University of Alberta

Characterization of the gene cluster encoding a non-ribosomal peptide synthetase for polymyxin biosynthesis in *Paenibacillus polymyxa* PKB1

by

Md. Shaheen

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science
in
Microbiology and Biotechnology

Department of Biological Sciences

Spring 2011
Edmonton, Alberta

Examining Committee

Dr. Susan E. Jensen, Department of Biological Sciences, University of Alberta

Dr. Julia M. Foght, Department of Biological Sciences, University of Alberta

Dr. Jonathan J. Dennis, Department of Biological Sciences, University of Alberta

Dr. John C. Vederas, Department of Chemistry, University of Alberta

Abstract

Polymyxins are cationic lipopeptide antibiotics, and are considered to be drugs of last resort to fight against multidrug resistant Gram negative bacteria. In this study the gene cluster encoding the non-ribosomal peptide synthetase which catalyses polymyxin biosynthesis was identified and sequenced in *Paenibacillus polymyxa* PKB1. The gene cluster consists of three biosynthetic genes and two ABC transporter-like genes. Disruption of a biosynthetic gene resulted in complete loss of antibiotic production whereas deletion of one or both transporter genes resulted in reduced antibiotic production. Two antibacterial compounds were identified in the wild type culture supernatants by HPLC analysis and agar diffusion bioassay. LC-MS and bioinformatic analysis suggested these compounds to be analogs of polymyxin B_2 and B_1, containing a D-DAB instead of L-DAB residue in position three of the molecules. (129 words)

Table of Contents

List of Tables

Page

List of Figures

Page

List of Abbreviations

A domain	Adenylation domain
aa	Amino acid
aac(3)IV	Apramycin resistance gene
Ala	Alanine
amp	Ampicillin resistance gene
apra	Apramycin
Asn	Asparagine
ATP	Adenosine triphosphate
Blast	Basic local alignment search tool
bp	Base pair
C domain	Condensation domain
cat	Chloramphenicol resistance gene
Cm	Chloramphenicol
Cy	Cyclization domain
D-	D- stereochemistry
D	Single letter code for Aspartic acid
DAB	2, 4 diamino butyric acid
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleoside triphosphate
E domain	Epimerization domain
EDTA	Ethylenediaminetetraacetic acid
FLP	Flip Recombinase
FRT	FLP recognition target
G	Guanine
GB	Glucose broth
Glu	Glutamic acid
HPLC	High performance liquid chromatography

Ile	Iso-leucine
IPTG	Isopropyl-β-D-thio-galactopyranoside
kan	Kanamycin
kb	Kilobase
kDa	kilo Dalton
kV	Kilo Volt
L-	L- stereochemistry
LB	Luria-Bertani broth
LC-MS	Liquid chromatography-mass spectrometry
Leu	Leucine
MIC	Minimum Inhibitory Concentration
MT	Methylation domain
NRPS	Non ribosomal peptide synthetase
NRP	Nonribosomal peptide
OD	Optical density
ORF	Open reading frame
Orn	Ornithine
oriT	Origin of transfer
Ox	Oxidation domain
PCR	Polymerase chain reaction
Phe	Phenyl alanine
PPi	Pyrophosphate
^{R}	Resistant
rpm	Rotations per minute
S	Serine
SAP	Shrimp alkaline phosphatase
SOB	Super optimal broth
SSC	Standard saline citrate
T domain	Thiolation domain
TAE	Tris-acetate-EDTA buffer
TE	Thioesterase or termination domain
Thr	Threonine

UV	Ultraviolet
w/v	Weight/volume
X-gal	5-bromo-4-chloro-3-indoyl-β-D- galactopyranoside
ZOI	Zone of inhibition
α	Alpha
β	Beta
Δ	Delta/ Deletion
λ	Lambda
μ	Micro

1. Introduction

Polymyxins are cationic lipopeptide antibiotics produced by Gram positive bacteria via non-ribosomal peptide synthetase (NRPS) mechanisms. Polymyxin was first isolated in 1947 from *Bacillus polymyxa* (Stansly and Schlosser 1947). At the same time polymyxin was reported to be isolated from *Bacillus aerosporus* by Ainsworth and co-workers and they named it aerosporin. 'Polymyxin', the name given by Stansly and Shepherd was taken as the official name for this group of antibiotics at a symposium of the New York Academy of Sciences in 1948 (Vogler and Studer 1966). The chemical composition and structure of polymyxin B_1 was first elucidated in 1954 by Hausmann and Craig (Hausmann and Craig 1954). Later in 1961 Vogler confirmed the structure (Paulus and Gray 1964; Wilkinson and Lowe 1964).

Although polymyxins are very potent antibiotics, they have not been in routine use due to their inherent toxicity to the human host and the availability of other active and less toxic antibiotics. However, the emergence of multidrug resistant Gram negative bacteria and their intrinsic susceptibility to polymyxins led to a reconsideration of these antibiotics. Although many studies have been reported on the chemical and therapeutic properties of polymyxins, studies on molecular mechanisms of their biosynthesis are still very limited. Only recently whole genome sequences of two *Paenibacillus polymyxa* strains have been made public and one of them has been reported to have polymyxin producing NRPS genes and produce polymyxin (Choi et al. 2009).

1.1 Polymyxin-producing bacteria

Polymyxins are mainly produced by different species of *Bacillus* (Storm et al. 1977). The first polymyxin reported to be isolated was from *B. polymyxa* (Stansly and Schlosser 1947). A number of other *Bacillus* species have been reported to produce different polymyxins, for example polymyxin E by *B. colistinus* (Evans et al. 1999) and *Paenibacillus amylolyticus* (DeCrescenzo Henriksen et al. 2007), polymyxin M by *Paenibacillus kobensis* (Martin et al. 2003) and polymyxins by *P. polymyxa* (Li et al. 2007). The polymyxin producing *Bacillus* strains were renamed *Paenibacillus* in 1993 based on their 16S rRNA gene sequence similarity (Ash et al. 1993; Jurtshuk et al. 1992). Other than polymyxins, strains of *Paenibacillus* produce many other peptides and lipopeptides with antimicrobial properties (Storm et al. 1977). As soil organisms, these antimicrobial metabolites likely provide *Paenibacillus* spp. with a growth advantage in the competitive soil environment. Some of the strains of *Paenibacillus* also display antagonistic activity against plant pathogenic fungi by producing various degradative enzymes and antifungal antibiotics. *P. polymyxa* was reported to produce fusaricidin, an antifungal antibiotic active against *Leptosphaeria maculans*, the causative agent of blackleg disease of Canola (Beatty and Jensen 2002).

1.2 Chemistry of polymyxins

Like typical non-ribosomal lipopeptides, polymyxins have both L- and D-amino acids. Polymyxins consist of 10 amino acids with a characteristic

polycationic heptapeptide ring. A typical structure of polymyxin B is shown in Figure 1.2.1. The peptide ring is linked to the fatty acid side chain by a linear tripeptide chain that forms an amide bond with a fatty acid moiety. 2, 4-Diaminobutyric acid (DAB) is a major building block of polymyxins and there are six DAB residues present in each molecule. Initial structure determination of polymyxin B suggested that one of the six DAB residues is a D isomer and most probably it was the amino acid at position 1 (Hausmann and Craig 1954). Later, it was proven that all six DAB residues are in the L configuration in polymyxin B (Paulus and Gray 1964; Wilkinson and Lowe 1964). Most recently, the same result was published for polymyxin M (Mattacin) produced by *P. kobensis* (Martin et al. 2003). In contrast, the first genetic characterization of a polymyxin synthetase gene cluster suggested that the source organism, *P. polymyxa* E681, must produce a polymyxin in which the third DAB residue is in D form, and hence they identified it as a different form of polymyxin, polymyxin A (Choi et al. 2009).

In all polymyxins the peptide ring is formed through cyclization of the carboxyl group of L-Thr at the position 10 with the α-amino group of the DAB residue at position 4, and the γ-amino group of this DAB residue is attached to the tripeptide side chain (Storm et al. 1977). The molecular weights of polymyxins vary among the different types with a range from 1129 to 1202 Da. The fatty acid side chains of polymyxins are usually heptanoic acid or octanoic acid or their methyl derivatives (Orwa et al. 2001; Price et al. 2007). Nonanoic acid as a fatty acid side chain has also been reported for some polymyxins (Govaerts et al.

2002b). The charged peptide moiety and the hydrophobic fatty acid tail give polymyxins an amphipathic nature. This amphipathic property allows them to disperse homogenously into both aqueous and non-aqueous media which leads to a high solubility in body fluids as well as in prokaryotic and eukaryotic lipid membranes (Hermsen et al. 2003).

1.3 Types of polymyxin

All of the polymyxins can be divided into five major groups and each group is divided into many sub-groups mainly based on variations in the fatty acid side chain. The major groups of polymyxin are named Polymyxin A, B, D, E (E is also known as colistin) (Stansly and Brownlee 1949), and M (known as mattacin) (Martin et al. 2003). All of the polymyxins share a common basic structure. They vary by the amino acid composition in certain positions of the peptide ring and by the fatty acid side chain length. Most of the variations have been described for the amino acids in position 6 and 7, among the types of polymyxins (Storm et al. 1977). The different types of polymyxins and their chemical compositions are given in Table 1.3.1.

1.4 Therapeutic properties and use of polymyxins

1.4.1 Spectrum and mode of action of polymyxins

Polymyxins are mainly effective against Gram negative bacteria, and most Enterobacteriaceae and non-fermentative species are found to be susceptible to polymyxins. Polymyxins generally do not have any effect on Gram-positive

bacteria (Evans et al. 1999; Hermsen et al. 2003). Although it is not apparent that the mechanism of action of polymyxin has anything to do with the presence of oxygen, anaerobes are resistant to polymyxins. *Pseudomonas aeruginosa* and *Acinetobacter* species highly resistant to most antibiotics were found to be intrinsically sensitive to polymyxin B (Gales et al. 2006). Polymyxins (Colistin) were also found to be active against *Haemophilus influenzae*, *Bordetella pertussis* and *Legionella pneumophila* (Li et al. 2005).

Although it is more than half a century since the polymyxins were discovered, their exact mechanism of action is not yet fully understood. Their degree of activity varies with the concentration; usually polymyxins are bacteriostatic in low concentrations and bacteriocidal at high concentrations. Polymyxins have a high affinity for lipopolysaccharide (LPS) of Gram negative bacteria, can bind to it strongly, and then cause permeability changes to the inner membrane, eventually resulting in cell death (Storm et al. 1977). Polymyxin molecules have a net charge of +5 which is greater than that of the Ca^{2+} and Mg^{2+} present in the LPS. Therefore, polymyxins bind to the LPS by competitively displacing these two ions which usually play a role in stabilizing the LPS in outer membrane. These initial electrostatic interactions stabilize the transient LPS-polymyxin complex, and allow the non-polar fatty acid side chain of polymyxin to come into close proximity with the outer membrane and interact further. This amphipathic interaction of polymyxin helps it to be deeply inserted into the outer membrane (Clausell et al. 2007; Hancock 1997; Hancock and Lehrer 1998; Schindler and Osborn 1979; Velkov et al. 2010). This overall effect of polymyxin

can be considered as a detergent-like action which produces a physico-chemical disruption of the cell envelope (Hermsen et al. 2003). Recent studies on the structure of polymyxin-LPS complexes by nuclear magnetic resonance (NMR) spectroscopy and isothermal titration calorimetry (ITC) illustrated in details the electrostatic and hydrophobic interactions between them. Upon initial contact of polymyxin with the LPS, the insertion of the fatty acid side chain and the interaction of the hydrophobic part of the peptide ring (usually the position 6 and 7 amino acids) loosen the packing of the fatty acid chains of adjacent lipid A molecules in the outer membrane. This in turn, leads to the loss of integrity of the outer membrane (Velkov et al. 2010). Polymyxins subsequently pass across the outer membrane by a hypothetical process called 'self-promoted uptake' (Hancock 1997; Hancock 1984). After passing through the outer membrane, polymyxins bind and disrupt the physical barrier of the phospholipid bi-layer of the inner membrane, leading to the formation of transient pores. This results in osmotic imbalance and eventually cell lysis (Clausell et al. 2007; Hancock 1997; Velkov et al. 2010). Other cell perturbing molecules including other antibiotics or toxic products may get into the cells through these transient pores and accelerate the process of cell death (Hancock 1997).

1.4.2 Resistance to polymyxins

Resistance to polymyxins can be acquired through mutation or adaptation. Most of the resistance studies have been carried out focusing on *P. aeruginosa*. Bacteria resistant to one type of polymyxin typically also exhibit resistance to

other types due to their structural homology (Evans et al. 1999). Since LPS is the anchoring point for polymyxins, most of the resistance mechanisms to polymyxins are achieved through modification of the LPS moiety especially the lipid A of the outer membrane (Zavascki et al. 2007). In *Salmonella enterica* serovar Typhimurium two two-component systems were reported to be involved in the resistance to polymyxins. The PhoP/PhoQ two-component system can sense low concentrations of Mg^{2+} in the outer membrane and become activated, and then activate the PmrA/PmrB two component system which then regulates the modification of lipid A with 4-amino-4-deoxy-L-arabinose or phosphoethanolamine. This modification leads to a reduction in net charge of the outer membrane, thus reduce the affinity for polymyxins (Delgado et al. 2006; Gunn et al. 1998; Gunn et al. 2000). The PmrA/PmrB two-component system is also regulated by the presence of Fe^{3+} and low pH (Kato et al. 2003; Perez and Groisman 2007). The genes required for the biosynthesis of 4-amino-4-deoxy-L-arabinose and its transfer to lipid A are usually present on an operon, *pmr,* containing genes *pmrHFILKJM*. This type of modification of lipid A has also been reported for other Gram-negative bacteria such as *P. aeruginosa, Burkholderia cepacia,* and *Yersinia pestis* (Falagas et al. 2010).

In *Burkholderia cenocepacia,* two different resistance mechanisms have been described for cationic lipopeptide antibiotics including polymyxins. A full length LPS containing heptose (reduced negative charge) is the first line of defense against polymyxins. The second line of defense is not associated with the membrane permeability. Rather it involves proteins preliminarily identified as

some stress response regulatory proteins and periplasmic proteases that act independently when the outer membrane is disrupted, although their mechanisms of action have not yet been identified (Loutet et al. 2010).

Vibrio cholerae has been reported to have a porin, OmpU, regulated by ToxR, which confers resistance to cationic antibiotics including polymyxin B sulphate (Mathur and Waldor 2004). The development of an efflux pump/ potassium system has also been reported for *Yersinia* spp. linked to the resistance to polymyxins (Bengoechea and Skurnik 2000).

Finally, *Klebsiella pneumoniae* showed resistance to polymyxin by simply overproducing capsular polysaccharide, providing a physical barrier between the cell surface and the antibiotic (Campos et al. 2004).

1.4.3 Clinical use of polymyxins

Polymyxin B and polymyxin E are the two polymyxins that have been used clinically; polymyxin E (colistin) was more widely used than polymyxin B (Li et al. 2006). Colistin methane sulfonate was first introduced clinically in 1962 to treat an infection caused by *Pseudomonas pyocyanea* (Evans et al. 1999). This methane sulfonate derivative of colistin was also used successfully to treat a *P. aeruginosa* infection in cystic fibrosis patients (Jensen et al. 1987). A parenteral formulation of polymyxin B is available as the sulphate salt although the intramuscular injection is no longer recommended because it causes severe pain at the site of administration. Polymyxin B sulphate is also available in combination with other antibiotics in topical, ophthalmic and otic formulations and the topical

formulation is used to prevent rather than to treat infections. Colistin is available as both sulphate and sulfomethyl salts but the sulphate salt of colistin is not recommended for parenteral use because of its higher toxicity compared to the sulfomethyl salt. Nebulized colistin is also recommended for chronic cystic fibrosis patients (Hermsen et al. 2003). Bacterial infections with carbapenem-resistant Enterobacteriaceae and *A. baumannii* remain a leading cause of post-transplantation morbidity and mortality. Polymyxin B is one of the few antibiotics available to treat carbapenem-resistant Gram-negative infections successfully (Patel et al. 2010). Recently carbapenemase-producing *K. pneumoniae* isolates have been reported to have higher polymyxin B MICs. However, a combination of polymyxin B with rifampin or doxicycline appeared to reduce the polymyxin MIC for this bacterium (Elemam et al. 2010). In the treatment of sepsis, it has been proven that removing LPS from the patient is the best option compared to the others because release of LPS is an early step which triggers the process of sepsis. In Japan polymyxin B is used as an embedded component in hemoperfusion cartridges which are used to remove LPS from the blood to block the initial process of sepsis (Rachoin et al. 2010).

1.4.4 Toxicity of polymyxins

Neurotoxicity and nephrotoxicity are the two major toxicities associated with polymyxin treatments (Li et al. 2006). Neurotoxicity due to polymyxins is usually less frequent than nephrotoxicity and is less severe. Both types of toxicities are dose dependent and usually resolve with the discontinuation of

treatment (Evans et al. 1999). The neurotoxicity of polymyxin is characterized by dizziness, muscular weakness, facial and peripheral paraesthesia, partial deafness, visual disturbances, vertigo, hallucinations and ataxia. Polymyxins interfere with the receptor sites of neurons and block the release of acetylcholine to the synaptic gap causing neuromuscular blockade. The neuromuscular blockade, in very severe cases, can cause apnea or respiratory failure, although these complications have not been observed in the last 15 years (Falagas and Kasiakou 2006). The common form of nephrotoxicity is damage to the convoluted tubule epithelium of the kidney. Acute tubular necrosis is associated with albuminuria, renal insufficiency, and increased serum creatinine and blood urea (Evans et al. 1999).

1.4.5 Non-therapeutic use of polymyxins

Polymyxins have been used as a standard in assays for detecting LPS binding affinity of many different compounds. At present, polymyxin integrated electrodes have been generated to detect LPS in many different samples *in situ*. Ferrocene-attached polymyxin B (PMB-Fc) can be used for the electrochemical detection of LPS at a concentration as low as 50 ng/ ml with a detection time of 5 min (Iijima et al. 2010). Hyperbranched poly (N-*iso*propyl acrylamide) (HB-PNIPAM) polymers, when conjugated with the antibiotic polymyxin-B (HB-PNIPAM-pmx), have been proposed for use as a dressing material to selectively remove Gram negative bacteria to reduce the bacterial load in wounds. This material is better than the currently available options in some aspects and the

inclusion of polymyxins added the ability to bind to the LPS of Gram negative bacteria (Shepherd et al. 2010).

