed of two cytoplasmic ATP-binding proteins (ClvFG) and a single membrane protein (ClvE)

(Siegers & Entian, 1995). There is also a hypothetical protein that is encoded by a gene downstream of *clvM* that has a potential PASTA domain which are also found in PBPs, as discussed previously.

GnaA shares 33% amino acid identity to GarA. The putative *gnaM* gene is split by an insertion element (although this could reflect an error during sequence assembly). Reconstitution of the *gnaM* gene *in silico* revealed 24% amino acid sequence identity to GarM and LigM, and 99% amino acid identity to the LanM present in *B. hansenii.*

To determine if the prepropeptides of *clvA* or *gnaA* could act as substrates for the actagardine biosynthetic machinery, both genes were synthesised and placed under control of the e*rmE** promoter in constructs identical to those used in the complementation experiments. These constructs were subsequently transferred into *S. lividans* TK24 containing the *garA* deleted cosmid. No halos were observed upon bioassay of agar-grown cultures against *M. luteus*, and no peaks were detected corresponding to gnavucin or michiganin A upon MALDI-ToF analysis of liquid culture supernatants. It is unclear whether these negative results are due to incompatibility between these prepropeptides and the actagardine biosynthetic machinery, or reflect the same issues observed with many of the previous failed attempts to achieve complementation of *gar* mutants, including *garA*.

5.5 Discussion

In this chapter an apramycin cassette was used to introduce in-frame deletion mutations into all of the proposed lantibiotic biosynthesis genes within the actagardine biosynthetic gene cluster present on the AG14 cosmid. The subsequent clones were subjected to bioassays and MALDI-ToF mass spectrum analysis. Out of the nine putative biosynthetic genes (*garA* to *garPBP*), five were likely to be essential for actagardine biosynthesis in the heterologous host *S. lividans* TK24.

Not surprisingly, all three of the proposed prepropeptide encoding and modification genes (*garA, garM* and *garO)* were likely to be essential for Ala(0)-actagardine biosynthesis. Deletion of *garA* in the natural producer *A. garbadinensis* also yielded a non-producing phenotype (Boakes *et al.*, 2009). Despite the complications caused by spontaneous oxidation, the tailoring gene *garO* is likely essential for the production of Ala(0)-actagardine in the heterologous host. Deletion of g*arO* appeared to have little effect on antimicrobial activity against *M. luteus*, consistent with published data (Boakes *et al.*, 2010). MALDI-ToF mass spectrum analysis of this mutant, and subsequently the wild type gene cluster (both in *S. lividans*), detected a peak corresponding to the biosynthesis of a new actagardine derivative, Ala(0)-DAA.

The three putative regulatory genes, *garR1, garR* and *garK,* are likely to be required for actagardine biosynthesis, but only *garR* could be confirmed (by successful complementation) as essential. Based on work on mersacidin, it is possible that GarR1 directly regulates expression of the *garAMO* operon. Unlike the mersacidin cluster, there is no obvious resistance mechanism that is regulated by GarRK, (apart from the ABC transporter GarTH) , but it is conceivable that GarRK, by analogy, regulate *garTH* expression (Guder *et al.*, 2002). The michiganin A gene cluster contains a *lanFEG* transporter system much like mersacidin. These transporters have been associated with immunity mechanisms for a number of lantibiotics, and are likely to be regulated by *lanRK* two component regulatory systems (Guder *et al.*, 2002, Chatterjee *et al.*, 2005).

The two transporter genes *garH* and *garT* appear to be unnecessary for actagardine production and export in the heterologous host, and presumably transporters encoded elsewhere in the host's genome are able to compensate for the loss of the lantibiotic transporters. Deletion of *garPBP* did not affect the biosynthesis of actagardine, while conversely, deletion of *ligQ*, encoding a putative immunity mechanism in *A. liguriae,* appeared to severely limit DAB production.

Attempts to complement the mutations, with one exception (*garR*), failed. This could reflect issues with expression levels at the ΦBT1 attachment site, or perhaps more likely, a complex mechanism of regulation that requires stoichiometric production of the various biosynthetic components.

Bioinformatic analysis identified two actagardine-like peptides, michiganin A and gnavucin, although their gene clusters differ markedly from those of their *Actinoplanes* counterparts. This is intriguing and raises the possibility that the actagardine "scaffold" is perhaps more common than had been previously recognised.

5.6 Summary

- The actagardine prepropeptide (GarA) and modification enzymes (GarM and GarO) are likely to be essential for actagardine biosynthesis.

- Deletion of *garO*, encoding a monooxygenase, resulted in the detection of a novel actagardine derivative, Ala(0)-deoxyactagardine; DAB was also identified in *A. liguriae*.

- Of the three regulatory genes (*garR1, garR and garK*), *garR* is clearly essential for biosynthesis.

- The putative lantibiotic transporters GarTH are not essential for actagardine biosynthesis and export. *S. lividans* TK24 appears to be able to compensate in their absence.

- GarPBP is also not essential for actagardine biosynthesis, and is likely not part of the *gar* gene cluster.

- LigQ, a putative immunity transporter, is required for wild type levels of DAB biosynthesis.

- Bioinformatic analysis identified two actagardine homologs, michiganin A and gnavucin. The gene cluster for michiganin A is markedly different to that of actagardine and DAB, yet produces a very similar lantibiotic.

6 Discussion

In this study, the actagardine and DAB biosynthetic gene clusters from *A. garbadinensis* and *A. liguriae* were subjected to detailed analysis. Heterologous expression of two cosmids, AG14 from *A. garbadinensis* and AL02 from *A. liguriae*, confirmed that they contained all of the genes required for the biosynthesis of actagardine and DAB, respectively. For the first time, this study confirmed that DAB can be synthesised by a heterologous host, *S. lividans* TK24. Moreover, the expression of the AG14 cosmid in the modified *S. coelicolor* strain M1146 suggested that deletion of secondary metabolite pathways from the host genome considerably improved the yield of lantibiotic production.

Analysis of the two cosmids identified a number of ORFs that encoded putative lantibiotic biosynthetic enzymes, many of which are also encoded by other type B lantibiotic gene clusters (Chatterjee *et al.*, 2005). This allowed the bioinformatic prediction of MGSs for both actagardine and DAB biosynthesis. Deletion of genes that flanked the putative actagardine MGS confirmed that they were not required for biosynthesis, even though some of them had some functional analogy with genes adjacent to the putative DAB MGS. Expression of three other *A. liguriae* cosmids, identified by Novacta Biosystems and with a centrally located *ligA*, also served to narrow down the number of potential DAB biosynthetic genes. Combining the computational analysis and cosmid reduction mutations, more confident MGSs were predicted for both actagardine and DAB production. To confirm the involvement of each gene in the actagardine MGS in lantibiotic biosynthesis, a deletion series was created, deleting each gene in turn and analysing its effects on actagardine production.

Nine genes were initially predicted to be involved in the biosynthesis of actagardine. Mutational analysis of the AG14 cosmid determined that *garA*, encoding the prepropeptide, appeared essential for the biosynthesis of actagardine in the *S. lividans* TK24 heterologous host. Deletion of *garA* has also been accomplished in the natural producer, resulting in the same loss of lantibiotic biosynthesis (Boakes *et al.*, 2010, Boakes *et al.*, 2009). The modification gene *garM* also appeared essential for production.

Three putative regulatory genes, *garR1*, *garR* and *garK*, all appeared essential for actagardine biosynthesis, but only *garR* could be confirmed as essential for actagardine biosynthesis in the heterologous host by genetic complementation. Deletion of *garR1* abolished actagardine biosynthesis, and by comparison to mersacidin and MrsR1, it has been suggested that GarR1 regulates the expression of *garA*. Unlike mersacidin (Guder *et al.*, 2002), however, the GarRK two component system does not appear to solely regulate the putative immunity mechanism since they are essential for the biosynthesis of actagardine, while the transport genes *garTH* are not. The michiganin A gene cluster does not have a homolog of *garR1*, and must regulate the expression of *clvA* by another mechanism. It does, however, have homologs of the two component LanRK regulators that are also encoded by the actagardine, DAB and mersacidin gene clusters. Using mersacidin as a model, the actagardine and DAB two component LanRK may regulate expression of the immunity response to the lantibiotics that they produce. As indicated above, surprisingly the two component transport system, *garTH,* was not required for actagardine export, at least not in the heterologous host, where alternative transporters can presumably carry out this function. Apart from these two genes, the only other potential immunity gene in the actagardine cluster is *garPBP.* While deletion of *garPBP* appeared to have no affect on actagardine biosynthesis in the heterologous host, it is conceivable that any phenotypic consequence was masked by the intrinsic resistance of *S. lividans* TK24 to actagardine (up to 100 µg/ml, Novacta Biosystems, personal communication). Expression of *garPBP* in an actagardine sensitive host, where its introduction provides immunity to actagardine, could confirm the involvement of GarPBP in self-resistance to actagardine.

The michiganin A gene cluster does not possess close *garTH* homologues, but instead contains the *clvFEG* transporter genes thought to be involved in immunity. However, there are putative ABC transporter genes in close proximity to the *clv* gene cluster that have not yet been studied in any detail. Many lantibiotic gene clusters, for example that for nisin, encode transporter proteins (such as LanTH) for export of the lantibiotic and in some cases some immunity, and in addition a dedicated immunity transporter (LanFEG), providing high levels of immunity to the lantibiotic.

The most unusual gene within the actagardine cluster is *garO*, predicted to encode a monooxygenase enzyme that adds an oxygen moiety to the C-terminal methylanthionine bride of actagardine. This gene is not essential for peptide production, and *S. lividans* TK24 clones expressing AG14Δ*garO* produce readily detectable quantities of the novel compound DAA. DAA has higher MICs against a number of pathogens indicating that the addition of an oxygen to the C-terminal methyl-lanthionine bridge does increase antimicrobial potency (Boakes *et al.*, 2009). However, the lack of detectable oxidation of michiganin A, and the lack of potential monooxygenase genes in the *clv* and *gna* gene clusters, suggests that this modification is not crucial in an ecological setting.

The AL02 cosmid (from *A. liguriae*) that encodes DAB production possesses close homologues of many of the genes found in the actagardine biosynthesis cluster, and they are likely to perform the same function in DAB biosynthesis. Notable differences are the additional gene *ligQ,* and the absence of a *garPBP* homologue. Unlike *garPBP*, *ligQ* is essential for wild-type levels of the heterologous production of DAB by *S. lividans* TK24 and is predicted to play a role in immunity of the producing organism to the lantibiotic.

The biosynthetic gene clusters of actagardine and DAB are highly homologous, and share a significant amount of sequence identity that rapidly decreases outside of the predicted MGSs (Boakes *et al.*, 2009). The DAB cluster appears to contain a non-functional LigO monooxygenase that is unable to oxidise the C-terminal methyl-lanthionine bridge of the DAB peptide. Computational analysis and comparison to GarO predicted a frame shift mutation towards the end of *ligO* that presumably abolished its function. This is consistent with the results of (Boakes *et al.*, 2010). The DAB prepropeptide substrate, which differs by two amino acids to actagardine, is amenable to oxidation by GarO when expressed in a AG14Δ*garA* mutant of *A. garbadinensis*, resulting in the production of a novel lantibiotic, actagardine B (Boakes *et al.*, 2010).

The DAB gene cluster, as mentioned, also contains the additional gene *ligQ* that encodes an ABC transporter that has no counterpart in the actagardine or michiganin A clusters. It is hypothesised that this gene plays a role in immunity of the host to DAB in

a similar way to the presumed immunity transporter ClvFEG encoded by the michiganin A gene cluster. While it cannot be directly implicated, GarPBP is a possible candidate for an immunity mechanism encoded by the actagardine cluster, if it has one at all. If this is an immunity mechanism, then it has one clear distinction from LigQ - it is not essential for lantibiotic biosynthesis.

BLAST searches using the GarA prepropeptide identified two actagardine homologs, ClvA and GnaA. ClvA is the prepropeptide of michiganin A, a lantibiotic from the plant pathogen *C. michiganensis*, while GnaA is a thus far unidentified actagardine analogue encoded within the newly sequenced genomes of *R. gnavus* and *B. hansenii* (Holtsmark *et al.*, 2006). While the amino acid sequences of these peptides vary considerably, all still maintain an overall structure similar to actagardine. The structure of the lantibiotic is key to its antimicrobial activity, perhaps more so than the amino acid sequence, since many peptide variants (of actagardine and mersacidin, for example) often retain reasonable levels of antimicrobial activity (Appleyard *et al.*, 2009, Boakes *et al.*, 2009). The leader sequences of the lantibiotics are also very different, some missing motifs conserved in other members, and the leader sequences also vary in size, with the gnavucin leader sequence being significantly shorter than the others. These differences most likely reflect variation in the modification enzymes of the individual natural producers.