1.4.6 Prospects of polymyxins

The polymyxins had largely fallen out of favor as chemotherapeutic agents due to their toxicity to the host, as suggested by clinical reports in the early 1970s. However, recent studies suggested that the toxicities associated with polymyxins might be less severe and less frequent than earlier reported (Arnold et al. 2007). The emergence of multi-drug resistant Gram-negative bacteria and their resistance to all currently available antibiotics except polymyxins has led to a re-evaluation of polymyxins as potential chemotherapeutic agents. Several outbreaks caused by multidrug resistant *P. aeruginosa* and *A. baumannii* have been reported worldwide recently and these bacteria have been found to be sensitive only to polymyxins (Bonomo and Szabo 2006; Go et al. 1994). Therefore polymyxins are now considered to be a drug of last resort to treat infections caused by these multidrug resistant Gram-negative bacteria. In one study it was found that the bacteria were cleared in 88% of patients infected with *P. aeruginosa* and *A. baumannii* after polymyxin B treatment, but 14% of the patients developed renal failure (Ouderkirk et al. 2003). Many studies have been reported to reduce the toxicity of polymyxins. At the very beginning of chemotherapeutic applications of polymyxin, the toxicity was reduced by making their sulphate or methane sulfonate salts. Polymyxins conjugated with high molecular weight polysaccharides like dextran 70 were reported to have reduced toxicity, with the

same endotoxin-neutralizing ability (Bucklin et al. 1995). Because the fatty acid side chain is considered to be responsible for polymyxin toxicity, a new approach has been taken to reduce polymyxin toxicity by producing polymyxin nonapeptide. The polymyxin B nonapeptide (PMBN) is a truncated polymyxin molecule having nine amino acids and no fatty acid side chain. PMBN has no antimicrobial activity itself, but it can interact with bacterial LPS (referred to as sensitizing activity). PMBN can be used as a combination therapy with other antibiotics as it helps other antibiotics to gain entry to the cells, and thus facilitates the ability of the other antibiotics to act on the cells (Tsubery et al. 2002).

Similar attempts have been reported by many other research groups to make analogs of polymyxin. Polymyxin B and E derivatives carrying three cationic charges instead of five have been generated and found to have bacterial sensitizing activity with a lower affinity to the host's kidney brush border membrane. These analogs can be used with other antibiotics with the reduced risk of nephrotoxicity. It has also been seen that, without compromising the net charge, removing DAB residues from the peptide ring and re-introducing the DAB residues in the side chain reduces the antibacterial activity of polymyxins. It has also been seen that changing only one amino acid in the second position of the polymyxin molecule can decrease the MIC by up to eight fold. This indicates that the changes in the amino acid compositions and their positions have a great effect on the biochemical properties of polymyxin (Vaara 2010; Vaara and Vaara 2010).

1.5 Non-ribosomal peptide synthetases

Non-ribosomal peptide synthetases (NRPS) produce a variety of peptide secondary metabolites in microorganisms like bacteria and fungi. These non-ribosomally synthesized peptide metabolites are very diverse in their chemical compositions, structures and biological activities. These peptides help the producing organisms to survive in adverse conditions like low nutrient habitats by accumulating nutrients or by allowing to them compete more successfully with other microorganisms by being toxic to them. A significant number of non-ribosomal peptides have great therapeutic value because they have antimicrobial activity and can be used as antibiotics and many of them are immunosuppressant. In fact, penicillin, the first antibiotic discovered, is produced by a NRPS mechanism, although the non-ribosomal mechanism of peptide synthesis was only reported in 1970 by Lipmann and co-workers (Lipmann 1971; Lipmann et al. 1971; Lipmann 1973).

One of the hallmarks of non-ribosomally synthesized peptides is the presence of non-proteinogenic residues like (D)-amino acids. More than 300 different building blocks can be incorporated into non-ribosomal peptides, whereas ribosomal peptide synthesis is limited to the standard 20 amino acids (Marahiel 1992). This wider range of building components provides endless possibilities of producing peptides with remarkable structural diversity by the NRPS mechanisms. Polymyxins have D-leucine or D-phenylalanine in the 6th amino acid position in their structure and the non-proteinogenic amino acid di-amino butyrate (DAB) in positions 1, 3, 4, 5, 8 and 9 (Figure 1.2.1).

NRPSs are multimodular enzyme systems in which the modules are organized to act as an assembly line and each module is responsible for the incorporation of one amino acid into the peptide product it generates (Marahiel et al. 1997; von Dohren et al. 1997). Schematic representation of the module and domain organization of a NRPS is given in Figure 1.5. 1. According to the assembly mechanisms used to produce peptide products, NRPS systems have been classified into three groups (Mootz et al. 2002). Type A NRPSs are designated as linear NRPSs and are analogous to type-I polyketide synthases (PKSs) (Fischbach and Walsh 2006; Staunton and Weissman 2001). In type A NRPS, each module is used only once and catalyzes only one round of peptide chain elongation to produce a fully assembled linear peptide product. The number and order of the substrate amino acids in the final peptide product can be predicted from the amino acid sequence of the type A NRPS since the order of the amino acids in the NRP is co-linear with the modules of its producing NRPS. The structural organization of the NRPSs will be discussed in detail in the next section. Type B NRPS systems are designated as iterative NRPSs. As the name implies, in this system the modules or domains are used more than once in the assembly of a single peptide product. Type B NRPSs first produce repeated short identical peptide sequences and then assemble these short sequences to make a final peptide product which is a multimer of the initial short sequences. Therefore the number of modules in the type B NRPS is equal to the number of amino acids in one initial short peptide sequence. Type C NRPSs are designated as non-linear NRPSs and have at least one unusual arrangement in their modules or core

domains. Type C NRPSs also differ from the type A NRPSs by the fact that they catalyze an unusual internal cyclization in the final peptide product or carry out a branch-point peptide biosynthesis. Specialized condensation or ligase domains that incorporate small molecules such as amines into the non-ribosomally assembled peptide can also be found in this type of NRPS (Mootz et al. 2002). Since the polymyxin molecule has a peptide ring, it can be assumed to be an example of a peptide that is produced by a type C NRPS.

1.6 Catalytic domains of NRPSs

There are three main domains, the adenylation (A) domain, the peptidyl carrier protein (PCP) domain and the condensation (C) domain required for the non-ribosomal biosynthesis of a peptide chain. Together they form a module named the minimal elongation module (Lautru and Challis 2004). The initiation (first) module of a NRPS usually does not contain a C domain, since it incorporates the first amino acid and there is not yet an acceptor amino acid to form the peptide bond. The growing amino acid chain is usually tethered to the PCP domain and to release the fully assembled peptide from the last module, an additional domain named the thioesterase (TE) domain or termination domain is usually found following the PCP domain of the last module. The TE domain catalyses a nucleophilic hydrolase reaction that leads to the release of a linear peptide. It can also catalyse an internal nucleophilic reaction resulting in the release of a peptide with concomitant intramolecular cyclization (Challis and Naismith 2004). Since non-ribosomal peptides are structurally very diverse, the

peptide synthetases can also contain additional optional domains along with the basic peptide elongation domains. These additional domains are present in many modules and modify the peptide products to give additional structural features like branched chains or cyclic structures or they can result in addition of certain other chemical groups to the peptide chain. Epimerization (E) domains are seen in certain modules following the PCP domain at the C terminus of the module. These E domains catalyse epimerisation reactions to convert (L) amino acids to their (D) forms either before or after peptide bond formation takes place. Cy domains can be found in some modules instead of the C domains. Cy domains catalyze a condensation reaction as well as the intramolecular heterocyclisation of serine, cysteine or threonine residues. Methyltransferase (MT) domains are present in the A domains in some NRPSs. The MT domain is responsible for the methylation of the amine group of many amino acids in non-ribosomally synthesized peptides. All of the domains in a module are connected through flexible linker regions.

Fatty acid synthases and polyketide synthases are usually found in homodimeric forms and the catalytic activity of the domains depends on the formation of this functional dimer. Homodimeric forms of domains in NRPSs have not been observed so far. However the first report of heterodimer formation by a NRPS domain came from the VibF of vibriobactin synthesis. VibF can act on acyl chains *in trans* as well as *in cis* and therefore suggests a different dimension of diversity in peptide biosynthesis (Hillson and Walsh 2003).

1.6.1 Adenylation domain

The chemical principle of non-ribosomal peptide biosynthesis is given in Figure 1.6.1.1. The adenylation (A) domain activates the appropriate amino acid as an aminoacyl adenylate by selecting it from the available substrate options (Dieckmann et al. 1995). This is the very first step in non-ribosomal peptide biosynthesis. The activation of an amino acid as an aminoacyl adenylate occurs with the hydrolysis of an ATP in the presence of Mg^{2+} (Dieckmann et al. 1995). The A domain shows diversity in substrate recognition and so plays a key role in the activation of both proteinogenic and non-proteinogenic amino acid residues. The direct activation of (D)-amino acids has even been reported for a few NRPSs (Li and Jensen 2008; Yin and Zabriskie 2006).

The A domain is a member of the adenylate-forming enzyme superfamily (Challis and Naismith 2004) and the enzymes of this superfamily share a common group of motifs, described as AMP-binding motifs (Bairoch and Boeckmann 1991). Therefore this superfamily of enzymes is also called an AMP-binding protein family and has been subdivided into four classes based on sequence similarities. Class I mainly consists of the eukaryotic acyl-CoA-synthetases, class II enzymes are mostly associated with non-ribosomal peptide synthetases, and acetyl-CoA-synthetases are grouped into class III. Class IV is the most structurally and functionally diverse group and it contain all other AMP-binding proteins including luciferases and other acyl-CoA synthetases of prokaryotic and plant origin (Fulda et al. 1994). The class II enzymes differ from the others by the fact that they do not transfer the respective activated amino acyl intermediates to

CoA (Pavela-Vrancic et al. 1994). In this case, an intra-enzymatical transfer of the activated amino acid to the thiol group of a 4'-phospho-pantetheine (ppan) cofactor covalently attached to carrier protein occurs and from there it undergoes a directed condensation reaction (Stein et al. 1994).

The crystal structures of four different A domains have been determined so far. These include the phenylalanine-activating A domain (PheA) present in the first module of the gramicidin S synthetase A of *Bacillus brevis* (Conti et al. 1997), the 2,3-dihydroxybenzoate (DHB)-activating A (DhbE) domain of *Bacillus subtilis* (May et al. 2002) and the D-alanyl carrier protein ligase (DltA) of *Bacillus cereus* (Du et al. 2008) and *B. subtilis* (Yonus et al. 2008). DhbE is an unusual adenylating enzyme because after activating the DBH, it transfers the activated substrate to a separate PCP domain that is encoded by the *dhbB* and there is no link between DhbE and DhbB proteins (May et al. 2002). DltA also resembles an A domain of a NRPS by catalyzing the ATP dependent activation of D-alanine and subsequent transfer of the activated substrate to the D-alanyl carrier protein DltC (Du et al. 2008). Although the sequences of the members of the adenylate forming enzyme superfamily differ substantially, all share a similar folding pattern consisting of two structural subunits. The large N-terminal subunit is about 420 amino acids in length and linked to the 110 amino acid long C-terminal subunit by a hinge-like region. The active site is located at the junction of the two subunits. Comparison of the amino acid sequence in the active site of DbhE and PheA as revealed by the crystal structure, along with the sequence alignment of other A domains showed that 10 amino acids (A1-A10) in the active

site were crucial for substrate recognition and activation (Conti et al. 1997). From the crystal structure of PheA, it was observed that these residues are located within a radius of ~5.5 Å from the phenylalanine substrate bound in the active site. Therefore a predictive method was described based on these highly conserved 10 amino acid sequences in the active site of the A domain and they were designated as forming a 'specificity-conferring code' for peptides produced via the NRPS mechanism (Challis et al. 2000; Stachelhaus et al. 1999). The specificity of uncharacterized A domains can be determined by aligning their sequences with those of other A domains of known specificity (Challis et al. 2000), or more simply by comparing only the 10 specificity-conferring amino acid sequences of the active site (Stachelhaus et al. 1999).

Very recently a new transductive support vector machine (TSVMs)-based approach has been described for functional sub-typing of the substrate specificities of the A domain based on the physico-chemical fingerprint of the amino acid residues at the active site. Instead of a 10 amino acid (A10)-based codon, a signature sequence of 34 amino acids, excluding the A10 core motif of the A domains, of known specificity has been used in this system to generate a database. Analysis of these data was used to generate an algorithm based on the physico-chemical properties (the number of hydrogen bond donors, polarity, volume, secondary structure preferences for beta-turns, beta-sheets and alpha-helices, hydrophobicity and the iso-electric point) of these amino acids. The substrate specificity of a new A domain is predicted by analyzing and comparing the sequence of the A domain with this database. Although the prediction

accuracy is very high, this method does not reduce the value of the 'specificity conferring code' method, as a direct way to predict specificity. Rather, a combination of both methods will give a powerful tool to predict accurately the substrate specificity of unknown A domains. Thus, this will make it possible to deduce a putative structure of a non-ribosomally synthesized peptide from the nucleotide sequence of its synthetase genes (Rausch et al. 2005).

1.6.2 Peptidyl carrier protein or thiolation domain

The next step in non-ribosomal peptide biosynthesis is to transfer the activated amino acid to the peptidyl carrier protein (PCP) domain. The PCP domain is about 80-100 amino acids in length and is the second domain of a minimum elongation module. This domain is called PCP because of its functional and structural similarity with the acyl carrier protein (ACP) of fatty acid and polyketide synthases (Marahiel 1997). This is also known as the thiolation (T) domain because it forms a carboxy-thioester intermediate with the activated amino acid (Stachelhaus et al. 1996a). The activated amino acid is transferred to the thiol moiety of a 20 Å long cofactor 4'-phospho-pantetheine (ppan) covalently tethered to a PCP domain. The cofactor ppan binds to the side chain of an invariant Ser (S) residue of PCP at position 45 located in a very conserved sequence LGG(HD)S(LI) known as motif 6 (Stachelhaus et al. 1996a). The binding of the cofactor ppan to a PCP domain converts it from its apo form to holo form, and the transfer of the cofactor is catalyzed by ppant ransferase (Schwarzer et al. 2003). The cysteamine thiol group of the cofactor ppan on the

PCP domain then becomes the binding site for the activated amino acid (Lambalot et al. 1996). This mode of substrate flow along the synthetase has been described as the multiple-carrier thiotemplate mechanism (Stachelhaus et al. 1996b; Stein et al. 1996). Dissection studies of the different domains of a module showed that the PCP domain is a segment of a synthetase enzyme and has no autonomous activity. It can only act in cooperation with other domains. The first A domain of gramicidin S synthetase was shown to activate the amino acid (Phe) but without the following PCP domain, it could not catalyze the formation of the thioester intermediate. This provided strong support for the modular structure of peptide synthetases and the multiple-carrier thiotemplate mechanism (Stein et al. 1996) .

1.6.3 Condensation domain

Condensation (C) domains are about 450 amino acids in length and are always present on the N-terminal ends of minimum elongation modules. Condensation domains catalyze the formation of peptide bonds in non-ribosomal peptide biosynthesis, resulting in the elongation of the peptide chain from N to C terminal. The number of peptide bonds in a fully assembled peptide product is usually equal to the number of C domains in its synthetase (Stachelhaus et al. 1998). Each C domain virtually has two sites, one acceptor site and one donor site and these two sites then interact with the ppan tethered activated amino acids on its two adjacent PCP domains. According to the multiple carrier thiotemplate model (Stein et al. 1996), a peptide bond is formed by the nucleophilic action on the thioester bond, through which the activated amino acid is tethered to the

acceptor PCP domain. Therefore, the growing peptide chain is tethered to the ppan on the donor PCP domain by the free amino group of the activated amino acid already tethered to it (Mootz and Marahiel 1997). The acceptor site has higher affinity for the activated amino acid on the PCP domain than the donor site, and that is how the direction of peptide elongation is maintained (Mootz et al. 2000; Mootz et al. 2002). Biochemical studies of many C domains demonstrated that the acceptor site exhibits selectivity towards both the stereochemistry (L- or D- isomer) and the nature of the side chain of the amino acid to be incorporated. Therefore, in addition to the A domain, the C domain acts as a second line of substrate selection fidelity (Ehmann et al. 2000).

The crystal structure of VibH is the only atomic model for a C domain available to date. VibH is an freestanding enzyme that catalyses the formation of a peptide bond between the α amino group of norspermidine and 2,3-dihydroxybenzoate (DBH) tethered to ppan on a separate protein, VibB, in the biosynthesis of the siderophore vibriobactin in *Vibrio cholerae* (Keating et al. 2002). However, from the degree of sequence homology between VibH and conventional C domains, it was predicted that all C domains will have similar structure (Keating et al. 2002).

VibH has an N- terminal and a C terminal domain, which are both very similar and have a characteristic αβα sandwich structure. Therefore VibH is described as a 'pseudo-dimeric' form of a monomeric protein (Keating et al. 2002).

Although a very weak sequence homology has been observed among the CoA dependent acyltransferases and the C and E domains of NRPSs, they share a conserved H HxxxDG motif (Keating et al. 2002). The His195 residue in the conserved motif of chloramphenicol acetyl transferase (CAT), homologous to the His126 of VibH, was found to be a general base to attack the thioester bond by activating the primary hydroxyl group of chloramphenicol as a nucleophile (Lewendon et al. 1994). However, studies of the conserved HHxxxDG motifs of many CoA-dependent acyltransferases revealed that not all of the structurally related acyltransferase enzymes use histidine residues for the same role in acyltransfercatalysis (Keating et al. 2002). A more defined role for these conserved histidine residues will require further study.

1.6.4 Thioesterase Domain

The last step in non-ribosomal peptide biosynthesis is the release of the fully assembled peptide product from the biosynthetic machinery, and this reaction is catalyzed by the thioesterase domain (TE) or termination domain. The cleavage of the fully assembled peptide product makes the NRPS ready for the next round of peptide biosynthesis (Schwarzer et al. 2003). Typically two types of TE domains can be seen in NRPS systems (Challis and Naismith 2004). Type I TE (TEI) domains are usually involved in this cleavage reaction to release newly synthesized peptides. TE domains are about 280 amino acids in length and are located at the C terminus of the last module of NRPSs systems (Kohli and Walsh 2003). The cleavage of the mature peptide product from the NRPS machinery is a

two step process. First, the full length peptide chain tethered to the ppan on the PCP domain of the last module is transferred from the ppan cofactor to the hydroxyl group of a highly conserved serine (S) residue located in the GxSxG core motif of the TE domain. This results in the formation of a peptidyl-O-TE intermediate. The next step is the deacylation of this intermediate by cleaving the ester bond with the nucleophilic attack of a water molecule to produce a linear peptide product, or by an intermolecular nucleophilic reaction to produce a cyclic or branched chain peptide (Bruner et al. 2002).