Actagardine (and derivatives thereof) could potentially have a role in the clinical environment and the lantibiotic is currently being used as a scaffold for an antibiotic being developed by Novacta Biosystems. It is therefore important to gain as much information as possible about the enzymes that are utilised in its biosynthesis, and the mechanism of regulation. The actagardine “scaffold” has been identified in five real or putative lantibiotics produced by sometimes quite different biosynthetic gene clusters. This resembles in nature, if not extent, the lacticin 481 family of lantibiotics that has at least 16 members that vary in both the lantibiotic prepropeptide sequence, and gene cluster composition and arrangement (Dufour *et al.*, 2007). It is likely that we are yet to discover other actagardine homologs and actagardine-like gene clusters. In the environment, the actagardine scaffold must have significant antimicrobial activity against a number of hosts for it to be made by such different microorganisms.

6.1 Future work - Shorter term:

Complementation of deleted genes should be further pursued. This would allow the unambiguous identification of genes essential for the biosynthesis of actagardine and DAB. One potential host for these studies would be *S. coelicolor* M1146. In this study *S. coelicolor* M1146 appeared able to produce Ala(0)-actagardine at much higher levels and at earlier time points than the host used in this study, *S. lividans* TK24. The complemented genes should also be configured to be expressed from their native promoters to mimmick, as closely as possible, expression levels in the wild type strain. Analysis of agar plugs using MALDI-ToF spectroscopy could also be used to increase the sensitivity of detection of Ala(0)-actagardine and Ala(0)-DAB production, and hence complementation.

To explore the actagardine homologs and their biosynthesis further, both the *gnaA* and *clvA* prepropeptides should be expressed in the *S. lividans* TK24 AG14Δ*garA* clone using Novacta Biosystems' pAGvarX construct (Boakes *et al.*, 2010, Boakes *et al.*, 2009). These constructs could avoid the previous problems encountered using the *ermE** promoter constructs and if expressed, could demonstrate the plasticity of the actagardine biosynthetic machinery. Production of gnavucin by this method would demonstrate that the GnaA prepropeptide is a viable substrate for lantibiotic biosynthesis, and would result in the production of a novel lantibiotic peptide. If this lantibiotic is produced, further experiments would be required to determine its antimicrobial activity and potential for clinical application. Heterologous production of michiganin A and gnavucin could provide new insights into the specificity of a number of the actagardine biosynthetic enzymes, including GarO. If the michiganin A or gnavucin peptide is oxidised, it would demonstrate the possibility of using the GarO monooxygenase as a general modification enzyme for actagardine-like lantibiotics with the prospect of improved antimicrobial activity.

qRT-PCR analysis of the actagardine biosynthetic gene cluster could identify the relative expression levels of each gene in both the wild type, and a number of deletion constructs. This work should be attempted in *S. coelicolor* M1146 mutants since the level of *gar* gene expression is likely to be higher, and to occur earlier in growth than in

S. lividans TK24. This would clarify the role of each of the three regulators in the gene cluster, the GarRK two component regulatory system and GarR1, in both actagardine biosynthesis and immunity. Similar experiments should be carried out with the DAB gene cluster, in particular to address the role of the essential LigQ.

Deletions should also be made in the natural producers of all of the genes analysed in this study that did not appear to be necessary for actagardine biosynthesis. The natural producer, where each biosynthetic protein has presumably evolved to perform a specific role in actagardine biosynthesis, may be less tolerant of gene deletions than the heterologous strain, where host-encoded functions might fortuitously compensate for deletions within the cluster. This is key to understanding the role of genes such as *garPBP,* where there is no clear evidence to suggest that it is a functional component of actagardine biosynthesis. Similarly, *garTH* should also be deleted in *A. garbadinensis*.

6.2 Future work - longer term:

The most novel gene in the actagardine biosynthesis cluster is GarO, and it should be studied in greater detail in the future. This enzyme has the potential to increase the effectiveness of clinically useful lantibiotics. The substrate specificity of this enzyme should be determined. For example, is the enzyme limited to recognising the entire structure of actagardine, or can it use the (methyl) lanthionine bridges found in many other lantibiotics, such as michiganin A, as substrate. The crystal structure of GarO should be determined, and its active site and binding pockets identified, thus aiding our understanding of how it functions and its ability to modify actagardine and DAB. Potentially, this information could lead to engineering of the enzyme, extending its substrate specificity, and potentially leading to improved therapeutic agents.

Efforts should also be made to confirm, or otherwise, the role and mechanism of action of GarPBP. If this enzyme is involved in immunity of *A. garbadinensis* to actagardine, it would be the first example of the direct involvement of a PBP in lantibiotic resistance. If this is the case, the mechanism by which it exerts actagardine resistance should be studied in detail.

This study has identified three regulators that are all essential for biosynthesis of actagardine in *S. lividans* TK24. It is essential to understand the role of each of these regulators as they do not fit any current model of lantibiotic regulation. This would enable knowledge-based approaches to improve the level of actagardine production, which may be crucial for its commercial application.

Further studies should also be made to confirm the mode of action of actagardine. It has been suggested that the molecule is able to interact with Lipid II, but this has never been unambiguously established. Understanding how actagardine binds to Lipid II could lead to knowledge-based optimisation of the lantibiotic. The latter is particularly attractive given the ribosomal origin of lantibiotics, where prepropeptide variants can be generated quickly and economically by oligonucleotide synthesis.

6.3 Final conclusions

Actagardine and its gene cluster represent a new subgroup of lantibiotic molecules that have a number of defining features. The structure of the lantibiotic is conserved in at least five different organisms, and may possibly be found elsewhere. While the peptide structures are similar, the leader sequences that are attached to them vary markedly. Some of the gene clusters contain bespoke components, such as the ClvFEG transporter in the michiganin A gene cluster, and the LanO monooxygenase found in the actagardine and DAB clusters. This study has identified and confirmed the minimal number of genes required to produce actagardine in a heterologous host. It has also paved the way for further studies into the complex role some of the genes play in its biosynthesis and regulation. The information gleaned in this study will hopefully be used to advance the field of actagardine research, and contribute to its use as a clinically useful antimicrobial agent.

7 References

Altena, K., A. Guder, C. Cramer & G. Bierbaum, (2000) Biosynthesis of the lantibiotic mersacidin: organization of a type B lantibiotic gene cluster. *Appl Environ Microbiol* 66: 2565-2571.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers & D. J. Lipman, (1990) Basic local alignment search tool. *J Mol Biol* 215: 403-410.