The first insights into the functional and structural properties of TE domains came from the crystal structure of the TE domain (SrfTE) of the surfactin producing NRPS in *B. subtilis*. Based on structural features, SrfTE is classified as a member of the α/β hydrolase family. From its structural homology to serine esterases, it was hypothesized that the Ser80, His207 and Asp107 of the SrfTE form the active site catalytic triad, which is involved in the macrocyclization and cleavage of the peptide product (Bruner et al. 2002). The crystal structure of another TE domain (FenTE) of a NRPS, producing the antibiotic fengycin in *B. subtilis* has recently been determined and the structure suggested that FenTE is also a member of α/β hydrolase family (Samel et al. 2006).

The co-crystal structure of SrfTE attached to the seven amino acid linear chain of surfactin, and biochemical data suggest that only the last two residues of surfactin (Leu7 and Leu6) bind specifically to the active site of the SrfTE domain and this binding is crucial for efficient macrocyclization. Residue substitution

mutation studies revealed that the Glu in position 1 in the surfactin molecule is also crucial for macrocyclization, although it was not required for the cleavage of the peptide product. This suggested that the Glu1 residue is important for the intramolecular nucleophilic attack to form macrocycliclactone from the peptidyl-O-TE intermediate (Tseng et al. 2002).

Many of the NRPS gene clusters so far analysed also contain Type II TE domains, although their exact function is not clear. They are called Type II TE domains because of their sequence homology to the TEIIs of vertebrate fatty acid synthases whereas they have only 10% sequence homology with the TEIs of NRPS systems. They do however share the GxSxG core motif (Schneider and Marahiel 1998). TEIIs are not mandatory for non-ribosomal peptide biosynthesis but are important because their removal decreases product yield significantly although it does not completely abolish the biosynthesis of the product (Schneider and Marahiel 1998). Recent work revealed that they are involved in hydrolytic removal of acetyl groups that inactivate and block the NRPS machinery by binding to the ppan on PCP domains. Therefore, they play role in regenerating the NRPS machinery by catalyzing the 'deblocking after mispriming' steps (Schwarzer et al. 2002).

1.6.5 Modifying Domains or Tailoring Domains

Other than the regular domains, there are some additional domains that can be present in some NRPSs (Figure 1.5.1). These additional domains are responsible for much of the structural diversity of non-ribosomally synthesized

peptides by carrying out the modification of amino acid substrates during peptide biosynthesis. Moreover, these chemical modifications can impart increased structural rigidity and stability against proteolytic degradation to the natural non-ribosomal peptides.

The cyclization (Cy) domain is about 450 amino acids in length and substitutes for the C domain in many NRPSs (Lautru and Challis 2004). No structure of a Cy domain is available to date. Cy domains exhibit a very high sequence homology with C domains and catalyze both condensation and cyclization reactions (Challis and Naismith 2004). Cy domains have a DxxxxD core motif instead of a NHxxxDG motif (core motif of C domains), and the conserved Asp (D) residues are crucial for both the condensation and heterocyclization reactions (Keating et al. 2002). It has been shown that the catalytic reaction of Cy domains is a two step process: first it catalyzes the formation of the peptide bond and then the cyclization of the thiol side chain of cysteine or hydroxyl side chain of threonine or serine (Walsh et al. 2001). Cyclodehydration of the peptide bond catalyzed by Cy domains generates rigid five-membered heterocycles, oxazolines from threonine and serine, and thiazolines from cysteine (Schneider et al. 2003).

In some NRPSs, another domain called an oxidation (Ox) domain is found to be strictly associated with the Cy domain. It is about 250 amino acids in length and can be found in two different positions with respect to the boundaries of NRPS modules: either at the C-terminal part of A domain of the same module, or downstream of the PCP domain of the previous module. Ox domains are found to

be involved in the oxidation of thiazoline and oxazoline rings formed by the Cy domain into aromatic thiazole or oxazole rings respectively, using flavin-mononucleotide as a cofactor (Du et al. 2000).

One of the structural features of many non-ribosomally synthesized peptides is the presence of methylated amino acids. The addition of the methyl group to the amino acids is usually catalyzed by a methyltransferase (MT) enzyme. MT domains can be classified into two types based on the site of methylation: N-methylation is carried out by N-methyltransferases (N-MT) and C-methylation is by C-methyltransferases (C-MT). Usually a MT domain is found to be inserted in the module responsible for the activation of the amino acid to be methylated, and methylation can take place either before or after the activation of the amino acid (Walsh et al. 2001). The N-MT is usually 420 amino acid in length and present immediately after the C terminal end of the A domain. Therefore, a typical domain order for this module would be C–A(MT)–PCP (Grunewald and Marahiel 2006). The occurrence of N-MT domains is more frequent than C-MT domains in NRPSs or in hybrid PKS/NRPSs (Walsh et al. 2001).

One of the examples of diversity in the composition of non-ribosomal peptides is the presence of D-amino acids. D-amino acids can be incorporated directly into non-ribosomal peptides or they can arise from the conversion of an already incorporated L-amino acid to its D-form. This stereochemical conversion is catalyzed by a domain called an epimerization (E) domain (Stachelhaus and Walsh 2000). E domains are about 450 amino acids in length and are located at the C-terminal end of the PCP domain of a module that is eventually responsible

for the incorporation of a D-amino acid. E domains exhibit a weak sequence homology to C domains but share a conserved HHxxxDG motif (Keating et al. 2002). Mutation studies on the tyrocidine synthetase system revealed that the second histidine residue in the conserved HHxxxDG motif is crucial for the epimerisation reaction (Stachelhaus and Walsh 2000). From their structural similarity, it is proposed that E domains might have arisen from the C domains through evolutionary processes (Challis and Naismith 2004). E domains can catalyze the epimerization reaction either before or after the formation of peptide bond, and the timing depends on the position of the E domain in the NRPS system. E domains located in an initiation module usually catalyze epimerization before condensation of the amino acids, whereas, E domains in chain elongation modules catalyze epimerization after the formation of the peptide bond (Luo et al. 2002).

It is not clear how substrate selectivity is ensured by E domains. However the most supported mechanisms of epimerization reaction catalyzed by E domain is called a two-base mechanism, in which one enzymic base deprotonates the Cα, and it is then reprotonated by the cognate enzymic acid. It is assumed that the E and C domains form a temporary dimer to carry out the epimerization and subsequent condensation reaction by recognizing the right stereoisomer of the amino acid (Luo et al. 2002).

1.6.6 Post assembly tailoring

In addition to the modification done by the integrated tailoring domains of the NRPS assembly line, some modifications can take place in the NRPs by independent enzymes temporally associated with NRPS system. Post assembly modification is very important for many NRP antibiotics because this modification can make the antibiotic bioactive. For example ACV, the immediate precursor of penicillin, is an acyclic tripeptide with no antibiotic activity. The post assembly tailoring enzyme is openicillin N synthase (IPNS) converts the acyclic ACV to isopenicillin N with the characteristic bicyclic ring required for its bioactivity (Roach et al. 1997). Post assembly modifications are also essential for two other very active antibiotics, vancomycin and daptomycin, to become bioactive (Nolan and Walsh 2009).

1.7 Proofreading

The transmission and expression of genetic information has to be very pricise to maintain the heredity and that is why they are carried out using a template-dependent synthesis process. Fidelity in the substrate selection according to the template is also very important to retain the information correctly. The enzymes involved in replication, transcription and translation follow a common mechanism to ensure fidelity. DNA and RNA polymerases can efficiently differentiate between structurally similar substrates and select the correct substrate, nucleoside triphosphates (NTPs) and aminoacyl-tRNAs (aa-tRNAs), respectively. The principle of this process of substrate selection is the complementary base pairing of the substrate or the substrate carrier, with the

respective template. In ribosomal peptide biosynthesis, using the mRNA as a template, ribosomes select the correct aminoacyl-tRNA (aa-tRNA) from a pool of available substrates on the basis of the complementary base pairing between codon on the mRNA and the anticodon on tRNA (Rodnina and Wintermeyer 2001). In ribosomal peptide biosynthesis, 20 amino acids are incorporated via 61 possible amino-acyl tRNAs, whereas in non-ribosomal peptide synthesis there are more than 300 different building blocks that can be incorporated (Marahiel 1992). Therefore the question is how NRPSs ensure fidelity during the non-ribosomal peptides biosynthesis.

The 10 amino acid positions in the active site of an A domain, the 'specificity conferring code', serve as a "codon" for the selection of the correct substrate avoiding the wrong one. Therefore an A domain will activate only one substrate that perfectly fit in the active site by interacting with the 'specificity conferring code'. This discrimination in substrate selection by A domains provides the first line of specificity to maintain the fidelity of non-ribosomal peptide biosynthesis (Challis et al. 2000).

The 'multiple carrier thio-template' mechanism of NRP synthesis provides another means to ensure fidelity. When the substrate amino acid passes through the domains in a module, recognition of the substrate on an upstream domain by a downstream domain is also very important for the fidelity of the process. The incorporation of correct amino acids by non-ribosomal peptide synthetases occurs in two steps: activation of substrate as an aminoacyl-adenylate followed by a transfer to ppan on a PCP domain to form a thioester intermediate. Therefore,

during the amino acid incorporation, fidelity can be ensured by rejecting either the mis-formed aminoacyl adenylate or the mis-acylated thioester (Pavela-Vrancic et al. 1999).

The ppan transferases have a broad tolerance to CoA derivatives and can transfer only ppan alone as well as acyl-ppan to PCP domains. Transfer of acyl-ppan to PCP domains is called mispriming of PCP domains. The misprimed PCP domains are not recognized by the immediate downstream C domain, resulting in blocking of the NRPS machinery (Schwarzer et al. 2003). Type II thioesterase (TEII) has been reported to be involved in the regeneration of misprimed NRPS machinery by catalyzing hydrolysis of these acyl groups (Yeh et al. 2004). The C domain also has a role in the fidelity of the non-ribosomal peptide biosynthesis process. An incorrectly loaded substrate cannot be recognized by the respective C domain which increases the half life of the unprocessed substrate enzyme intermediate. The increased half life of the unprocessed substrate intermediate can be kinetically detected by a TEII which then hydrolytically removes the incorrect acyl group from the PCP domain. This process of proofreading is very similar to that of ribosomal peptide synthesis (Yeh et al. 2004).

1.8 Non-ribosomally synthesized lipopetides

Non-ribosomally synthesized peptides very often have an N-terminally attached fatty acid chain and hence are called lipopeptides. This fatty acid part gives them more structural diversity and amphipathic properties. Non-ribosomally synthesized lipopeptides have received considerable attention for their surfactant-

type antimicrobial properties. However in NRPSs, no domains are found to be involved in fatty acid biosynthesis. Therefore, it has been suggested that the fatty acid moieties of lipopeptides are synthesized by separate PKSs and are incorporated into peptide products by direct transfer to the first activated amino acid of the peptide chain. Although the enzymes catalyzing the transfer of the fatty acid moiety have not been identified so far, a condensation domain at the N terminal end of the first NRPS module was reported to be found in all lipopeptide synthetase enzyme complexes. Therefore, it is assumed that this condensation domain in the first module of the NRPS catalyzes the incorporation of the fatty acid (Schwarzer et al. 2003).

The lipopeptides produced by various species of bacteria, especially various strains of *Bacillus,* can be classified into two subgroups based on their types of fatty acids. The first group includes the lipopeptides, surfactin, lichenysin and fengycin which contain β-hydroxy fatty acids. The members of the second group are iturin, mycosubtilin and bacillomycin which contain a β-amino fatty acid. Gene clusters responsible for synthesis of several members of both groups have been completely sequenced and characterized (Duitman et al. 1999; Moyne et al. 2004). The first insights into the synthesis of the second group of lipopeptides came from the sequence of the mycosubtilin operon. The mycosubtilin operon displays a typical modular arrangement of peptide synthetases along with a functional domain homologous to fatty acid synthases and PKS. An amino transferase domain was also found at the N-terminal end of the first module of the NRPS (the second ORF of the operon), and this domain

was proposed to be responsible for the activation of a long chain fatty acid. However no genes have been identified to account for the synthesis and incorporation of the β-hydroxy fatty acid moieties in group one lipopeptides in either the peptide synthetase operon or the flanking regions (Konz et al. 1999; Steller et al. 1999). In the case of daptomycin, two genes *dptE* and *dptF* were assumed to be involved in N-terminal lipidation of the peptide part. The deduced product of *dptE* was shown to contain a conserved motif of adenylate forming enzymes and was assumed to catalyze the acylation of the fatty acid chain which is then transferred to the first amino acid of daptomycin by DptF. DptF, the deduced product of *dptF,* was found to be homologous to acyl carrier proteins (ACP). This mode of free fatty acid activation was also shown in *Mycobacterium tuberculosis.* The product of one of the *fadD* genes in *M. tuberculosis* was found to catalyze the formation of acyl-adenylates of long chain fatty acids (Trivedi et al. 2004). However the exact mechanisms of the transfer of the fatty acid to the peptide chain have not yet been made clear. Polymyxins have N-terminally attached heptanoic acid or octanoic acids and so they do not fit clearly into either of these two subgroups.

1.9 Hybrid PKS/ NRPS

Non-ribosomal peptide synthetases and polyketide synthases are involved in the synthesis of many secondary metabolites with important therapeutic and antimicrobial properties. However, recently it has been shown that these two multimodular enzyme systems can work together to synthesis hybrid natural

products. The hybrid systems of these two different modular enzyme complexes are of special interest for combinatorial biosynthesis, because the combination of PKS and NRPS activities can lead to the endless possibility of synthesis of natural products with even greate diversity in structure and activity.

Although PKSs and NRPSs share a similar modular arrangement, the biosynthetic principle of PKSs is even more similar to that of fatty acid biosynthesis. PKSs can be classified into three types but the hybrid NRPS/PKS systems contain mainly type I PKS components. Like NRPS, a typical elongation module of a type I PKS consists of three core domains: acyltransferase (AT), acyl carrier protein (ACP) and ketoacyl synthase (KS). Polyketide synthesis usually starts with the transfer of a two, three or four carbon building block such as acetyl-CoA, propionyl-CoA or butyryl-CoA to the ACP domain of the PKS system. The transfer of these already activated substrates is catalyzed by the AT domain. In contrast to the A domain of a NRPS, the AT domain of the PKS does not need to activate the substrate, only catalyze the transfer reaction. Following transfer, acyl-substrate is tethered to a ppan cofactor, covalently attached to the ACP domain by a thioester bond and this step is very similar to the step carried out by PCP of a NRPS. Then the KS domain catalyzes the condensation of the two acyl-substrates tethered on two adjacent ACP domains. The condensation reaction is catalyzed in two steps: first the incoming acyl-substrate binds to the KS domain by forming a thioester bond with a conserved cysteine residue and then the KS domain decarboxylates the acyl-substrate on the acceptor ACP domain and transfers the incoming residue onto it to elongate the chain. Like

NRPSs, PKSs have some tailoring enzymes like ketoreductases (KR), dehydratases (DH) and enoyl reductases, involved in the reduction of the β-keto acyl intermediates. All three are required for the complete reduction to yield a saturated fatty acid chain and lack of any of them results in only partial reduction. When the full length product is formed, a thioesterase domain (TE) catalyzes the cleavage of the final product from the biosynthetic machinery with or without macrocyclization (Keating and Walsh 1999; Staunton and Weissman 2001).

Integrated NRPS/PKS systems are classified into three different types. In the first type the unreleased product of a PKS is directly transferred to a NRPS multi-enzyme. In this case the gene encoding the PKS is not linked to the gene encoding the NRPS. Examples of this type include rapamycin, epothilon and mycobactin. An acyl transferase domain (AT) is assumed to be involved in transfer of the immediately synthesized polyketide to the N terminal condensation domain of the first module of the NRPS. The second type is characterized by integrated genes where both NRPS and PKS domains are found within a single protein. The examples of this type are mycosubtilin, the antibiotic TA, and yersiniabactin. In this case a putative amino transferase domain is found to be involved in the transfer of polyketide chain to the NRPS machinery (Tillett et al. 2000). In the third type of hybrid system, a fatty acid intermediate is synthesized by PKS and transformed into a CoA derivative which is then transferred to the N terminal condensation domain of the first module of the NRPS system. One of the examples of this system is surfactin, in which the β-hydroxy fatty acid is

transferred by SrfD, an acyltransferase enzyme from β-hydroxymyristoyl-CoA to the Glu module of NRPS (Steller et al. 2004)

1.10 Rational design

The in-depth analysis of the different NRPS systems and other processes provided a good understanding of how non-ribosomal peptides and polyketides are produced in nature by modular assembly line enzyme machinery. This profound knowledge of modular organization of NRPSs and PKSs has led to a newer area research to make novel compounds by manipulating the modular assembly line of the biosynthetic enzymes or by applying a combination of chemical and biological approaches. The NRPSs and PKSs employ very similar systems to produce peptides and polyketides respectively. Domains with different catalytic fuctions of these enzyme complexes are responsible for the activation of specific precursor units, their subsequent modification, maturation and release of the final product. This type of catalytically active functional modular arrangement implies that any synthetase that contains the appropriate number of these modules in the correct order could synthesize any defined product (Kopp and Marahiel 2007). Sequence analysis of many bacterial and fungal non-ribosomal peptide synthetases has shown that there are certain levels of sequence homology present in every different functional domain. As each adenylation domain is responsible for activating a particular amino acid, changing an A domain can lead to the incorporation of a different amino acid in the final peptide product (Schneider et al. 1998). So by knowing the roles of the different component amino acids in the

peptide, it will be possible to design a new peptide with more bioactivity but less toxicity by directly changing its biosynthesis enzyme template. Biochemical and genetic studies of several biosynthesis systems have revealed the minimal size of each functional domain and the interdomain and intermodule linking regions. This knowledge gives researchers the opportunity to locate and amplify a particular domain encoding region responsible for substrate recognition and activation, from various peptide synthetase and polyketide synthase encoding genes in bacterial and fungi (Stachelhaus et al. 1995).