Appleyard, A. N., S. Choi, D. M. Read, A. Lightfoot, S. Boakes, A. Hoffmann, I. Chopra, G. Bierbaum, B. A. Rudd, M. J. Dawson & J. Cortes, (2009) Dissecting structural and functional diversity of the lantibiotic mersacidin. *Chem Biol* 16: 490-498.

Arioli, V., M. Berti & L. G. Silvestri, (1976) Gardimycin, a new antibiotic from Actinoplanes. III. Biological properties. *J Antibiot (Tokyo)* 29: 511-515.

Asaduzzaman, S. M. & K. Sonomoto, (2009) Lantibiotics: diverse activities and unique modes of action. *J Biosci Bioeng* 107: 475-487.

Baltz, R. H., (1998) Genetic manipulation of antibiotic-producing Streptomyces. *Trends Microbiol* 6: 76-83.

Bibb, M. J., P. R. Findlay & M. W. Johnson, (1984) The relationship between base composition and codon usage in bacterial genes and its use for the simple and reliable identification of protein-coding sequences. *Gene* 30: 157-166.

Bierman, M., R. Logan, K. O'Brien, E. T. Seno, R. N. Rao & B. E. Schoner, (1992) Plasmid cloning vectors for the conjugal transfer of DNA from Escherichia coli to Streptomyces spp. *Gene* 116: 43-49.

Blaesse, M., T. Kupke, R. Huber & S. Steinbacher, (2003) Structure of MrsD, an FAD-binding protein of the HFCD family. *Acta Crystallogr D Biol Crystallogr* 59: 1414-1421.

Boakes, S., A. N. Appleyard, J. Cortes & M. J. Dawson, (2010) Organization of the biosynthetic genes encoding deoxyactagardine B (DAB), a new lantibiotic produced by Actinoplanes liguriae NCIMB41362. *J Antibiot (Tokyo)*.

Boakes, S., J. Cortes, A. N. Appleyard, B. A. Rudd & M. J. Dawson, (2009) Organization of the genes encoding the biosynthesis of actagardine and engineering of a variant generation system. *Mol Microbiol* 72: 1126-1136.

Breukink, E., C. van Kraaij, R. A. Demel, R. J. Siezen, O. P. Kuipers & B. de Kruijff, (1997) The C-terminal region of nisin is responsible for the initial interaction of nisin with the target membrane. *Biochemistry* 36: 6968-6976.

Breukink, E., I. Wiedemann, C. van Kraaij, O. P. Kuipers, H. Sahl & B. de Kruijff, (1999) Use of the cell wall precursor lipid II by a pore-forming peptide antibiotic. *Science* 286: 2361-2364.

Brotz, H., G. Bierbaum, K. Leopold, P. E. Reynolds & H. G. Sahl, (1998) The lantibiotic mersacidin inhibits peptidoglycan synthesis by targeting lipid II. *Antimicrob Agents Chemother* 42: 154-160.

Brotz, H., G. Bierbaum, P. E. Reynolds & H. G. Sahl, (1997) The lantibiotic mersacidin inhibits peptidoglycan biosynthesis at the level of transglycosylation. *Eur J Biochem* 246: 193-199.

Brotz, H. & H. G. Sahl, (2000) New insights into the mechanism of action of lantibiotics--diverse biological effects by binding to the same molecular target. *J Antimicrob Chemother* 46: 1-6.

Burrage, S., T. Raynham, G. Williams, J. W. Essex, C. Allen, M. Cardno, V. Swali & M. Bradley, (2000) Biomimetic synthesis of lantibiotics. *Chemistry* 6: 1455-1466.

Chakicherla, A. & 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: 23533-23539.

Chan, W. C., H. M. Dodd, N. Horn, K. Maclean, L. Y. Lian, B. W. Bycroft, M. J. Gasson & G. C. Roberts, (1996) Structure-activity relationships in the peptide antibiotic nisin: role of dehydroalanine 5. *Appl Environ Microbiol* 62: 2966-2969.

Chandrapati, S. & D. J. O'Sullivan, (1999) Nisin independent induction of the nisA promoter in Lactococcus lactis during growth in lactose or galactose. *FEMS Microbiol Lett* 170: 191-198.

Chatterjee, C., M. Paul, L. Xie & W. A. van der Donk, (2005) Biosynthesis and mode of action of lantibiotics. *Chem Rev* 105: 633-684.

Chen, P., F. Qi, J. Novak & P. W. Caufield, (1999) The specific genes for lantibiotic mutacin II biosynthesis in Streptococcus mutans T8 are clustered and can be transferred en bloc. *Appl Environ Microbiol* 65: 1356-1360.

Chen, P., F. X. Qi, J. Novak, R. E. Krull & 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: 139-144.

Chen, W., T. Huang, X. He, Q. Meng, D. You, L. Bai, J. Li, M. Wu, R. Li, Z. Xie, H. Zhou, X. Zhou, H. Tan & Z. Deng, (2009) Characterization of the polyoxin

biosynthetic gene cluster from Streptomyces cacaoi and engineered production of polyoxin H. *J Biol Chem* 284: 10627-10638.

Cherepanov, P. P. & 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: 9-14.

Chung, Y. J., M. T. Steen & J. N. Hansen, (1992) The subtilin gene of Bacillus subtilis ATCC 6633 is encoded in an operon that contains a homolog of the hemolysin B transport protein. *J Bacteriol* 174: 1417-1422.

Claesen, J. & 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: 16297-16302.

Cooper, L. E., A. L. McClerren, A. Chary & W. A. van der Donk, (2008) Structure-activity relationship studies of the two-component lantibiotic haloduracin. *Chem Biol* 15: 1035-1045.

Coronelli, C., G. Tamoni & G. C. Lancini, (1976) Gardimycin, a new antibiotic from Actinoplanes. II. Isolation and preliminary characterization. *J Antibiot (Tokyo)* 29: 507-510.

Corvey, C., T. Stein, S. Dusterhus, M. Karas & K. D. Entian, (2003) Activation of subtilin precursors by Bacillus subtilis extracellular serine proteases subtilisin (AprE), WprA, and Vpr. *Biochem Biophys Res Commun* 304: 48-54.

Cotter, P. D., L. H. Deegan, E. M. Lawton, L. A. Draper, P. M. O'Connor, C. Hill & R. P. Ross, (2006) Complete alanine scanning of the two-component lantibiotic lacticin 3147: generating a blueprint for rational drug design. *Mol Microbiol* 62: 735-747.