The basic principle for combinatorial biosynthesis of a natural product is to combine components of the synthesizing machinery of different metabolites rationally to produce a new molecule of interest. Therefore, in combinatorial biosynthesis, the choice of the fusion site between modules or domains is very important. Biochemical and structural studies have led to precise identification of the domain borders as well as the linker regions in between the modules (Mootz et al. 2000).

The linker regions between the domains within a module are different than the linker regions between the modules in a NRPS. Highly conserved residues were found in the linker regions and these linker regions offer an advantage by providing a location to cut and link domains or modules from different NRPS to form a totally new NRPS. A minimal elongation module of a NRPS consists of three core domains present in the order C-A-PCP, suggesting three different fusion points for recombination of domains: i) intramodular fusion between the C and A domain, ii) intramodular fusion between the A and PCP domain and iii)

intermodular fusion between the PCP and C domains (CAPCP-CAPCP) (Mootz et al. 2000). Although it is possible to swap individual domains, better productivity of the engineered NRPS is found when a whole module is changed. The possible reason may be that the cognate PCP or C domain is most compatible with its A domain.

The second approach for synthesizing non-natural peptide using NRPS machinery is called a chemoenzymatic process. The basic principle of chemoenzymatic biosynthesis of a product is to modify a chemically synthesized product by using certain biologically active enzymes *in vitro*. It was demonstrated that the excised TE domain from a modular biosynthetic system was able to catalyze the specific macrocyclization reaction of chemically synthesized acyl substrates *in vitro*. This offers great potential to diversify the natural products synthesized by various NRPSs and PKSs (Kohli et al. 2002). In chemoenzymatic synthesis of daptomycin, a linear precursor peptide chain synthesized entirely by solid state peptide synthesis, a chemical synthesis process was converted to daptomycin by stereo- and regioselective macrocyclization reaction catalyzed by an isolated TE domain. To imitate the natural conditions for TE domain to catalyze the reactions, the artificial peptide chain was activated by tethering the C-terminal end to N-acetylcysteamine (SNAC), which mimics a ppan molecule (Grunewald et al. 2004).

1.11 Transporter Genes

Transporter-like genes have been reported for almost all NRPS systems in bacteria and fungi (Guillemette et al. 2004; Menges et al. 2007; Pearson et al. 2004). Transporter genes usually are present in the flanking region of the NRPS gene cluster and are transcriptionally oriented in both directions, either in the same or opposite, compared to the NRPS genes. Most of these transporter genes fall into the ATP-binding cassette (ABC) superfamily. ABC transporters usually transport a variety of molecules ranging from very small (ions, amino acids) to large molecules (peptides, polysaccharides) in or out of the cells. Most of the ABC transporters transport molecules unidirectionally through the cell membrane and are called primary transporters. ABC transporters are classified into two families based on their polarity of transport: export or import. Those involved in the export of various peptides or polysaccharides and hydrophobic antibiotics fall into the ABC-A transporter family and those involved in the import of various solutes fall into the ABC-B transporter family. The transporter proteins of the ABC-B family are components of periplasmic binding protein-dependent (BPD) uptake systems. Proteins of each family can be further divided into two sub-families based on their secondary structure: ABC-A1, ABC-A2 and ABC-B1, ABC-B2. A characteristic structural feature of the ABC transporter system is that it contains a membrane spanning (MS) domain and an ABC module. The MS domains and ABC modules can be present in the same protein or in separate proteins and it is believed that they function as homodimers (Saurin et al. 1999).

The ABC modules are considered to be the most conserved regions, which consist of the $Walker_A$, $Walker_B$ motifs and the linker peptide (Walker et al. 1982).

1.12 Research proposal

Although polymyxins were discovered more than half a century ago, molecular characterization of the genes encoding the NRPS for polymyxin synthesis is very limited. Only one gene sequence of a polymyxin gene cluster is so far available in GenBank, and it is an out come of the whole genome sequencing of a natural isolate, *P. polymyxa* E681. Very recently a whole genome sequence of another strain of the same species, *P. polymyxa* SC2, was deposited in the GenBank, but preliminary examination of the sequence for the genes encoding the NRPS for polymyxin production indicates that they may contain errors and need further editing.

It is well established that polymyxins are produced via a non-ribosomal peptide synthetase mechanism. Structural insights into the mechanism of non-ribosomal peptide synthesis have been obtained in recent years and this knowledge can be exploited for the biosynthesis of new or modified natural products. Therefore, it should be possible to make natural non-toxic analogs of polymyxins by introducing changes into the NRPS by modification of the genes. Combinatorial biosynthesis, semi-enzymatic biosynthesis, or swapping domains or modules of NRPSs have been reported for many antibiotics as ways to change their composition. Therefore, decoding the genetic information for polymyxin

biosynthesis is very important as a first step to take this antibiotic to a new level for combating multidrug resistant Gram negative bacteria.

Therefore, the main goal of this research was to identify the genes encoding the NRPS that catalyzes the production of polymyxins in *P. polymyxa* PKB1. *P. polymyxa* PKB1 is an environmental isolate, identified based on its biochemical and growth characteristics and its 16S r RNA gene sequence similarity. *P. polymyxa* PKB1 has been reported to produce polymyxins and the antifungal antibiotic fusaricidin (Beatty and Jensen 2002; Li et al. 2007).

Table 1.3.1: Different types of polymyxins and their chemical compositions

Types of polymyxins		Amino acids position										Fatty acids	Mass
		1	2	3	4	5	6	7	8	9	10		
Polymyxin A	Polymyxin A	L - DAB	L - Thr	D - DAB	L - DAB	L - DAB	D - Leu	L - Thr	L - DAB	L - DAB	L - Thr	6-MOA	1156
Polymyxin B	Polymyxin B_1	L - DAB	L - Thr	L - DAB	L - DAB	L - DAB	D - Phe	L - Leu	L - DAB	L - DAB	L - Thr	6-MOA	1202
	Polymyxin B_2	L - DAB	L - Thr	L - DAB	L - DAB	L - DAB	D - Phe	L - Leu	L - DAB	L - DAB	L - Thr	6-MHA	1188
	Polymyxin B_3	L - DAB	L - Thr	L - DAB	L - DAB	L - DAB	D - Phe	L - Leu	L - DAB	L - DAB	L - Thr	OA	1188
	Polymyxin B_4	L - DAB	L - Thr	L - DAB	L - DAB	L - DAB	D - Phe	L - Leu	L - DAB	L - DAB	L - Thr	HA	1174
	Polymyxin B_5	L - DAB	L - Thr	L - DAB	L - DAB	L - DAB	D - Phe	L - Leu	L - DAB	L - DAB	L - Thr	NA	1202
	Polymyxin B_6	L - DAB	L - Thr	L - DAB	L - DAB	L - DAB	D - Phe	L - Leu	L - DAB	L - DAB	L - Thr	3-OH-6-MOA	1218

Polymyxin D	Polymyxin D_1	L - DAB	L - Thr	D - Ser	L- DAB	L- DAB	D- Leu	L - Thr	L- DAB	L- DAB	L - Thr	6-MOA	1143
	Polymyxin D_2	L - DAB	L - Thr	D - Ser	L- DAB	L- DAB	D- Leu	L - Thr	L- DAB	L- DAB	L - Thr	OA	1129
Polymyxin E	Polymyxin E_1	L - DAB	L - Thr	L - DAB	L - DAB	L - DAB	D - Leu	L- Leu	L - DAB	L - DAB	L - Thr	6-MOA	1168
	Polymyxin E_2	L - DAB	L - Thr	L - DAB	L - DAB	L - DAB	D - Leu	L- Leu	L - DAB	L - DAB	L - Thr	6-MHA	1154
	Polymyxin E_3	L - DAB	L - Thr	L - DAB	L - DAB	L - DAB	D - Leu	L- Leu	L - DAB	L - DAB	L - Thr	OA	1154
	Polymyxin E_4	L - DAB	L - Thr	L - DAB	L - DAB	L - DAB	D - Leu	L- Leu	L - DAB	L - DAB	L - Thr	HA	1140
	Polymyxin E_7	L - DAB	L - Thr	L - DAB	L - DAB	L - DAB	D - Leu	L- Leu	L - DAB	L - DAB	L - Thr	7-MOA	1168
Polymyxin M		L - DAB	L - Thr	L - DAB	L - DAB	L - DAB	D - Leu	L - Thr	L - DAB	L - DAB	L - Thr	6-MOA	1156

(The most common areas of amino acid variations are shown in grey shaded columns. DAB= 2,4-diaminobutyric acid, 6-MOA= 6 methyl octanoic acid, 6-MHA= 6 methyl heptanoic acid, OA= octanoic acid and HA= heptanoic acid, NA= Nonanoic acid (Govaerts et al. 2002a; Govaerts et al. 2002b; Martin et al. 2003; Orwa et al. 2001; Parker and Rathnum 1975; Vogler and Studer 1966).

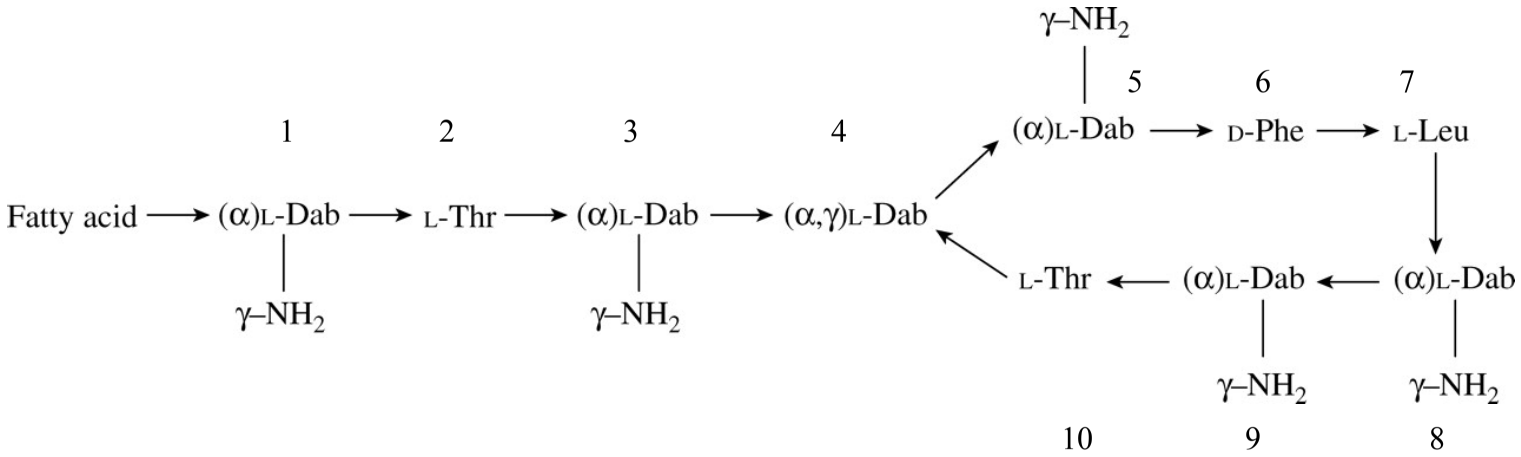


Figure 1.2.1: Structure of polymyxin B. The numbers indicate the position of the amino acids in the molecule.

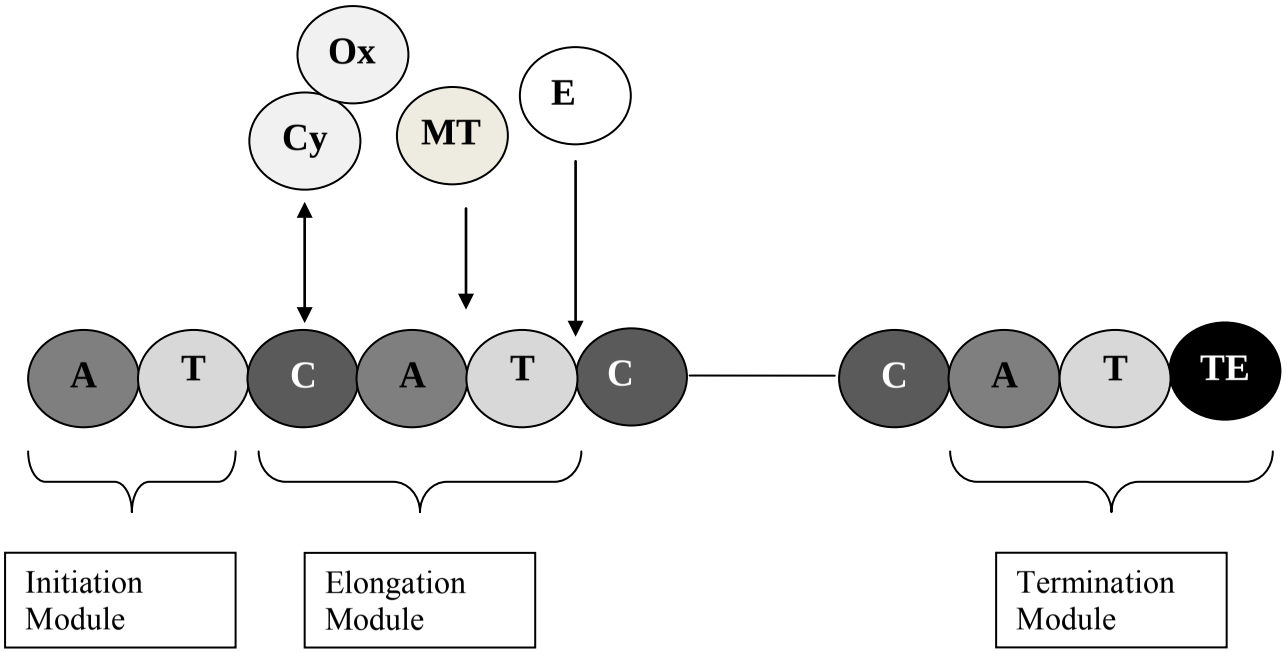


Figure 1.5.1: Schematic representation of module and domain organization of NRPSs. The organization of three modules is shown here: initiation, elongation and the termination module. A minimum elongation module contains one A domain, one T domain and one C domain.The relative position of tailoring domains (Cy, Ox, MT, and E) in modules are also shown (Adapted from Lautru and Challis 2004).

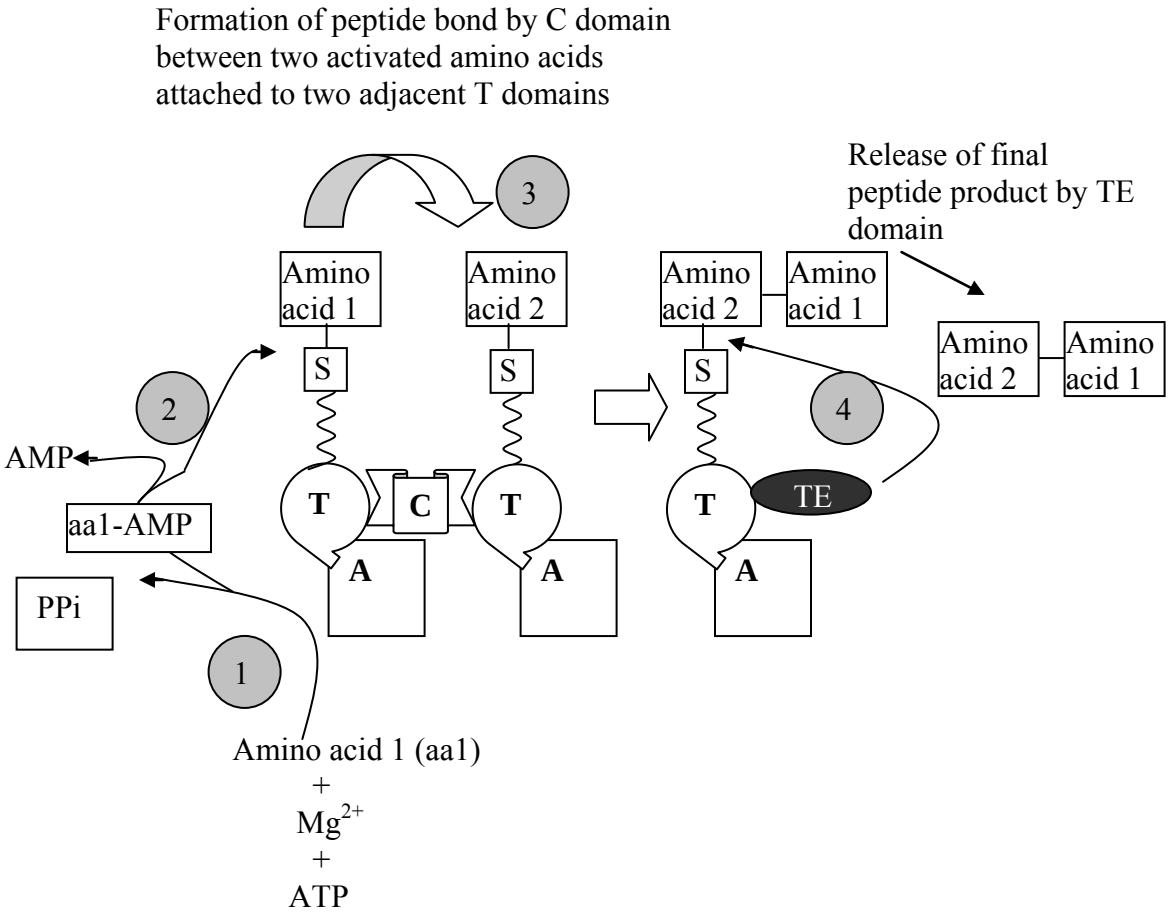


Figure 1.6. 1.1: The overall reaction mechanisms for non-ribosomal peptide biosynthesis. The condensation, adenylation, thiolation and thioesterase NRPS domains are labelled C, A, T and TE respectively, and the panthetheinyl arms of thiolation domains are indicated by wavy lines. The numbers in the grey circles indicate the reaction steps catalyzed by individual domains (Adapted from Marahiel and Essen 2009).