Dabard, J., C. Bridonneau, C. Phillipe, P. Anglade, D. Molle, M. Nardi, M. Ladire, H. Girardin, F. Marcille, A. Gomez & M. Fons, (2001) Ruminococcin A, a new lantibiotic produced by a Ruminococcus gnavus strain isolated from human feces. *Appl Environ Microbiol* 67: 4111-4118.

Datsenko, K. A. & 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: 6640-6645.

de Ruyter, P. G., O. P. Kuipers, M. M. Beerthuyzen, I. van Alen-Boerrigter & W. M. de Vos, (1996) Functional analysis of promoters in the nisin gene cluster of Lactococcus lactis. *J Bacteriol* 178: 3434-3439.

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

Diederichs, K., J. Diez, G. Greller, C. Muller, J. Breed, C. Schnell, C. Vonrhein, W. Boos & W. Welte, (2000) Crystal structure of MalK, the ATPase subunit of the trehalose/maltose ABC transporter of the archaeon Thermococcus litoralis. *EMBO J* 19: 5951-5961.

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

Dodd, H. M., N. Horn & M. J. Gasson, (1995) A cassette vector for protein engineering the lantibiotic nisin. *Gene* 162: 163-164.

Driessen, A. J., H. W. van den Hooven, W. Kuiper, M. van de Kamp, H. G. Sahl, R. N. Konings & W. N. Konings, (1995) Mechanistic studies of lantibiotic-induced permeabilization of phospholipid vesicles. *Biochemistry* 34: 1606-1614.

Dufour, A., T. Hindre, D. Haras & J. P. Le Pennec, (2007) The biology of lantibiotics from the lacticin 481 group is coming of age. *FEMS Microbiol Rev* 31: 134-167.

Eustaquio, A. S., B. Gust, U. Galm, S. M. Li, K. F. Chater & L. Heide, (2005) Heterologous expression of novobiocin and clorobiocin biosynthetic gene clusters. *Appl Environ Microbiol* 71: 2452-2459.

Flinspach, K., L. Westrich, L. Kaysser, S. Siebenberg, J. P. Gomez-Escribano, M. Bibb, B. Gust & L. Heide, (2010) Heterologous expression of the biosynthetic gene clusters of coumermycin A(1), clorobiocin and caprazamycins in genetically modified Streptomyces coelicolor strains. *Biopolymers* 93: 823-832.

Foulston, L. C. & M. J. Bibb, (2010) Microbisporicin gene cluster reveals unusual features of lantibiotic biosynthesis in actinomycetes. *Proc Natl Acad Sci U S A*.

Fredenhagen, A., G. Fendrich, F. Marki, W. Marki, J. Gruner, F. Raschdorf & 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: 1403-1412.

Gartemann, K. H., O. Kirchner, J. Engemann, I. Grafen, R. Eichenlaub & A. Burger, (2003) Clavibacter michiganensis subsp. michiganensis: first steps in the

understanding of virulence of a Gram-positive phytopathogenic bacterium. *J Biotechnol* 106: 179-191.

Gentschev, I. & W. Goebel, (1992) Topological and functional studies on HlyB of Escherichia coli. *Mol Gen Genet* 232: 40-48.

Giffard, C. J., H. M. Dodd, N. Horn, S. Ladha, A. R. Mackie, A. Parr, M. J. Gasson & D. Sanders, (1997) Structure-function relations of variant and fragment nisins studied with model membrane systems. *Biochemistry* 36: 3802-3810.

Gomez-Escribano, J. P. & M. J. Bibb, (2010) Engineering *Streptomyces coelicolor* for heterologous expression of secondary metabolite gene clusters. *Microbial Biotechnology*.

Gomez, A., M. Ladire, F. Marcille & M. Fons, (2002) Trypsin mediates growth phase-dependent transcriptional tegulation of genes involved in biosynthesis of ruminococcin A, a lantibiotic produced by a Ruminococcus gnavus strain from a human intestinal microbiota. *J Bacteriol* 184: 18-28.

Gravesen, A., B. Kallipolitis, K. Holmstrom, P. E. Hoiby, M. Ramnath & S. Knochel, (2004) pbp2229-mediated nisin resistance mechanism in Listeria monocytogenes confers cross-protection to class IIa bacteriocins and affects virulence gene expression. *Appl Environ Microbiol* 70: 1669-1679.

Guder, A., T. Schmitter, I. Wiedemann, H. G. Sahl & G. Bierbaum, (2002) Role of the single regulator MrsR1 and the two-component system MrsR2/K2 in the regulation of mersacidin production and immunity. *Appl Environ Microbiol* 68: 106-113.

Gust, B., G. L. Challis, K. Fowler, T. Kieser & K. F. Chater, (2003) PCR-targeted Streptomyces gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. *Proc Natl Acad Sci U S A* 100: 1541-1546.

Gust, B., G. Chandra, D. Jakimowicz, T. Yuqing, C. J. Bruton & K. F. Chater, (2004) Lambda red-mediated genetic manipulation of antibiotic-producing Streptomyces. *Adv Appl Microbiol* 54: 107-128.

Hata, D. J. & D. S. Smith, (2004) Blood group B degrading activity of Ruminococcus gnavus alpha-galactosidase. *Artif Cells Blood Substit Immobil Biotechnol* 32: 263-274.

Heidrich, C., U. Pag, M. Josten, J. Metzger, R. W. Jack, G. Bierbaum, G. Jung & 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: 3140-3146.

Holtsmark, I., D. Mantzilas, V. G. Eijsink & 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: 5814-5821.

Hong, H. J., M. I. Hutchings, L. M. Hill & M. J. Buttner, (2005) The role of the novel Fem protein VanK in vancomycin resistance in Streptomyces coelicolor. *J Biol Chem* 280: 13055-13061.

Hopwood, D. A., T. Kieser, H. M. Wright & M. J. Bibb, (1983) Plasmids, recombination and chromosome mapping in Streptomyces lividans 66. *J Gen Microbiol* 129: 2257-2269.

Horn, M. J., B. Jones & S. J. Ringel, (1941) Isolation of Mesolanthionine from various alkali-treated proteins. *J. Biol. Chem*: 141-149.