2. Materials and Methods

2.1 Bacterial and fungal strains, plasmids, reagents and media

All of the bacterial and fungal strains used in this study are described in Table 2.1.1. *Paenibacillus polymyxa* PKB1 and the antifungal indicator organism, *Leptosphaeria maculans,* were provided by the Alberta Research Council (Vegreville, Alberta). All of the genomic libraries of *P. polymyxa* PKB1 in cosmid and fosmid vectors and the plasmid constructs containing the $Apra^{R}Cm^{R}$*oriT* antibiotic resistance cassette used for PCR-targeted gene disruption were obtained from the previous study of fusaricidin production in *P. polymyxa* PKB1 (Li, 2007; Li et al. 2007). All other bacterial strains and plasmids used in this study were commercially available (Table 2.1.2).

The chemicals used in this study were reagent grade and the growth media were microbiological grade, purchased from Sigma and Difco Laboratories, unless otherwise mentioned. The restriction enzymes and other DNA manipulating and cloning products used were manufactured by Roche, New England Biolabs and Fermentas Life Sciences. All of the oligonucleotide primers used were prepared by Integrated DNA Technologies, Inc. (IDT) and supplied through Molecular Biology Services unit (MBSU, Department of Biological Sciences, University of Alberta).

Hybond-N nylon membranes used for colony lifts and Southern analysis were manufactured by Amersham Bioscience (Buckinghamshire, UK) and the 0.45µ HAWP disc filters used for mating of bacteria and filtration of solvents for HPLC analysis were manufactured by Millipore.

2.2 Growth conditions for bacteria

E.coli strains were routinely grown in LB broth (1.0% tryptone, 0.5% yeast extract, 0.5% NaCl) or super optimal broth (SOB; 2.0% Tryptone, 0.5% yeast extract, 0.05% NaCl, 2.5 mM KCl, pH 7.0) medium unless otherwise mentioned, at 37^{0}C for 16 h in a shaker at 250 rpm (Model G-25, New Brunswick Scientific Co., Edison, NJ) or tube roller (Bellco, Biotechnology). Antibiotics were added to the medium when appropriate, alone or in combination, at the following concentrations: ampicillin 100 µg/ml; apramycin 50 µg/ml; chloramphenicol 25 µg/ml; kanamycin 50 µg/ml; streptomycin 50 µg/ml and spectinomycin 100 µg/ml to select certain strains of bacteria or maintain antibiotic resistance genes. LBA (LB agar) medium was prepared by adding 1.5% agar to the LB.

E. cloni Replicator^{TM} strains containing the fosmid clones were grown in TB medium (Bacto tryptone 11.8 g/l, yeast extract 23.6 g/l, dipotassium hydrogen phosphate 9.4 g/l, potassium dihydrogen phosphate 2.2 g/l, glycerol 2 g/l) containing 12.5 µg/ml chloramphenicol and 1x Replicator induction (L-arabinose solution, Lucigen, Middleton, WI) solution for 16 h at 37^{0}C with shaking at 225 rpm. L-arabinose was used to induce the production of TrfA replicator protein of *E. cloni* to activate the medium-copy origin replicator (*oriV*) of the fosmid vector to increase its number to 20-50 copies per cell.

P. polymyxa PKB1 strains were routinely grown in GB medium (1.0% glucose, 1.0% peptone, 0.2% beef extract, 0.1% yeast extract and 0.5% NaCl) or Brain Heart Infusion (BHI) at 37^{0}C in a tube roller for 16 h. In some instances *P.*

polymyxa was grown in the same medium at 28 ^{0} C in a rotary shaker. When required 1.5 % agar was added to GB medium or BHI and used with or without antibiotics as required.

L. maculans was grown on potato dextrose agar (PDA) at 28^{0}C for 3 d for use as an indicator strains for antifungal bioassay as well as for making spore stocks.

2.3 Storage of bacterial strains

All of the bacterial strains used in this study were stored in 20% glycerol at -80^{0}C. *E. coli* strains were grown in 2 ml LB or SOB media at 37 ^{0}C unless otherwise mentioned for 16 h in a tube roller. After incubation, 1.5 ml of culture was harvested at 5000 rpm for 1 min in bench-top centrifuge (Eppendorf 5424), 900 µl of the culture supernatant was removed and 400 µl of 50% glycerol was added to make a final glycerol concentration 20%, and stored at -80^{0}C. *E. cloni* fosmid-bearing clones were grown in TB medium for 16 h at 37 ^{0}C and a 20% glycerol stock culture was prepared following the same procedure described above and stored at -80^{0}C.

P. polymyxa PKB1 was grown in GB medium at 37 ^{0}C for 16 h in a tube roller and stored in 20% glycerol at -80^{0}C.

L. maculans grown on PDA at 28^{0}C for 3 d was scraped with a spatula, re-suspended in 20% glycerol in 1.5 ml microcentrifuge tube and stored at -80^{0}C.

2.4 Making electro competent cells

2.4.1 *E. coli* DH5α

The *E. coli* DH5α cells were grown in 2 ml SOB medium for 16 h at 37^{0}C and 2 ml of this was used as a pre-culture to inoculate 200 ml of SOB medium in a 500 ml flask. The culture was grown at 37^{0}C at 250 rpm until the OD_{600} reached 0.5-0.6. The culture was then chilled on ice and harvested by centrifuging at 3000 × g for 5 min at 4 ^{0}C. The cells were washed first with 200 ml of 4 ^{0}C 10% glycerol and then with 100 ml followed by 50 ml of 4 ^{0}C 10% glycerol. The washed cells were re-suspended in 3-4 ml of ice-cold 10% glycerol and dispensed in 40 µl aliquots, flash frozen in a dry ice-ethanol bath and stored at -80^{0}C for future use.

2.4.2 *E. coli* BW25113/pIJ790

Electrocompetent *E. coli* BW25113/pIJ790 cells used in the Polymerase Chain Reaction (PCR) targeted mutagenesis were grown in 2 ml of SOB medium containing chloramphenicol (25 µg/ml), at 30^{0}C for 16 h and 0.1 ml of this was used as a pre-culture to inoculate 50 ml of SOB medium containing chloramphenicol (25 µg/ml), and 20 mM $MgSO_4$. The culture was incubated at 30^{0}C, to maintain the temperature sensitive pIJ790 plasmid, for about 3 to 4 h until the OD_{600} reached about 0.6. The SOB medium was also supplemented with 10mM L-arabinose when the λ-RED functions encoded on pIJ790 were to be expressed. The cell culture was then chilled to 4^{0}C and harvested by centrifuging at 3000 × g for 5 min at 4^{0}C. The cells were washed with 50 ml of ice-cold 10%

glycerol followed by two washes with 25 ml of ice-cold 10% glycerol. The washed cells were re-suspended in 150 µl of ice-cold 10% glycerol. Electrocompetent cells of *E. coli* BW25113/pIJ790 were prepared fresh every time before use for transformation.

2.4.3 *E. coli* ET12567/pUZ8002

Electrocompetent *E. coli* ET12567/pUZ8002 cells used for intergeneric conjugation were grown in 5 ml of LB medium containing kanamycin (25 µg/ml) and chloramphenicol (25 µg/ml) to maintain selection for pUZ8002 and the *dam* mutation at 37^{0}C for 16 h and 2 ml of this was used as a pre-culture to inoculate 200 ml of LB medium containing the same antibiotics. The culture was grown at 37^{0}C for about 3 to 4 h until the OD_{600} reached about 0.6. The cell culture was then chilled to 4^{0}C and harvested by centrifuging at 3000 × g for 5 min at 4^{0}C. The cells were washed first with 200 ml of 4^{0}C 10% glycerol and then with 100 ml followed by 50 ml of 4^{0}C 10% glycerol. The washed cells were re-suspended in 3-4 ml of 4^{0}C 10% glycerol. The washed cell suspension was then dispensed in 40 µl aliquots, flash frozen in a dry ice-ethanol bath and stored at -80^{0}C for future use.

2.5 Transformation of competent cells

In this study both electrocompetent and chemically competent *E. coli* strains were used. For the transformation of frozen competent *E. coli* stains, the cells were slowly thawed on ice before use. An appropriate amount of linear or

circular DNA was added to a microcentrifuge tube containing 40 µl of competent cells, mixed well gently and held at 4 ^{0}C for 2-3 min. In the case of electrocompetent cells the entire contents of the microcentrifuge tube was transferred carefully with a micropipette into the bottom of a 2 mm gap size, disposable pre-sterilized electroporation cuvette (Molecular Bioproduct Inc, Thermo Fisher Scientific). The cuvette was placed in a BIO-RAD Gene Pulser®II and an electric pulse was applied to the cells with the instrument setting: 25 µF, 200Ω and 2.5 kV. In the case of chemically competent cells, the cells mixed with the appropriate amount of DNA were held at 4^{0}C for 30 min and then a heat shock was applied for 30 sec at 42^{0}C. Immediate after the electric pulse or heat shock, 1 ml of LB or SOB medium was added to the cell-DNA mixture and the suspension was transferred to a pre-sterilized microcentrifuge tube and allowed to recover by incubating at 37 ^{0}C for 1 h. An appropriate volume of the cell suspension was spread on a solid agar medium containing appropriate antibiotic(s) and was incubated at 37^{0}C unless otherwise recommended for 16-24 h to select for the desired transformants. In those cases with blue-white selection, 80 µl of X-gal (5-bromo-4-Chloro-3-β-D-galactopyranoside) was spread along with or without 40µl of 100nM IPTG (isopropyl-β-thiogalactopyranoside) on each plate before applying the cell suspension.

2.6 Isolation of DNA from bacterial culture

2.6.1 Preparation of plasmids, cosmids or fosmids

Appropriate bacterial strains with desired plasmids, cosmids or fosmids were grown under appropriate conditions. Usually 1 ml of culture was harvested by centrifugation in an Eppendorf microcentrifuge at 10,000 rpm for 1 minute. The supernatant was removed and cells were subjected to alkaline lysis (Birnboim method) (Birnboim and Doly 1979). The crude plasmid, cosmid or fosmid preparations were treated with RNase A, extracted with buffered phenol-chloroform and washed twice with chloroform-iso amyl alcohol mixture. Two volume of 95% ethanol was added to the final plasmid, cosmid or fosmid preparation to precipitate the DNA. The precipitated DNA was collected by centrifuging at 15,000 rpm in an Eppendorf microcentrifuge for 10 min, washed with 70% ethanol and re-dissolved in sterilized TE buffer or sterile double distilled water.

2.6.2 Isolation of genomic DNA from *P. polymyxa* PKB1

A modified version of a protocol originally used for the isolation of genomic DNA from *Streptomyces* (Hoopwood et al. 1985) was used for preparing the genomic DNA from *P.polymyxa* PKB1. The culture was grown in 50 ml BHI at 28 ^{0}C for 16 h at 250 rpm in a rotary shaker and the cells were harvested by centrifuging at 12,000 × g for 10 min at 4^{0}C. The culture supernatant was removed and the cell pellet was washed twice with an equal volume of 10.3% sucrose solution and re-suspended in 4 ml of freshly prepared lysozyme solution (25 mM Tris-HCl, 25 mM EDTA, 0.3 M sucrose, pH 8.0, lysozyme 4 mg/ml and RNase A 100 µg/ml). The sample was incubated at 37^{0}C and after 45 min SDS

and proteinase K was added to a final concentration of 2.0% and 0.2 mg/ml, respectively, and incubated at the same temperature for another 15 min. The cell lysate was then extracted twice with equal volume of buffered phenol-chlorofom-isoamyl alcohol (25:24:01 v/v) solution and subsequently washed twice with equal volume of chloroform-isoamyl alcohol (24:1 v/v). The aqueous portion containing the genomic DNA was transferred into a new tube and two volume of ice cold 95% ethanol and 0.1 volume of 3M sodium acetate were added to precipitate the DNA. The tube was gently inverted until the DNA was seen as a fibrous material floating in the liquid, and then kept in at 4^0C for 30 min. The genomic DNA was collected by centrifuging the tube at 12,000 × g for 10 min. The DNA pellet was washed with 70% ethanol, air dried and re-dissolved in 0.5 ml of 10 mM Tris-HCl buffer (pH 8.5) and stored at 4^0C.

For screening of different mutant *P. polymyxa* strains, Sigma® GenElute™ Bacterial Genomic DNA preparation kit was used according to the manufacturer's instructions using 1.5 ml of BHI culture grown at 37^0C for 16 h.

2.7 DNA methods

2.7.1 Polymerase chain reaction

Polymerase Chain Reactions (PCR) were routinely carried out in 200 µl thin walled PCR tubes using a T-Gradient thermocycler (Biometra, Goettingen, Germany). A typical reaction mixture contained 50 ng of plasmid or cosmid DNA or 150 ng of genomic DNA as template, 10 pmol of each primer, 0.25 µl of Taq DNA polymerase (Produced by Fermentation Service Unit, Department of

Biological Sciences, University of Alberta) in 1x Taq buffer (50 mM Tris-HCl, pH 9.0, 1.5 mM $MgCl_2$, 0.4 mM β-mercaptoethanol, 0.1 mg/ml bovine serum albumin, 10 mM ammonium sulphate, 0.2 mM of each dNTP) in a final volume of 12.5 µl. Primers used in this study are shown in Table 2.7.1.1.

The amplification of antibiotic resistance cassettes for PCR-targeted gene disruption was scaled up to give a total volume of 250 µl distributed in five PCR tubes.

2.7.2 Gel electrophoresis

Agarose gel electrophoresis was routinely used to fractionate DNA fragments of different sizes, to check the status of restriction enzyme digestion, purity of DNA preparations and to identify PCR products. Although 0.8% agarose gels in 1xTAE buffer (40 mM Tris-Acetate and 1 mm EDTA, pH 8.0) were routinely used, depending on the sizes of expected DNA fragments, 0.5%, 1.0% and 1.5% agarose gels were also used in some instances. The DNA samples loaded in a gel were fractionated by electrophoresis at 50-100V through the gel submerged in 1X TAE buffer.

After electrophoresis, gels were stained with ethidium bromide (2µg/ml) for 5-10 min, and observed under UV light. The sizes of DNA bands were estimated by comparing with molecular weight markers. Lambda phage DNA samples digested with *Bst*EII or *Pst*I were routinely used as molecular weight markers unless otherwise mentioned. Commercial molecular weight markers manufactured by Fermentas Inc. were used in some special instances.

2.7.3 Gel purification of DNA

Restriction endonuclease digested DNA fragments or PCR product(s) fractionated by electrophoresis on standard 0.8% or 1.0 % agarose gels were routinely extracted by using QIAquick Gel Extraction Kits (Qiagen, Mississauga, ON). The desired bands of DNA from ethidium bromide stained agarose gels, visualized under UV light were excised carefully with a blade and transferred into sterile 1.5 ml microcentrifuge tubes. The DNA was then extracted from the gel by following the instructions provided by the manufacturer and stored in elution buffer which came with the kit at -20^{0}C.

2.7.4 Quantification of DNA

DNA preparations (plasmid, cosmid, fosmid or genomic DNA), gel purified DNA fragments and PCR primers were quantified and checked for purity either by using a UV-spectrophotometer (UNICAM UV/Vis Spectrophotometer UV3, ATI Unicam, Cambridge, UK) or a NanoDrop® ND-1000 Spectrophotometer (NanoDrop Technologies, Inc., Wilmington, DE). When measuring with the Nanodrop, sterile Milli-Q™ water was read as blank and then 1.5 to 2.0 µl of DNA preparation was read. The instrument calculates the concentration of DNA and provides results as ng of DNA/µl and also provides A_{260}:A_{280} ratio of the sample as a measure of purity. When UV/Vis Spectrophotometer was used, sterile Milli-Q™ water was used to blank the instrument, and the DNA sample diluted with Milli-Q™ water was scanned at 220 nm to 320 nm. Absorbance at 260 nm was measured and the concentration of

DNA was calculated by multiplying the absorbance with the dilution factor and 33 and expressed as µg of DNA /ml.

2.8 Cloning of DNA

2.8.1 Cloning of endonuclease digested fragments

Plasmids were routinely used as vectors for cloning of DNA fragments. The required plasmid was digested with the appropriate restriction endonuclease(s) under the conditions specified by the enzyme manufacturer. On some occasions the digested plasmid was also treated with Shrimp Alkaline Phosphatase (SAP) (Fermentas Inc.) which was inactivated as per the manufacturer's instruction before the linearized plasmid was used for ligation reactions. The progression of digestion was checked by agarose gel electrophoresis. The DNA fragment that was to be cloned was digested with the same enzyme or with another enzyme that gave compatible ends for ligation. The desired DNA fragment was fractionated by gel electrophoresis and identified by comparing with molecular marker. The DNA fragment was gel purified, quantified and mixed with linear vector (suitable plasmid) in 3:1 ratio. A typical 20 µl ligation mix contained 1 µl of T4 DNA ligase in 1x T4 DNA ligation buffer and incubated at 16^0C for 16 h. After the ligation, 1-3 µl of the reaction mixture was transformed into *E. coli* DH5α electrocompetent cells and the resultant transformants were isolated on LB plates containing appropriate antibiotic(s).

Shot-gun cloning was used when all the fragments of a large DNA clone were required to be cloned. In this instance the fosmid clones were digested with

appropriate restriction enzymes and the progression of digestion was checked by gel electrophoresis. The digested DNA was extracted with an equal volume of buffered phenol-chlorofom followed by two equal volume chloroform washes to remove the restriction enzymes. The DNA fragments were precipitated and re-dissolved in a suitable volume of sterile double distilled water. The concentration of the digested DNA was determined and ligated with linearized pUC119 (digested with the same restriction enzyme) according to the procedure describe above. After the ligation reaction, 3 µl of the reaction mix was transformed into *E. coli* DH5α electrocompetent cells and the resultant transformants were isolated by blue-white selection on LB plate containing ampicillin 100 µg/ml.

2.8.2 Direct cloning of PCR products

The cloning of a PCR product in a plasmid vector was carried out using the pCR® 2.1-TOPO kit (Invitrogen, Carlsbad, CA). A PCR product was directly cloned into TOPO vector by following the manufacturer instruction and the resulting transformants were isolated on LB agar plates by blue-white selection.

2.9 DNA sequencing and sequence analysis

Sequencing of DNA was routinely carried out using the BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, California, USA). A reaction mixture of 20 µl was prepared with 250ng of plasmid DNA or 100 ng of gel purified DNA fragments, 4 µl of BigDye reaction ready mix, 4 µl of 2.5x BigDye sequencing buffer, 5 pmol of primer. The sequencing reaction was carried

out for 30 cycles using a T-Gradient thermocycler with conditions set as follows: dissociation temperature 96^{0}C for 30 sec, annealing temperature varied depending upon the primer used, and elongation at 60^{0}C for 1 to 2 min depending upon the length of the template. After the reaction cycle the reaction mixture was transferred to a 1.5 ml microcentrifuge tube and 2 µl of BigDye salt solution and 80 µl 4^{0}C 95% ethanol were added and held at 4^{0}C for 20 min. The sample was then centrifuged at 15,000 rpm in an Eppendorf microcentrifuge for 5 min and the pellet was washed with 1 ml of 70% ethanol, and air dried for 5 min. The microcentifuge tube containing the dried pellet was sent to MBSU facility (Department of Biological Sciences, University of Alberta) to obtain the sequence information. The nucleotide sequence was analyzed and compiled using computer programs including GeneTool 2.0, Lasergene 7.0 and online BLAST (http://blast.ncbi.nlm.nih.gov/).

2. 10 Colony lifts

In order to screen the clones from the *P. polymyxa* PKB1 genomic library prepared in pSMART fosmid during the fusaricidin study by Jingru Li, clones were patched with sterile toothpicks on LB agar plates containing ampicillin 100 µg/ml. The 12x12 inch^2 plates were divided into 625 sections to keep track of individual clones. The LB plates were incubated at 37^{0}C for 16 h and chilled at 4^{0}C for 30 min before the colony lift. Hybond-N nylon membranes of appropriate size were carefully placed on the agar surface of the LB plates to transfer the growth of colonies to the membranes. The position of the membranes was marked

on the back of the plate and one corner of the membrane was trimmed, so that the membrane can be superimposed after processing. After 1 min the membranes were removed from the plates and were subjected to a series of treatments to lyse the cells and release the fosmid DNA onto the membranes. The membranes were first placed with the colony side up onto a 3MM paper saturated with 10% SDS for 1 min. The membranes were then soacked with denaturation solution (0.5 M NaOH, 1.5 M NaCl) for 5 min followed by soaking with neutralization solution (1.5 M NaCl, 0.5 M Tris-HCl, pH 7.5) for 2×3 min. The membranes were washed vigorously in 2x SSC solution (0.3 M NaCl, 0.03 M tri-sodium citrate) to remove the cell debris and air dried. The membranes were then baked at 80^0C for 2 h to fix the DNA. Once the DNA was fixed the membranes were stored at room temperature in a sealed plastic bag.

2.11 Transfer of DNA onto nylon membranes for Southern analysis

In order to identify *P. polymyxa* genes involved in polymyxin biosynthesis, fragments of DNA from fosmid clones were subjected to Southern analysis. The fosmid clones were digested with restriction endonucleases, fractionated by agarose gel electrophoresis, and stained minimally with ethidium bromide so that any hybridizing bands could be tracked after Southern analysis. Before transferring DNA fragments from the agarose gels to nylon membranes, the gels were treated by soaking in depurination solution (0.125 M HCl) for 10 min with gentle agitation followed by soaking in denaturation buffer (1.5 M NaCl, 0.5 M NaOH) for 30 min with gentle agitation. Gels were then washed with

neutralization buffer (1.5 M NaCl, 0.5 M Tris-HCl, pH 7.5) for 30 min with gentle agitation and transferred onto 3M blotting paper soaked with 20x SSC solution. Hybond –N- nylon membranes were placed on the agarose gels and the fractionated DNA was transferred by capillary blotting overnight onto the membranes. The membranes were then marked for proper position, transferred to a vacuum oven and baked for 2 h at 80^{0}C to fix the DNA. The membranes were used immediately or stored at room temperature in plastic bags for future use.

2.12 Preparation and Labelling of DNA probes

In this study a non-radioactive method was used to label DNA probes by using the DIG High Prime DNA Labelling and Detection Starter Kit I (Roche Applied Science). In this method digoxigenin (DIG), a steroid hapten, was used to label the DNA probes. The hybridized probes were detected by immunoassay with anti-digoxigenin-AP Fab fragments and visualized with the colorimetric substrate. To label probes, gel purified DNA fragments of interest were quantified and 1 µg amounts were taken and diluted to a volume of 20 µl with sterile double distilled water. The DNA was then labelled by following the manufacturer's instructions with Dig-High Prime according to random primed labelling technique to make DIG-labelled DNA.

2.13 Hybridization of nylon blots

Nylon membranes containing the fixed DNA were transferred and washed into a hybridization bottle (Robbins Scientific) containing DigEasy hybridization

solution for 30 min in a hybridization chamber. Membranes were then incubated with DIG-labelled DNA probes and processed according to the manufacturer's instructions. After processing, the DNA bands that hybridized with the specific probes were identified by superimposing with the ethidium bromide stained photo of the original gel or with the plate containing the colonies (In the case of colony hybridization).

2.14 Generation of *P. polymyxa* PKB1 mutants

P. polymyxa mutants were generated using REDIRECT© technology: a PCR-targeting system originally developed for *Streptomyces* adapted for *Paenibacillus* by Jingru Li (Li et al. 2007). In this method a cosmid clone carrying the gene(s) of interest was transferred into *E. coli* BW25113 with pIJ790 plasmid encoding λ-RED mediated homologous recombination functions. A PCR amplified antibiotic resistance cassette was used to delete or disrupt the gene(s) of interest present on the cosmid clone.

2.14.1 Preparation of cassette for gene disruption

In this study a PCR amplified Apra^{R}Cm^R*oriT* antibiotic resistance cassette was used for PCR-targeted gene disruption. The PCR product was generated by using a gel purified 2.5 kb *Bgl*II fragment from pJL9 plasmid as template and two primers specific to the gene(s) of interest and the antibiotic resistant cassette. The forward primer was 59 bp long with the first 39 bp homologous to the region upstream of the gene to be disrupted including the first codon and the remaining

20 bp was designed to target the disruption cassette. The reverse primer was 58 bp long with the first 39 bp homologous to the region downstream of the gene to be disrupted (may include the stop codon) and the remaining 19 bp was designed to target the disruption cassette. The amplified PCR product contained the apramycin resistance gene *aac(3)IV* (originally from pIJ773 plasmid), chloramphenicol resistant gene (originally from pC194 plasmid), *oriT* of the plasmid RP4, and two FRT sites flanking at the two ends. The 50 µl reaction mix was set up with 50ng of template DNA, 50 pmol of each primer, 50 µmol of each dNTPs, 1 µl of Taq DNA polymerase, 1 µl of 1/50 dilution of pfu polymerase and 1x Taq buffer (200 mM Tris-Hcl pH 8.3, 500 mM KCl, 25 mM $MgCl_2$). The reaction mix was subjected to 30 cycle of amplification under the following PCR conditions: 95^{0}C for 3 min, 10 cycle of 95^{0}C for 45 sec, 50^{0}C for 45 sec and 72^{0}C for 2 min; 20 more cycle of 95^{0}C for 45 sec, 55^{0}C for 45 sec and 72^{0}C for 2 min with a final elongation at 72^{0}C for 5 min.

2.14.2 PCR-targeted gene disruption

The antibiotic resistance cassette used for PCR-targeted gene disruption was an Apra^{R}Cm^R*oriT* disruption cassette, amplified by PCR using a gel-purified 2.5 kb *Bgl*II fragment of pJL9 as template. The pJL9 was constructed during the study of fusaricidin (Li et al. 2007). The amplified linear PCR product was introduced by electroporation into *E. coli* BW25113/pIJ790 containing a cosmid clone containing the target genes. The plasmid pIJ790 encodes the λ-RED function which promoted the homologous recombination between the target genes

and the PCR amplified antibiotic resistant cassette. The resulting mutants were screened on agar plates with appropriate antibiotics at 37 ^{0}C to eliminate the temperature-sensitive pIJ790 plasmid. The cosmid clones in which the target genes were replaced with the disruption cassette were isolated and transformed into *E. coli* ET12567/pUZ8002 strain to transfer the mutation to *P. polymyxa* PKB1 by intergeneric conjugation.

2.14.3 Transfer of mutated cosmid by intergeneric conjugation

In this study intergeneric conjugation was done by using *E. coli* ET12567, a non-methylating strain which contains a non-transmissible plasmid pUZ8002, as a donor cell. To prepare for the intergeneric conjugation, *E. coli* ET12567/pUZ8002 containing the mutant cosmid was grown in 50 ml of BHI with appropriate antibiotics at 37^{0}C and wild type *P. polymyxa* PKB1 was grown in 50 ml of BHI at 28^{0}C at 250 rpm for about 4-6 h until both cultures reached an OD_{600} of 0.4 to 0.6. The cultures were then harvested by centrifuging at 3000 × g for 5 min and washed separately three times with equal volumes of BHI. Both the donor and recipient cells were re-suspended in 1 ml of BHI and then mixed together in 3:1 ratio. The mixture (a 0.2 ml amount) was then transferred onto a 0.22 µm HAWP filter (13 mm in diameter) (Millipore Corporation, Bedford, MD) placed on a BHI agar plate and incubated at 28 ^{0}C for 16 -18 h. After the incubation the filters were transferred from the BHI plate into a tube and the cell mixture was re-suspended in 1ml of BHI broth. Two hundred microlitre aliquots of the culture were spread on GB agar plates containing chloramphenicol (5

µg/ml) a nd pol ymyxin B (25 µ g/ml) a nd i ncubated a t 37 ^{0}C f or 24 h. Chloramphenicol w as u sed t o s elect t he *P. polymyxa* PKB1 ex-conjugants and polymyxin B t o c ounter s elect the donor *E. coli* ET12567/pUZ8002 s train. Chloramphenicol r esistant colonies that a ppeared a fter the i ncubation w ere patched on a fresh BHI plate containing higher concentrations of chloramphenicol (10 µ g/ml) a nd pol ymyxin B (50 µ g/ml) t o c onfirm t he c hloramphenicol resistance of t he *P.polymyxa* conjugants a nd t o e liminate any remaining *E. coli*. The ex-conjugants w ere c hecked for t he presence of t he mutation i n the ge nome by PCR.

2.15 Production of Antibiotics

To assess a ntibiotic p roduction, the w ild t ype and m utant s trains of *P. polymyxa* PKB1 w ere gr own i n G lucose-Starch-Calcium c arbonate (G SC) medium (glucose 20 g/ l, s tarch 20 g/ l, $(NH_4)_2SO_4$ 20 g/l, ye ast e xtract 10 g/ l, K_2HPO_4 2.6 g/ l, $FeSO_4 \cdot 7H_2O$ 0.1 g/ l, M $gSO_4 \cdot 7H2O$ 0. 5 g/ l, NaCl 0.25 g/ l, $CaCO_3$ 9.0 g/ l) i n a r otary s haking i ncubator f or 3 d at 3 0^{0}C a t 250 rpm. T he culture w as ha rvested by c entrifuging a t 3 000 × g a t 4 ^{0}C f or 10 min and t he supernatant was mixed with methanol to give a final concentration of methanol of 20% (v/v) and held at 4^{0}C for 30 min. The methanol-culture supernatant mix was centrifuged at 15,000 rpm in be nch-top c entrifuge (Eppendorf 5424) and the supernatant was ana lyzed for anti bacterial ac tivity by agar dif fusion bioassay against *E. coli* strains a nd a lso a nalyzed by high pe rformance l iquid chromatography (HPLC).

2.16 Concentrating bioactive compounds

Bioactive compounds from the culture supernatants were concentrated to improve the HPLC analysis. Ten milliliters of culture supernatant containing 20% methanol was passed through a Sep-Pak C_{18} cartridge (Waters Corporation) that had been activated by passing 5 ml of methanol followed by 5 ml of water. The loaded cartridge was then washed with 5 ml of water and the sample was eluted with 3 ml of 80% methanol followed by 2 ml of 100 % methanol. The two methanol eluants were combined and then dried completely under a stream of compressed air while being held at 4^0C. The residue was re-dissolved in 0.5 ml of the same mobile phase as that used for HPLC analysis.

2.17 Agar diffusion bioassay

2.17.1 Antibacterial activity

Production of antibacterial compounds by wild type and different mutant *P. polymyxa* PKB1 strains was checked by agar diffusion bioassay against *E. coli* DH5α and *E. coli* ESS strains as representative of Gram negative bacteria. *E. coli* DH5α and *E. coli* ESS strains were grown in 2 ml amount of LB at 37^0C for 16 h and 100 µl of the cultures was added to 100 ml of molten LB agar cooled down to 45^0C and plates were prepared. An 8 mm sterilized borer was used to punch holes in the agar medium and 100 µl of test solution was applied in each well. The plates were then incubated at 37^0C for 16 h and the zones of inhibition were measured.

2.17.2 Antifungal activity

The antifungal activity of fusaricidin produced by *P. polymyxa* PKB1 strains was checked by agar diffusion bioassay against *L. maculans* as an indicator fungus. PDA plates were spread with 20 µl of spore stock of *L. maculans,* wells were punched with a sterile 10 mm borer and 150 µl of culture supernatants were added to each well. The plates were incubated at 28^{0}C for 3 d and the zones of inhibition were measured.

2.18 High performance liquid chromatography

High performance liquid chromatography (HPLC) was used for the separation and identification of bioactive compound present in the culture supernatant and also was used to determine the relative concentration of the compound present in wild type and different mutants. The HPLC analysis was carried out using an Alliance 2695 separation module (Waters, Milford, MA) with Photodiode Array (PDA) detector (Waters Model 996) controlled by Waters Millenium^{32} Software (version 3.20; Waters, Milford, MA). The sample was analyzed by injecting 25-100 µl into a reverse phase Phenomenex® Bondclone C_{18} column (8 mmx100 mm, 10 µm) (Phenomenex, Torrance, CA) with the mobile phase consisting 0.1 M disodium phosphate (pH 3.0 with phosphoric acid) and acetonitrile (77:23) at a flow rate of 2.0 ml/min. Each sample was run for 20-25 min, the absorbance was determined at 212 nm and the data was processed by Waters Millenium^{32} Software (version 3.20; Waters, Milford, MA). All the water used in HPLC system has low conductivity produced by Milli-Q™ ion exchange

filtration system. Fractions are collected for each 30 sec (equivalent to 1 ml of eluent) of a runtime of 25 min with fraction collector (FRAC-100, Pharmacia) and 150 µl of each fraction was assayed for antimicrobial activity against *E. coli* by agar diffusion bioassay.

2.19 Liquid chromatography–mass spectrometry (LC-MS)

Liquid chromatography–mass spectrometry (LC-MS) analysis of culture supernatants of wild type and different mutant *P. polymyxa* was carried out using an Alliance 2695 HPLC system (Waters Corporation) with PDA detector (Waters Model 996) coupled with a single quadrupole spectrometer (Micromass ZMD-2) controlled by MassLynx software. Samples were analyzed by injecting 5 µl to 10 µl into a reverse phase XTerra®MS C_{18} column (2.1 mmx 150 mm, 5 µm) with the mobile phase consisting 0.1% formic acid (pH 3.0 with ammonium hydroxide) and acetronitrile (77:23) at a flow rate 0.2 ml/min for 20 min. The column was kept at a constant temperature of 25°C in a heated column chamber. The UV data for each sample were collected online with a PDA detector set at 212 nm before introduction into the mass spectrometer. The mass spectral data were acquired for the total ion current of the samples ionized over time with electrospray ionization operated in a positive ion mode. The other MS parameters set were source temperature 150^{0}C, desolvation temperature 300^{0}C, cone voltage 15V, and nitrogen gas flow 15 liter/minute. The nitrogen gas was provided from a Whatman nitrogen gas generator with a constant pressure of 90 psi. Full mass spectra were acquired over an extensive mass range of m/z 450 to 1400 for 20

min with a scan time of 1.0 second and inter scan time of 0.1 second to obtain the total information of the sample composition. The spectra were analyzed using Masslynx™ 4.0 (Waters corporation) software.

Table 2.1.1: Organisms used in this study.

Strains	Description/ Relevant information	Source/ Reference
E. cloni	Host for fosmid library	Lucigen
E. coli DH5α	Host for plasmid and cosmid vectors used for general cloning experiments. Indicator for agar diffusion bioassay of polymyxin	Gibco BRL
E. coli ESS	Super sensitive strain of *E. coli* for agar diffusion bioassay of polymyxin	A.L Demain, Drew Univ.
E. coli BW25113	Host for PCR-targeted mutagenesis containing pIJ790 plasmid carrying gene encoding λ RED mediated homologous recombination functions	Gust et al. 2003
E. coli ET12567/pUZ8002	Methylation-deficient host used for inter-generic conjugation containing pUZ8002 carrying gene encoding the plasmid mobilization functions	Kieser et al. 2000
E. coli DH5α/ BT340	Host containing a FLP recombination temperature sensitive plasmid BT340	Gust et al. 2003
One Shot® TOP10	Host for pCR® 2.1-TOPO® vector used for direct cloning of PCR products	Invitrogen
P. polymyxa PKB1	Wild type strain	Alberta Research Council
P. polymyxa Δ*pmxC*$_{PKB1}$	*pmxC*$_{PKB1}$ deletion mutant (lacks one of the transporter-like genes)	This study
P. polymyxa Δ*pmxD*$_{PKB1}$	*pmxD*$_{PKB1}$ deletion mutant (lacks one of the transporter-like genes)	This study
P. polymyxa Δ*pmxC+D*$_{PKB1}$	*pmxC+D*$_{PKB1}$ deletion mutant (lacks both transporter-like genes)	This study
P. polymyxa Δ*pmxE*$_{PKB1}$	*pmxE*$_{PKB1}$ disrupted mutant, cannot produce polymyxins	This study
P. polymyxa Δ*pmxC+D*$_{PKB1}$-IF	*pmxC+D*$_{PKB1}$ in-frame deletion mutant (lacks antibiotic resistance cassette and both transporter-like genes)	This study
P. polymyxa Δ*fusA*	*fusA* disrupted mutant, cannot produce fusaricidin	Li et al. 2007
L. maculans	Indicator organism for fusaricidin bioassay	Alberta Research Council

Table 2.1.2: Cloning vectors and recombinant plasmids used in this study.