Hsu, S. T., E. Breukink, G. Bierbaum, H. G. Sahl, B. de Kruijff, R. Kaptein, N. A. van Nuland & A. M. Bonvin, (2003) NMR study of mersacidin and lipid II interaction in dodecylphosphocholine micelles. Conformational changes are a key to antimicrobial activity. *J Biol Chem* 278: 13110-13117.

Hsu, S. T., E. Breukink, E. Tischenko, M. A. Lutters, B. de Kruijff, R. Kaptein, A. M. Bonvin & N. A. van Nuland, (2004) The nisin-lipid II complex reveals a pyrophosphate cage that provides a blueprint for novel antibiotics. *Nat Struct Mol Biol* 11: 963-967.

Iwamoto, K., T. Hayakawa, M. Murate, A. Makino, K. Ito, T. Fujisawa & T. Kobayashi, (2007) Curvature-dependent recognition of ethanolamine phospholipids by duramycin and cinnamycin. *Biophys J* 93: 1608-1619.

Izaguirre, G. & J. N. Hansen, (1997) Use of alkaline phosphatase as a reporter polypeptide to study the role of the subtilin leader segment and the SpaT transporter in the posttranslational modifications and secretion of subtilin in Bacillus subtilis 168. *Appl Environ Microbiol* 63: 3965-3971.

Jones, G. & P. Dyson, (2006) Evolution of transmembrane protein kinases implicated in coordinating remodeling of gram-positive peptidoglycan: inside versus outside. *J Bacteriol* 188: 7470-7476.

Jung, G. & H.-G. Sahl, (1991) *Nisin and novel lantibiotics : proceedings of the First International Workshop on Lantibiotics, April 15-18, 1991, Physikzentrum Bad Honnef, F.R.G,* p. xvii, 490 p. ESCOM, Leiden.

Kern, D., B. F. Volkman, P. Luginbuhl, M. J. Nohaile, S. Kustu & D. E. Wemmer, (1999) Structure of a transiently phosphorylated switch in bacterial signal transduction. *Nature* 402: 894-898.

Kettenring, J. K., A. Malabarba, K. Vekey & B. Cavalleri, (1990) Sequence determination of actagardine, a novel lantibiotic, by homonuclear 2D NMR spectroscopy. *J Antibiot (Tokyo)* 43: 1082-1088.

Kieser, T., (2000) *Practical streptomyces genetics,* p. 613 p. John Innes Foundation, Norwich.

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

Kleerebezem, M., (2004) Quorum sensing control of lantibiotic production; nisin and subtilin autoregulate their own biosynthesis. *Peptides* 25: 1405-1414.

Kluskens, L. D., A. Kuipers, R. Rink, E. de Boef, S. Fekken, A. J. Driessen, O. P. Kuipers & G. N. Moll, (2005) Post-translational modification of therapeutic peptides by NisB, the dehydratase of the lantibiotic nisin. *Biochemistry* 44: 12827-12834.

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

Koponen, O., M. Tolonen, M. Qiao, G. Wahlstrom, J. Helin & 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: 3561-3568.

Kuipers, A., E. de Boef, R. Rink, S. Fekken, L. D. Kluskens, A. J. Driessen, K. Leenhouts, O. P. Kuipers & 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: 22176-22182.

Kuipers, A., J. Meijer-Wierenga, R. Rink, L. D. Kluskens & G. N. Moll, (2008) Mechanistic dissection of the enzyme complexes involved in the biosynthesis of lacticin 3147 and nisin. *Appl Environ Microbiol*.

Kuipers, O. P., M. M. Beerthuyzen, R. J. Siezen & W. M. De Vos, (1993) Characterization of the nisin gene cluster nisABTCIPR of Lactococcus lactis. Requirement of expression of the nisA and nisI genes for development of immunity. *Eur J Biochem* 216: 281-291.

Kupke, T., C. Kempter, G. Jung & 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: 11282-11289.

Levengood, M. R., G. C. Patton & W. A. van der Donk, (2007) The leader peptide is not required for post-translational modification by lacticin 481 synthetase. *J Am Chem Soc* 129: 10314-10315.

Li, B., J. P. Yu, J. S. Brunzelle, G. N. Moll, W. A. van der Donk & S. K. Nair, (2006) Structure and mechanism of the lantibiotic cyclase involved in nisin biosynthesis. *Science* 311: 1464-1467.

Machaidze, G. & J. Seelig, (2003) Specific binding of cinnamycin (Ro 09-0198) to phosphatidylethanolamine. Comparison between micellar and membrane environments. *Biochemistry* 42: 12570-12576.

MacNeil, D. J., K. M. Gewain, C. L. Ruby, G. Dezeny, P. H. Gibbons & T. MacNeil, (1992) Analysis of Streptomyces avermitilis genes required for avermectin biosynthesis utilizing a novel integration vector. *Gene* 111: 61-68.

Majchrzykiewicz, J. A., J. Lubelski, G. N. Moll, A. Kuipers, J. J. Bijlsma, O. P. Kuipers & 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: 1498-1505.

Makino, A., T. Baba, K. Fujimoto, K. Iwamoto, Y. Yano, N. Terada, S. Ohno, S. B. Sato, A. Ohta, M. Umeda, K. Matsuzaki & T. Kobayashi, (2003) Cinnamycin (Ro 09-0198) promotes cell binding and toxicity by inducing transbilayer lipid movement. *J Biol Chem* 278: 3204-3209.

Malabarba, A., R. Pallanza, M. Berti & B. Cavalleri, (1990) Synthesis and biological activity of some amide derivatives of the lantibiotic actagardine. *J Antibiot (Tokyo)* 43: 1089-1097.

Marchler-Bauer, A., J. B. Anderson, F. Chitsaz, M. K. Derbyshire, C. DeWeese-Scott, J. H. Fong, L. Y. Geer, R. C. Geer, N. R. Gonzales, M. Gwadz, S. He, D. I. Hurwitz, J. D. Jackson, Z. Ke, C. J. Lanczycki, C. A. Liebert, C. Liu, F. Lu, S. Lu, G. H. Marchler, M. Mullokandov, J. S. Song, A. Tasneem, N. Thanki, R. A. Yamashita, D. Zhang, N. Zhang & S. H. Bryant, (2009) CDD: specific functional annotation with the Conserved Domain Database. *Nucleic Acids Res* 37: D205-210.

Marcille, F., A. Gomez, P. Joubert, M. Ladire, G. Veau, A. Clara, F. Gavini, A. Willems & M. Fons, (2002) Distribution of genes encoding the trypsin-dependent lantibiotic ruminococcin A among bacteria isolated from human fecal microbiota. *Appl Environ Microbiol* 68: 3424-3431.

Maris, A. E., M. R. Sawaya, M. Kaczor-Grzeskowiak, M. R. Jarvis, S. M. Bearson, M. L. Kopka, I. Schroder, R. P. Gunsalus & R. E. Dickerson, (2002)

Dimerization allows DNA target site recognition by the NarL response regulator. *Nat Struct Biol* 9: 771-778.

Matthews, R. G. & C. W. Goulding, (1997) Enzyme-catalyzed methyl transfers to thiols: the role of zinc. *Curr Opin Chem Biol* 1: 332-339.

McArthur, M. & M. J. Bibb, (2008) Manipulating and understanding antibiotic production in Streptomyces coelicolor A3(2) with decoy oligonucleotides. *Proc Natl Acad Sci U S A* 105: 1020-1025.

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

Meyer, C., G. Bierbaum, C. Heidrich, M. Reis, J. Suling, M. I. Iglesias-Wind, C. Kempter, E. Molitor & 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: 478-489.

Mulders, J. W., I. J. Boerrigter, H. S. Rollema, R. J. Siezen & W. M. de Vos, (1991) Identification and characterization of the lantibiotic nisin Z, a natural nisin variant. *Eur J Biochem* 201: 581-584.

Nagao, J., S. M. Asaduzzaman, Y. Aso, K. Okuda, J. Nakayama & K. Sonomoto, (2006) Lantibiotics: insight and foresight for new paradigm. *J Biosci Bioeng* 102: 139-149.

Neis, S., G. Bierbaum, M. Josten, U. Pag, C. Kempter, G. Jung & H. G. Sahl, (1997) Effect of leader peptide mutations on biosynthesis of the lantibiotic Pep5. *FEMS Microbiol Lett* 149: 249-255.

Neuwald, A. F., L. Aravind, J. L. Spouge & E. V. Koonin, (1999) AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. *Genome Res* 9: 27-43.

Okeley, N. M., M. Paul, J. P. Stasser, N. Blackburn & W. A. van der Donk, (2003) SpaC and NisC, the cyclases involved in subtilin and nisin biosynthesis, are zinc proteins. *Biochemistry* 42: 13613-13624.

Pag, U. & H. G. Sahl, (2002) Multiple activities in lantibiotics--models for the design of novel antibiotics? *Curr Pharm Des* 8: 815-833.

Paget, M. S., L. Chamberlin, A. Atrih, S. J. Foster & 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: 204-211.

Parenti, F., H. Pagani & G. Beretta, (1976) Gardimycin, a new antibiotic from Actinoplanes. I. Description of the producer strain and fermentation studies. *J Antibiot (Tokyo)* 29: 501-506.

Patton, G. C., M. Paul, L. E. Cooper, C. Chatterjee & W. A. van der Donk, (2008) The Importance of the Leader Sequence for Directing Lanthionine Formation in Lacticin 481. *Biochemistry*.

Patton, G. C. & W. A. van der Donk, (2005) New developments in lantibiotic biosynthesis and mode of action. *Curr Opin Microbiol* 8: 543-551.

Paul, M., G. C. Patton & W. A. van der Donk, (2007) Mutants of the zinc ligands of lacticin 481 synthetase retain dehydration activity but have impaired cyclization activity. *Biochemistry* 46: 6268-6276.

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

Qiao, M., T. Immonen, O. Koponen & 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: 75-80.

Qiao, M. & P. E. Saris, (1996) Evidence for a role of NisT in transport of the lantibiotic nisin produced by Lactococcus lactis N8. *FEMS Microbiol Lett* 144: 89-93.

Relyea, H. A. & W. A. van der Donk, (2006) Nature's way to make the lantibiotics. *J Chem Educ* 83: 1769-1778.

Rince, A., A. Dufour, S. Le Pogam, D. Thuault, C. M. Bourgeois & J. P. Le Pennec, (1994) Cloning, expression, and nucleotide sequence of genes involved in production of lactococcin DR, a bacteriocin from lactococcus lactis subsp. lactis. *Appl Environ Microbiol* 60: 1652-1657.

Rink, R., L. D. Kluskens, A. Kuipers, A. J. Driessen, O. P. Kuipers & G. N. Moll, (2007) NisC, the cyclase of the lantibiotic nisin, can catalyze cyclization of designed nonlantibiotic peptides. *Biochemistry* 46: 13179-13189.

Rodriguez, E., Z. Hu, S. Ou, Y. Volchegursky, C. R. Hutchinson & R. McDaniel, (2003) Rapid engineering of polyketide overproduction by gene transfer to industrially optimized strains. *J Ind Microbiol Biotechnol* 30: 480-488.

Rogers, L. A., (1928) The Inhibiting Effect of Streptococcus Lactis on Lactobacillus Bulgaricus. *J Bacteriol* 16: 321-325.

Rogers, L. A. & E. O. Whittier, (1928) Limiting Factors in the Lactic Fermentation. *J Bacteriol* 16: 211-229.

Rutherford, K., J. Parkhill, J. Crook, T. Horsnell, P. Rice, M. A. Rajandream & B. Barrell, (2000) Artemis: sequence visualization and annotation. *Bioinformatics* 16: 944-945.

Sahl, H. G. & G. Bierbaum, (1998) Lantibiotics: biosynthesis and biological activities of uniquely modified peptides from gram-positive bacteria. *Annu Rev Microbiol* 52: 41-79.

Sambrook, J. & D. W. Russell, (2001) *Molecular cloning : a laboratory manual,* p. 1 v. (various pagings). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Schmid, D. G., F. Majer, T. Kupke & G. Jung, (2002) Electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry to reveal the substrate specificity of the peptidyl-cysteine decarboxylase EpiD. *Rapid Commun Mass Spectrom* 16: 1779-1784.

Schmitz, S., A. Hoffmann, C. Szekat, B. Rudd & G. Bierbaum, (2006) The lantibiotic mersacidin is an autoinducing peptide. *Appl Environ Microbiol* 72: 7270-7277.

Schneider, E. & S. Hunke, (1998) ATP-binding-cassette (ABC) transport systems: functional and structural aspects of the ATP-hydrolyzing subunits/domains. *FEMS Microbiol Rev* 22: 1-20.

Schnell, N., K. D. Entian, U. Schneider, F. Gotz, H. Zahner, R. Kellner & G. Jung, (1988) Prepeptide sequence of epidermin, a ribosomally synthesized antibiotic with four sulphide-rings. *Nature* 333: 276-278.