Vectors	Description/ Relevant information	Source/ Reference
pCR2.1 TOPO	Linear phagemid vector for direct cloning of PCR product	Invitrogen
pIJ773	A pBluescript plasmid containing gene disruption cassette consists of the apramycin resistant gene *aac(3)IV* and the *oriT* flanked by two FRT sites.	Gust et al. 2003
pIJ790	Plasmid conaining the λ RED recombination genes, Cm^R	Gust et al. 2003
pJL9	pIJ773 modified with incorporation of *cat* gene from pC194 between the *oriT* and the *aac(3)IV*.	Li et. al. 2007
pSL1180	General cloning vector	Amersham
pSMART-FOS	Copy control fosmid vector used for making genomic library	Invitrogen
pUC119	General cloning vector with blue/ white selection, Amp^R	Vieira and Messing 1987
pUZ8002	Non-transmissible, lack cis-acting function but can transfer function for in-trans conjugation of *oriT* containing replicon.	Gust et al. 2000
Col-8	Completely sequenced cosmid clone containing middle portion of the gene cluster for polymyxin production	Li 2007
Col-9	Cosmid clone containing some area of polymyxin gene cluster	Li 2007
Col-9 Δ$pmxC_{PKB1}$	Col-9 with no $pmxC_{PKB1}$	This study
Col-9 Δ$pmxD_{PKB1}$	Col-9 with no $pmxD_{PKB1}$	This study
Col-9 Δ$pmxC$+D_{PKB1}	Col-9 with no $pmxC$+D_{PKB1}	This study
Col-9 Δ$pmxE_{PKB1}$	Col-9 with no $pmxE_{PKB1}$	This study
Col-9 Δ$pmxC$+D_{PKB1}-IF	Col-9 with in-frame deletion of $pmxC_{PKB1}$ and $pmxD_{PKB1}$	This study

Table 2.7.1.1: Oligonucleotide PCR primers used in this study.

Primers	Sequence 5' to 3' direction	Comments
pSHA01	GCG GAC AGG CTG GGC ATC AC	Forward primer at the 5' end of the polymyxin gene cluster in Col-8 to sequence out with other cosmid clones
pSHA02	CCG GCC CAA GTA CTC AAT ATT TCC	Reverse primer outside the 5' end of the Col-8 based on *pmxA*$_{E681}$ sequence to use with pSHA01
pSHA03	GCG CCA GCA GCA GGT GAC ACT C	Forward primer inside the 3' end of the Col-8 to sequence out
pSHA04	CGC ATA CGT CAG CCG GTC ATT CTC	Reverse primer outside the 3' end of the Col-8 based on *pmxE*$_{E681}$ sequence to use with pSHA03
pSHA05	GAA GCT CAA GAT GCC GCG AAA G	Forward primer to sequence further out after sequence obtain with pSHA03&04
pSHA06	CGG CGA CTG GAT ATT AGC GAT TTT C	Reverse primer to sequence further out after sequence obtain with pSHA03&04
pSHA07	CGT GTA CGA GAA TGA CCG GCT GAC	Forward primer used for sequencing the 3' end of cluster by primer walking
pSHA08	CAT CCA GTG GTG CAG ACG AGA CG	Reverse primer used for sequencing the 3' end of cluster by primer walking
pSHA09	AAC GGA GGT ACG ATG GTG ATT GTG	Forward primer used for sequencing and making the 1.1-kb probe
pSHA10	GGC CAT TTC CTC AAC CGT CG	Reverse primer used for sequencing and making the 1.1-kb probe
pSHA11	CGA GCA GGC CTA CAG CGA CGA TG	Forward primer used for sequencing the 5' end of the gene cluster
pSHA12	GCC GGT CAG CGC GTT TCC TT	Reverse primer used for sequencing the 5' end of the gene cluster

pSHA13	CAG CAA TCG GAC GAG CAG AAA G	Forward primer used for sequencing the 3' end of the gene cluster
pSHA14	CCA CCA GCA AGT CCA CCG AAC	Reverse primer used for sequencing the 3' end of the gene cluster
pSHA15	TAC CGA GTC TGC CGA ACG CAG GGA GGT GGC GTA CAC ATG ATT CCG GGG ATC CGT CGA CC	Forward REDIRECT© primer for PCR-targeted deletion of *pmxD*$_{PKB1}$
pSHA16	ATT CGC AAG CAG GAA GCC GCC CGC CGT AAA GCG GGT CTA TGT AGG CTG GAG CTG CTT C	Reverse REDIRECT© primer for PCR-targeted deletion of *pmxD*$_{PKB1}$
pSHA17	CCG CGT CAA GGT ATA ATC GTC G	Reverse primer used to sequence 5' end by primer walking
pSHA18	AAT GTT CAA AGC TTC GGG GAT ATC	Forward primer used to sequence 3' end by primer walking
pSHA19	TGG CGG CGG TGC TGT ATC TC	Forward primer used to check 9A sub-clone sequence
pSHA20	GCG TGA CGG ACA ACT TCT TCG AC	Forward primer used to check 9A sub-clone sequence
pSHA21	ATG GAC AAT GGA CTG AAG ATG GAC	Forward primer used to sequence 3' end by primer walking
pSHA22	TCA AGC GGC AGA TCG AAG CGG AGG CCG GGG TAC TGC ATG ATT CCG GGG ATC CGT CGA CC	Forward REDIRECT© primer for PCR-targeted deletion of *pmxC*$_{PKB1}$
pSHA23	CGA GTT CTT TCA CTT GCG AGA GCC ATC CGC CCT TTT TCA TGT AGG CTG GAG CTG CTT C	Reverse REDIRECT© primer for PCR-targeted deletion of *pmxC*$_{PKB1}$
pSHA24	ACG ACG ATT TCC GGC AGT TTC	Reverse primer used to sequence 5' end by primer walking
pSHA25	TCG CCA GCT CCT CGT CGT AG	Reverse primer used to check 9A sub-clone sequence
pSHA26	CTG AAG CGT GAA GTA GAG GTA ATC	Forward primer used to sequence out after pSHA21 sequence

pSHA27	CGG AGA GCG CGG AGG CCT TTT G	Forward primer that used with RED-SEQ-UP primer to check the location of Apra^{R}Cm^{R}*oriT* cassette in place of *pmxD*$_{PKB1}$
pSHA28	CAC GCT GGC CGG TTT TTC GCT TAC	Reverse primer that used with RED-SEQ-DOWN primer to check the location of Apra^{R}Cm^{R}*oriT* cassette in place of *pmxD*$_{PKB1}$
pSHA29	GGG ATT TGC TGT CGC GGG TGA AC	Forward primer to check the presence of *pmxD*$_{PKB1}$
pSHA30	CCG CCC GAC AAC GAC GCT CC	Reverse primer to check the presence of *pmxD*$_{PKB1}$
pSHA31	CAA AAG CCA TCT GGA TCC TAA GC	To check upstream of 5' end of cluster (primer position:10 bp down stream of cluster)
pSHA32	CGG ACG TGC TCA TGG TGG ACT CG	Forward primer that used with RED-SEQ-UP primer to check the location of Apra^{R}Cm^{R}*oriT* cassette in place of *pmxC*$_{PKB1}$
pSHA33	ACC CGC GAC AGC AAA TCC CCA T	Reverse primer that used with RED-SEQ-DOWN primer to check the location of Apra^{R}Cm^{R}*oriT* cassette in place of *pmxC*$_{PKB1}$
pSHA34	CGC GGA TCA ACA ACG ACC TGC AG	Forward primer to check the presence of *pmxC*$_{PKB1}$
pSHA35	GCG GCT TTG GCG GCT TCG ATA AT	Reverse primer to check the presence of *pmxC*$_{PKB1}$
pSHA36	TAT GGA TTA ACG CAA GCC CAG CGC CGA ATA TGG TTC ATG ATT CCG GGG ATC CGT CGA CC	Forward REDIRECT© primer for PCR-targeted deletion of *pmxE*$_{PKB1}$
pSHA37	CAG CCG TAT AAG AGG TTC GGA TCA ATC CGG TCA TCC TTC TGT AGG CTG GAG CTG CTT C	Reverse REDIRECT© primer for PCR-targeted deletion of *pmxE*$_{PKB1}$
pSHA38	CCG TGA ATT TAC TGA ACA GGA AGC	Forward primer used to sequence downstream of 3’ end obtained with pSHA26

pSHA39	CGC CCG GAA CCC CTA CAG GAA CC	Forward primer that used with RED-SEQ-UP primer to check the location of Apra^{R}Cm^{R}*oriT* cassette in place of *pmxE*$_{PKB1}$
pSHA40	GCT GGC GCT CGG GGG TTA TTA CG	Reverse primer that used with RED-SEQ-DOWN primer to check the location of Apra^{R}Cm^{R}*oriT* cassette in place of *pmxE*$_{PKB1}$
pSHA41	ACG CAG TGG TTC GAA GAG CCG GAG	Forward primer to check the presence of *pmxE*$_{PKB1}$
pSHA42	GCC GTG TTG TTC GCG TAA ATC CTG G	Reverse primer to check the presence of *pmxE*$_{PKB1}$
pSHA43	AGC CGC GTC AAG GTA TAA TCG TC	48 bp downstream from the beginning of *pmxA*$_{PKB1}$
pSHA44	CCC GAG CAT CCG CGT GAA TC	Forward primer to check sequence immediately after 3' end of Col-8 sequence
pSHA45	CGC TTC CTT CAC GCC ACC TAC AC	Reverse primer to check sequence immediately after 3' end of Col-8 sequence
Kan replace fwd	AGA TCT GAT CAA GAG ACAGGA TGA GGA TCG TTT CGC ATG ATT CCG GGG ATC CGT CGA CC	Forward REDIRECT© primer for PCR-targeted deletion of *kan* of the SuperCos1
Kan replace rev	TCG CTT GGT CGG TCA TTT CGA ACC CCA GAG TCC CGC TCA TGT AGG CTG GAG CTG CTT C	Reverse REDIRECT© primer for PCR-targeted deletion of *kan* of the SuperCos1
CAN 68	CGC CAG GGT TTT CCC AGT CAC GAC	Forward primer specific to the MCS of pUC119 for sequencing the end of the insert cloned into pUC119
CAN 69	GAG CGG ATA ACA ATT TCA CAC AGG A	Reverse primer specific to the MCS of pUC119 for sequencing the end of the insert cloned into pUC119
RED-SEQ-UP	CTG CAG GTC GAC GGA TCC	Reverse primer specific to near the 5' end of Apra^{R}Cm^{R}*oriT* cassette to check it location in cosmid or genomic DNA

RED-SEQ-DWN	CGA AGC AGC TCC AGC CTA C	Forward primer specific to near the 3' end of Apra^{R}Cm^{R}*oriT* cassette to check it location in cosmid or genomic DNA
M13 Forward	GTA AAA CGA CGG CCA G	Forward primer in *lacZα* gene in pCR® 2.1-TOPO® vector to sequence the end of the PCR product cloned into it.
M13 Reverse	CAG GAA ACA GCT ATG AC	Reverse primer in *lacZα* gene in pCR® 2.1-TOPO® vector to sequence the end of the PCR product cloned into it.

3. Results and Discussion

3.1 Background from previous studies

Paenibacillus polymyxa PKB1 is an environmental isolate studied for its ability to produce antimicrobial peptides (Beatty and Jensen 2002). During a previous investigation of the fusaricidin synthetase gene cluster, a genomic library of *P. polymyxa* PKB1 was constructed by partial digestion of genomic DNA with S*au*3A1 and packaging in SuperCos1, a cosmid vector (Li 2007). Screening of the cosmid genomic library with a PCR product probe, designed to detect Thr-specific non-ribosomal peptide synthetase (NRPS) gene fragments (part of adenylation domains) yielded many positive clones, as expected, since Thr is present in both fusaricidin and polymyxin. Initially, one of the positive clones, Col-8, was believed to contain genes encoding the NRPS for fusaricidin biosynthesis because it was the only clone that hybridized with both Thr- and Asn-specific NRPS gene fragment probes. The fusaricidin molecule contains both Thr and Asn. Therefore, as a part of the fusaricidin study, the Col-8 insert was completely sequenced, but it was found to contain putative polymyxin-producing genes rather than fusaricidin-producing genes.

Further analysis as part of the present project showed that the cluster appeared to be incomplete in Col-8 and the order of the genes suggested that a rearrangement may have occurred. Southern analysis of a Col-8 digest in comparison with a genomic DNA digest revealed Col-8 has non-contiguous chromosomal fragments. Although Col-8 appeared to be rearranged, restriction digestion pattern mapping and Southern analysis of the remaining hybridizing

clones revealed four cosmid clones, Col-9, Col-10, Col-15, and Col-20, to contain fragments overlapping with that of Col-8 and therefore to be candidates for containing a putative polymyxin-producing NRPS.

3.2 Analysis of the cosmid clones

The five cosmid clones (Col-8, Col-9, Col-10, Col-15 and Col-20) obtained from the previous study were transformed into *E. coli* DH5α cells to analyze them further. The *E. coli* cells containing the different cosmid clones were grown in LB medium with appropriate antibiotics at 37^0C for 16 h in a tube roller. The cosmids were isolated from the resulting cultures and digested with *Eco*R1, and restriction digestion patterns were compared (Figure 3.2.1). A 14.1-kb *Eco*R1 fragment was found to be common to all five clones and a 7.5-kb fragment was common to four clones but not Col-8. Since the complete nucleotide sequence of Col-8 was known, the *Eco*R1 sites in the Col-8 sequence were analyzed in silico using the software Laser Gene 7 and Gene Tools 2. The in-silico *Eco*R1 restriction map matche the actual *Eco*R1 digest of Col-8 except for the absence of the 6.9-kb fragment which is present in the actual digest but not in the in silico map. This 6.9-kb fragment represents the linear cosmid vector that was edited out of the Col-8 sequence during making the contig.

At this time there was no sequence of polymyxin encoding genes available for comparison within GenBank. After a detailed literature search a patent application was found claiming to identify a gene cluster encoding polymyxin biosynthesis genes, one of the outcomes of full genome sequencing of a natural

isolate of *P. polymyxa* E681 (Choi et al. 2009). In this patent application the sequences of three ORFs encoding NRPS genes, named *pmxA*, *pmxB* and *pmxE,* were given. In the present study the genes from *P. polymyxa* E681 will be designated as pmx_{E681}. The sequences of these genes, obtained by scanning a printout of a pdf file using character recognition software, were analyzed and compared with the nucleotide sequence of Col-8 using the software Gene Tools 2.0 and the NCBI BLAST program (on-line), and the corresponding regions were found to be very similar with a nucleotide sequence homology of about 92%. It was also found from the sequence comparison that the *Eco*R1 fragment from the 3' end of the insert in Col-8, which was 5.1 kb in size, should probably have been a 7.5-kb fragment as found in $pmxE_{E681}$ (Figure 3.2.2). Since all of the other cosmid clones except Col-8 have a 7.5-kb fragment in common, it was considered to be the same fragment as found in the simulated *Eco*R1 digest of $pmxE_{E681}$. Further sequence analysis and comparison to pmx_{E681} sequence revealed that Col-8 has only the middle portion of the putative polymyxin gene cluster and it appears to lack about 5.5 kb of sequence from the 3' end of *pmxE* and 2 kb from the 5' end of *pmxA*.

3.3 Primer walking with cosmid clones

Primer walking was chosen as the initial approach to sequence beyond the Col-8 boundaries. Primers were designed based on the known sequence of Col-8 and on the pmx_{E681} sequence available in the patent application, as the sequences of the pmx_{E681} genes were very similar to those of Col-8. Two primers, pSHA03

(homologous to an area upstream of the 3`end of the 5.1 *Eco*R1 fragment of Col-8) and pSHA04 (homologous to an area in the 7.5-kb *Eco*R1 fragment from $pmxE_{E681}$), were designed and used to amplify a PCR product with Col-15 and Col-20 as templates. A PCR product of the expected size of 750 bp was observed, gel purified and sequenced with the same primers. The sequence of the PCR product from both Col-15 and Col-20 was found to be homologous to the sequence of the corresponding region of $pmxE_{E681.}$ More primers were designed and used to amplify and sequence the remaining part of the 7.5-kb *Eco*R1 fragment using the cosmid clones Col-15 and Col-20 as templates. Since the sequences of the *P. polymyxa* PKB1 PCR products were found to be highly similar to the $pmxE_{E681}$ sequences, the reverse primer of a set of primers was designed based solely on $pmxE_{E681}$ sequence to amplify and sequence a fragment extending from within, to beyond the 3' end of the 7.5-kb *Eco*R1 fragment using the *P. polymyxa* PKB1 cosmid clones as template. The sequence revealed the presence of an *Eco*R1 site in a position, which supported the proposal that the 5.1-kb *Eco*R1 fragment of Col-8 was originally a 7.5-kb *Eco*R1 fragment in the PKB1 genome. Since the $pmxE_{E681}$ sequence is very similar to the sequence obtained by primer walking from Col-15 and Col-20, it was speculated that the downstream sequence would also be very similar, but a comparison of *Eco*R1, *Sal*I, *Sca*I and *Bst*Z171 restriction maps of these two cosmid clones with the in silico restriction map of the $pmxE_{E681}$ sequence did not support the speculation.

Primers were also designed to obtain additional DNA sequence information from upstream of the discontinuity located near the 5' end of the Col-

8 insert in order to complete the *pmxA* gene sequence, but no useful information was obtained using any of the cosmid clones as template. This suggested that these other cosmid clones may also contain discontinuities in this region. The discontinuities and rearrangements observed in Col-8 were not a new feature in the PKB1 cosmid library. This ambiguity was also observed during the study of fusaricidin in *P. polymyxa* PKB1 (Li 2007).