Siegers, K. & K. D. Entian, (1995) Genes involved in immunity to the lantibiotic nisin produced by Lactococcus lactis 6F3. *Appl Environ Microbiol* 61: 1082-1089.

Siegers, K., S. Heinzmann & K. D. Entian, (1996) Biosynthesis of lantibiotic nisin. Posttranslational modification of its prepeptide occurs at a multimeric membrane-associated lanthionine synthetase complex. *J Biol Chem* 271: 12294-12301.

Siezen, R. J., O. P. Kuipers & W. M. de Vos, (1996) Comparison of lantibiotic gene clusters and encoded proteins. *Antonie Van Leeuwenhoek* 69: 171-184.

Simon, R., U. Priefer & A. Puhler, (1983) A Broad Host Range Mobilization System for In Vivo Genetic Engineering: Transposon Mutagenesis in Gram Negative Bacteria. *Nat Biotech* 1: 784-791.

Sola, M., F. X. Gomis-Ruth, L. Serrano, A. Gonzalez & M. Coll, (1999) Three-dimensional crystal structure of the transcription factor PhoB receiver domain. *J Mol Biol* 285: 675-687.

Somma, S., W. Merati & F. Parenti, (1977) Gardimycin, a new antibiotic inhibiting peptidoglycan synthesis. *Antimicrob Agents Chemother* 11: 396-401.

Stein, T., S. Borchert, P. Kiesau, S. Heinzmann, S. Kloss, C. Klein, M. Helfrich & K. D. Entian, (2002) Dual control of subtilin biosynthesis and immunity in Bacillus subtilis. *Mol Microbiol* 44: 403-416.

Stein, T., S. Heinzmann, S. Dusterhus, S. Borchert & 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: 822-828.

Stein, T., S. Heinzmann, I. Solovieva & 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: 89-94.

Sun, Z., J. Zhong, X. Liang, J. Liu, X. Chen & L. Huan, (2009) Novel mechanism for nisin resistance via proteolytic degradation of nisin by the nisin resistance protein NSR. *Antimicrob Agents Chemother* 53: 1964-1973.

Terrak, M., E. Sauvage, A. Derouaux, D. Dehareng, A. Bouhss, E. Breukink, S. Jeanjean & M. Nguyen-Disteche, (2008) Importance of the conserved residues in the peptidoglycan glycosyltransferase module of the class A penicillin-binding protein 1b of Escherichia coli. *J Biol Chem* 283: 28464-28470.

van de Kamp, M., L. M. Horstink, H. W. van den Hooven, R. N. Konings, C. W. Hilbers, A. Frey, H. G. Sahl, J. W. Metzger & F. J. van de Ven, (1995) Sequence analysis by NMR spectroscopy of the peptide lantibiotic epilancin K7 from Staphylococcus epidermidis K7. *Eur J Biochem* 227: 757-771.

van den Berg van Saparoea, H. B., P. J. Bakkes, G. N. Moll & A. J. Driessen, (2008) Distinct contributions of the nisin biosynthesis enzymes NisB and NisC and transporter NisT to prenisin production by Lactococcus lactis. *Appl Environ Microbiol* 74: 5541-5548.

van den Hooven, H. W., H. S. Rollema, R. J. Siezen, C. W. Hilbers & O. P. Kuipers, (1997) Structural features of the final intermediate in the biosynthesis of the lantibiotic nisin. Influence of the leader peptide. *Biochemistry* 36: 14137-14145.

van der Meer, J. R., H. S. Rollema, R. J. Siezen, M. M. Beerthuyzen, O. P. Kuipers & W. M. de Vos, (1994) Influence of amino acid substitutions in the nisin leader peptide on biosynthesis and secretion of nisin by Lactococcus lactis. *J Biol Chem* 269: 3555-3562.

van Kraaij, C., E. Breukink, M. A. Noordermeer, R. A. Demel, R. J. Siezen, O. P. Kuipers & B. de Kruijff, (1998) Pore formation by nisin involves translocation of its C-terminal part across the membrane. *Biochemistry* 37: 16033-16040.

Vertesy, L., W. Aretz, A. Bonnefoy, E. Ehlers, M. Kurz, A. Markus, M. Schiell, M. Vogel, J. Wink & H. Kogler, (1999) Ala(0)-actagardine, a new lantibiotic from cultures of Actinoplanes liguriae ATCC 31048. *J Antibiot (Tokyo)* 52: 730-741.

Weigel, L. M., D. B. Clewell, S. R. Gill, N. C. Clark, L. K. McDougal, S. E. Flannagan, J. F. Kolonay, J. Shetty, G. E. Killgore & F. C. Tenover, (2003) Genetic analysis of a high-level vancomycin-resistant isolate of Staphylococcus aureus. *Science* 302: 1569-1571.

Widdick, D. A., H. M. Dodd, P. Barraille, J. White, T. H. Stein, K. F. Chater, M. J. Gasson & 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: 4316-4321.

Wiedemann, I., T. Bottiger, R. R. Bonelli, A. Wiese, S. O. Hagge, T. Gutsmann, U. Seydel, L. Deegan, C. Hill, P. Ross & 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: 285-296.

Xie, L., L. M. Miller, C. Chatterjee, O. Averin, N. L. Kelleher & W. A. van der Donk, (2004) Lacticin 481: in vitro reconstitution of lantibiotic synthetase activity. *Science* 303: 679-681.

Yanai, K., T. Murakami & M. Bibb, (2006) Amplification of the entire kanamycin biosynthetic gene cluster during empirical strain improvement of Streptomyces kanamyceticus. *Proc Natl Acad Sci U S A* 103: 9661-9666.

Yeats, C., R. D. Finn & A. Bateman, (2002) The PASTA domain: a beta-lactam-binding domain. *Trends Biochem Sci* 27: 438.

You, Y. O. & W. A. van der Donk, (2007) Mechanistic investigations of the dehydration reaction of lacticin 481 synthetase using site-directed mutagenesis. *Biochemistry* 46: 5991-6000.

Zimmermann, N. & G. Jung, (1997) The three-dimensional solution structure of the lantibiotic murein-biosynthesis-inhibitor actagardine determined by NMR. *Eur J Biochem* 246: 809-819.

Zimmermann, N., J. W. Metzger & G. Jung, (1995) The tetracyclic lantibiotic actagardine. 1H-NMR and 13C-NMR assignments and revised primary structure. *Eur J Biochem* 228: 786-797.