3.4 Screening of fosmid library

Since none of the cosmid clones were found to be sufficient to complete the sequence of the entire polymyxin gene cluster, a second genomic library of *P. polymyxa* PKB1, prepared in the copy control fosmid, pSMART was considered to be an alternative. This pSMART fosmid genomic library was also made during the study of the fusaricidin synthetase gene cluster (Li 2007), and was relatively small containing only 550 clones. Two different probes were used to screen the entire fosmid library to identify the clones possibly containing the areas needed to complete the sequence of the entire NRPS gene cluster for polymyxin biosynthesis (Figure 3.4.1). The first probe was a 1.0 kb DNA fragment obtained from a 1.11-kb PCR product generated with primers pSHA09 and pSHA10 using Col-15 as a template. The PCR product was cloned into the Topo 2.1 vector, transformed into *E. coli* DH5α, and grown in LB with appropriate antibiotics. The clone containing the PCR product was digested with *Eco*R1 to completion and the gel purified *Eco*R1 digested PCR product was labeled non-radioactively to use it as a probe. Thus this probe is homologous to the last 1.0 kb from the 3' end of the

7.5-kb *Eco*R1 fragment, which was common in Col-9, Col-10, Col-15 and Col-20. Fifteen clones were found to hybridize with this probe, suggesting that these clones may contain the 3' end of the polymyxin producing NRPS gene cluster. The second probe was the 1.67-kb *Eco*R1 fragment from Col-8, which was considered to be the last piece of legitimate sequence near the 5' end of the polymyxin producing NRPS gene cluster, present in the Col-8. The 1.67-kb *Eco*R1 fragment was sub-cloned into pSL1180, transformed into *E. coli* DH5α, and grown in LB with appropriate antibiotics to provide increased amounts of the fragment. Eventually the fragment was excised from the sub-clone and labeled non-radioactively to use as a probe. The boundary of the legitimate sequence of the polymyxin producing NRPS gene cluster in Col-8 was determined by comparing to the sequence of pmx_{E681}. Eleven clones were found to hybridize with this probe, and 10 of them are common to the clones that hybridized with the first probe. This indicated that 10 clones hybridized with both of the probes and should therefore contain almost the entire or the entire gene cluster encoding the NRPS for polymyxin biosynthesis. This is an unusually high number of clones apparently containing the desired DNA insert to be found in a relatively small library (550 clones). However, cultures of all 16 of the hybridizing fosmid clones were grown in TB medium containing 1x induction solution (used to increase the fosmid copy number) with appropriate antibiotics and used to isolate fosmids. The fosmid preparations were checked by PCR with the same primers (pSHA09 and pSHA10) used to generate the 1.11-kb probe (Figure 3.4.2). Thirteen clones were found to generate a strong band for the 1.11-kb PCR product, indicating that the

majority of the clones identified from the hybridization results were 'true'. Finding a large number of positive clones in screening the PKB1 genomic library is not unprecedented. Similar results were also seen in the study of fusaricidin. A possible explanation for this could be that *P. polymyxa* PKB1 produces many non-ribosomal peptides, and the NRPSs are organized in modules and domains which have conserved motifs with very similar nucleotide sequences, leading to cross hybridization with the probes. However, it is really very unusual to find the same region of the genome represented in 10 different clones in a small library since the calculated number of clones that would be expected from this library is about one tenth of the number found experimentally.

3.5 Selecting the right fosmid clones

It was not feasible to proceed with all 16 clones for sequence analysis; therefore, to narrow down the selection, all of the positive clones were digested with *Eco*R1 and their restriction patterns were compared to the *Eco*R1 restriction patterns of the five cosmid clones analyzed earlier (Figure 3.5.1). Based on the presence of the 14.1-kb and/or 7.5-kb *Eco*R1 fragments, and the amount of the PCR product generated with pSHA09 and pSHA10, four clones: 1B3, 2E5, 3A8 and 6C12, were chosen as candidates for containing the 3' end of the *pmxE* gene encoding NRPS for polymyxin biosynthesis. The PCR products generated with pSHA09 and pSHA10 using these four clones as templates were sequenced and were found match the sequence of the PCR product generated with the same primers using Col-15 as a template which reinforces the finding that these clones

have genes encoding NRPS for polymyxin biosynthesis. Fosmid 3A8 was chosen as the best candidate for containing the 3' end of the gene cluster and downstream because it did not have the 14.1-kb *Eco*R1 fragment but did have the 7.5-kb *Eco*R1 fragment. This indicated that 3A8 was more likely to contain the DNA sequence from downstream of the 7.5-kb *Eco*R1 fragment. On the other hand, the 4F9 clone was considered to be the best candidate to contain the 5' end and upstream of the gene cluster since it did not give any PCR product with pSHA09 and pSHA10 and it did not have either 7.5-kb or the 14.1-kb *Eco*R1 fragments.

3.6 Shot-gun cloning

Shot-gun cloning and end sequencing of sub-clones was considered to be one of the fastest ways to obtain sequence information from the positive fosmid clones. To sequence fosmid clones 4F9 and 3A8, shot-gun cloning was carried out in the plasmid vector pUC119. The sequencing information from the sub-clones was obtained by end-sequencing the cloned DNA with primers specific to the multi cloning site (MCS) of pUC119.

3.6.1 Shot-gun subcloning of the fosmid clone 3A8 and sequencing

In order to do the shot-gun cloning, the 3A8 fosmid clone was digested to completion with *Eco*R1 at 37º C for 16 h and the fragments were cloned into pUC119. The cloning vector pUC119 was linearized by digesting with *Eco*R1 followed by a Shrimp alkaline phosphatase (SAP) treatment to prevent self-ligation. The resulting sub-clones were transformed into *E. coli* DH5α and the

colonies were isolated on a LB plate containing appropriate antibiotics. Forty-three sub-clones were picked and grown in LB medium containing appropriate antibiotics. Plasmid DNA was prepared from these clones and digested with *Eco*R1 to check the size of the inserts. Based on the sizes of the inserts 13 sub-clones were end-sequenced with primers specific to the MCSs of the vector. The sub-clone 9A, containing a 6.3-kb insert, was found to have sequence homologous to the region immediately following the 7.5-kb *Eco*R1 fragment of the Col-15 cosmid clone, by comparing the sequence with the corresponding region of *pmxE*$_{E681}$. The sequence of the other end of the sub-clone showed no similarity to genes encoding NRPS, which indicated that 9A should contain the remaining 3' end of the gene cluster and beyond.

3.6.2 Sequencing of 9A, a sub-clone from fosmid clone 3A8

The 9A sub-clone was digested with several restriction endonucleases based on the sites available in the MCS of pUC119. *Hin*dIII and *Pst*I were found to give fragments suitable for end-sequencing, and *Pst*I was chosen. The 9A sub-clone was digested with *Pst*I and fragments were again sub-cloned into pUC119. This time the pUC119 was linearized with *Pst*I and treated with SAP before the ligation reaction. Sixteen clones were picked and clones containing the three inserts of the expected size were obtained and end-sequenced using the same primers specific to the MCS of pUC119. The sequence analysis of these clones, when aligned with the *pmxE*$_{E681}$ sequence, indicated that 9A contains almost the entire 3' end of the polymyxin synthetase gene cluster but homologous sequence

stopped at 221 bp short of the expected end of the cluster and switched abruptly to unrelated sequence. This indicated that the sub-clone 9A and hence the fosmid clone 3A 8 contained a non-contiguous fragment of PKB1 genome. This hypothesis was also supported by the fact that the site where the discontinuity appeared was a *Sau*3A1 site, and the PKB1 genomic DNA library was made by partial digestion with *Sau*3A1.

3.6.3 Shot-gun sub-cloning of the fosmid clone 4F9 and sequencing

The fosmid clone 4F9 was digested with *Eco*R1 to completion and the fragments were cloned into pUC119. As described previously, the vector was linearized with the same restriction enzyme, *Eco*R1, and treated with SAP to prevent self ligation. The clones were transformed into *E. coli* DH5α and the colonies were isolated on LB plates containing the appropriate antibiotic. Thirty eight colonies were picked and grown in LB medium containing appropriate antibiotics. Plasmid DNA was prepared from these clones and digested with *Eco*R1 to check the sizes of the inserts. Based on the sizes of the inserts, 16 clones were end-sequenced with the primers, CAN68 and CAN69. The remaining 5' end of the gene cluster encoding NRPS for polymyxin biosynthesis was obtained from one of these sub-clones and the sequence was identified as part of the gene cluster encoding the NRPS for polymyxin biosynthesis by comparison to the homologous area of the pmx_{E681} gene cluster. Unfortunately, a discontinuity was again found in the sub-clone at 23 bp upstream of the 5' end of the gene cluster.

3.7 Searching for other fosmid clones to complete the gene cluster encoding the NRPS for polymyxin biosynthesis

Since the two fosmid clones chosen initially, 3A8 and 4F9, gave only the 5' end of the gene cluster and 23 bp of upstream sequence (4F9), and did not contain the sequence information needed to complete the 3' end of the gene cluster encoding the NRPS for polymyxin biosynthesis (3A8), new fosmid clones were sought. To identify suitable clones to complete the sequence of the gene cluster, all of the positive fosmid clones from the first round of screening were explored either by Southern analysis or by PCR. Unfortunately, internal rearrangement was found to be a major problem for each fosmid and cosmid clone studied.

3.7.1 Searching for other fosmid clones to obtain sequence information beyond the 5' end of the gene cluster

Although there were 16 positive clones found in the initial screening, there was no good candidate containing sequence from beyond the 5' end of the polymyxin synthetase gene cluster. To find an alternative fosmid clone that might be useful to get sequence information from further upstream of the gene cluster, PCRs were carried out with the primers pSHA11 and pSHA12 (Figure 3.7.1.1). These two primers were originally designed to amplify a region 1.1 kb downstream from the beginning of the polymyxin synthetase gene cluster and to provide sequence information to join the 5' end of the gene cluster obtained from fosmid clone 4F9 to the 5' end of the region within the Col-8 sequence derived

from the polymyxin gene cluster. Unfortunately, no fosmid clone except 4F9 generated the expected 1.1-kb PCR product.

3.7.2 Searching for other fosmid clones to obtain the 3' end of the gene cluster

Since the 3A8 fosmid clone was found to have a non-contiguous fragment of PKB1 genomic DNA, the other positive fosmid clones were checked by Southern analysis to identify an alternative fosmid clone containing the 3' end of the gene cluster and sequence downstream. All of the 16 fosmid clones and three cosmid clones were digested again with *Eco*R1 to completion, fractionated in a 0.8% agarose gel and transferred to a nylon membrane. Southern analysis was carried out with a 9A sub-clone specific probe (Figure 3.7.2.1). The probe was a 1.0-kb PCR product generated with the primers pSHA13 and pSHA14 which were used to sequence an internal region of one of the sub-clone of 9A. The specificity of the primers was first checked with fosmid clones 6C12, 1B3 and 2E5, and all of them generated the same 1.0-kb PCR product. The PCR product generated with clone 9A as template was then gel-purified and was labelled non-radioactively. Upon Southern analysis of the digested DNA, many bands from different fosmid clones and cosmid clones were found to hybridize with the probe. As a positive control, the expected non-contiguous 6.3-kb band from 3A8 was labelled with the probe. Again hybridization of this high number of bands from different clones can be explained by the fact that NRPSs have modules with very high sequence homology which makes them very difficult to investigate by

Southern analysis. Four clones were originally chosen from the 16 positive fosmid clones as good candidates for containing the 3' end of the gene cluster encoding NRPS for polymyxin biosynthesis, and all of these four fosmid clones produced the 1.0-kb PCR product that was used as a probe. Of these four clones, two of them, 6C12 and 2E5 (Figure 3.7.2.1) shared a common 5.1-kb band that was considered to possibly represent the true size of the adjacent fragment. A new primer, pSHA18, was designed to hybridize upstream of the discontinuity point of the 9A sub-clone (inside of the polymyxin gene cluster) and it was used for sequencing with both of the fosmids as templates. Sequence information obtained with pSHA18 from both fosmid clones, 6C12, and 2E5, was found to be identical and homologous to the corresponding region of the pmx_{E681} gene cluster. Since the sequence information from both the fosmid clones was identical, the 5.1-kb *Eco*R1 fragment was considered to be the true representative of the PKB1 genome and hence, represents the 3' end of the gene cluster encoding NRPS for polymyxin biosynthesis.

3.8 Sequence analysis of the polymyxin synthetase gene cluster

The complete sequence of the gene cluster encoding the NRPS for polymyxin biosynthesis in *P. polymyxa* PKB1 was obtained by piecing together the sequence information from several cosmid and fosmid clones from PKB1 genomic libraries. None of the cosmid and fosmid clones were found to contain the whole gene cluster and most of them were found to be rearranged. It seemed that NRPS genes were more likely to be ligated together during construction of

the genomic library than other genomic DNA because NRPS containing clones always seemed to be over-represented in the libraries. Perhaps this is somehow due to the high homology of the domain sequences.

Cosmid clone Col-8 was sequenced previously during a study on fusaricidin biosynthesis (Li 2007) and was found to contain most of the polymyxin synthetase gene cluster. However, it is also notable that while Col-8 contained the middle portion of the gene cluster, both of the flanking regions included non-continuous pieces of PKB1 genomic DNA. This rearrangement at the end of the cosmid or fosmid inserts was found in almost all of the clones even though the PKB1 genomic libraries were prepared three times in different hosts with special precautions meant to minimize such problems.

The whole gene cluster consists of five open reading frames and is 40.8 kb in size. Three ORFs (*pmxA*$_{PKB1}$, *pmxB*$_{PKB1}$ and *pmxE*$_{PKB1}$) are the NRPS encoding genes, and the remaining two (*pmxC*$_{PKB1}$ and *pmxD*$_{PKB1}$) encode putative ABC transporter-like genes (Figure 3.8.1). Only 23 bp of sequence upstream from the 5'end of the cluster was obtained due to the lack of appropriate clones and extensive DNA rearrangement problems. About 1.6 kb of DNA was sequenced downstream of the 3' end of the gene cluster by primer walking using fosmid 6C12 and 2E5 as templates and an ORF was found which showed similarity only to hypothetical proteins deposited in the data bank. The protein has a very high homology with the hypothetical protein of *Geobacillus* sp. Y412MC10 (GenBank Accession No.: 003243619) and most similar proteins, as identified by BLAST analysis were not apparently related to NRPS gene clusters.

As discussed earlier, identification of the *pmx*$_{E681}$ polymyxin gene cluster was the outcome of a complete genome sequencing project, and the genome sequence information was only deposited in the GenBank very recently (Accession No. CP000154.1). Complete genome information of yet another *P. polymyx* strain, *P. polymyxa* S2C, was published only in Gen Bank in September, 2010. Since the *pmx*$_{E681}$ sequence is very similar to the *pmx*$_{PKB1}$ sequence, the region downstream from the 3' end of the *pmx*$_{E681}$ gene cluster was examined in the newly released genome sequence and found not to encode a hypothetical protein similar to that located downstream of the 3' end of the PKB1 polymyxin gene cluster.

Since it was not possible to obtain sufficient DNA sequence information flanking the *pmx*$_{PKB1}$ gene cluster to look for additional genes with functions possibly related to polymyxin biosynthesis, the flanking regions of the *pmx*$_{E681}$ were checked for relevant genes. There were no genes found immediately adjacent to the *pmx*$_{E681}$ gene cluster that showed any similarity to genes involved in fatty acid synthesis or in formation of the amide bond between the fatty acid and peptide chain. Similarly no regulatory genes were apparent in the region. Therefore the polymyxin gene cluster, at least in *P. polymyxa* E681, appears to comprise only three NRPS encoding genes and two ABC-transporter encoding genes.

3.9 Modular organization of polymyxin synthetase

The biosynthetic genes for polymyxin production are divided into three ORFs: *pmxA*$_{PKB1}$, *pmxB*$_{PKB1}$ and *pmxE*$_{PKB1}$. *pmxA*$_{PKB1}$ is 14.8 kb in size and encodes 4,918 amino acids, *pmxB*$_{PKB1}$ follows immediately after *pmxA*$_{PKB1}$ and is 3.3 kb, encoding 1,102 amino acids. *pmxE*$_{PKB1}$ is the largest of the all ORFs and consists of 18,759 bp encoding 6,253 amino acids. Sequence analysis of all three ORFs showed that they are the parts of a NRPS. Analysis of the translated product of each ORF with online software programs (http://nrps.igs.umaryland.edu/nrps/ and http://www.nii.res.in/nrps-pks.html) that analyze the domain and module arrangement of NRPS/PKS revealed that these three ORFs contain NRPS domains for production of a peptide composed of ten amino acids (Figure 3.9.1).

3.9.1 Condensation domains

The condensation (C) domain in a NRPS module is responsible for the formation of the peptide bond between two adjacent amino acids activated by the two adenylation domains. NRPS analysis showed that *pmxE*$_{PKB1}$ of polymyxin synthetase contained four condensation domains, *pmxA*$_{PKB1}$ has four and *pmxB*$_{PKB1}$ has one condensation domain. In addition to these nine condensation domains *pmxE*$_{PKB1}$ has one more condensation domain which is not very homologous to other C domains found in the polymyxin synthetase. This N terminal C domain is not very common but has been reported for hybrid NRPS systems which produce lipo-peptides in which this C domain is responsible for the

formation of an amide bond between the first amino acid of the peptide product and a fatty acid molecule. The attachment of the fatty acid by the N terminal C domain is essential and considered to be an early step in lipopeptide biosynthesis. This N terminal C domain is usually present at the very beginning of the NRPS gene immediately upstream of an A domain and followed by a T domain in the initiation module. This kind of initiation module has been reported for NRPS systems, like the gene cluster for arthrofactin synthetase in *Pseudomonas* sp. (Roongsawang et al. 2003). In *P. polymyxa* PKB1 this type of N terminal C domain has been reported for the NRPS system for fusaricidin biosynthesis too (Li et al. 2007). The amino acid sequence homology of the N terminal C domain of polymyxin synthetase with the corresponding C domain of *fusA* is very high (97%). The core motifs of all of the C domains were aligned with the core motifs of a consensus C domain predicted by comparing C domains from many NRPS (Marahiel et al. 1997). The core motifs of all C domains align very well with the consensus core motif sequences and it was also observed that the C domains that are responsible for catalyzing the peptide bonds between L and D amino acid have a higher homology to each other than to other C domains (Figure 3.9.1.1).

3.9.2 Adenylation (A) domain

In an NRPS system adenylation (A) domains are responsible for selecting the right amino acid from the pool and for the activation of the amino acid to incorporate it into the final peptide product. It was reported that in non-ribosomally-synthesized peptides, substrate substitution may be possible at

defined positions and the amino acid composition of the peptide products can even vary depending upon the availability and the relative concentration of the substrate amino acids in the growth medium (Ruttenberg and Mach 1966). However, comparing the amino acids forming the substrate binding pocket of A domain and the crystal structure of the Phe-activating A domain from gramicidin S synthetase allowed researchers to determine the "specificity codon" of an A domain (Stachelhaus et al. 1999; Lautru and Challis 2004). There are several online algorithms available that can analyze a NRPS domain and determine the substrate specificity of the A domain by comparing the substrate binding pocket amino acids to those of other A domain sequences available in GenBank.

In this study two programs, http://nrps.igs.umaryland.edu/nrps/ and http://www.nii.res.in/nrps-pks.html, were used to analyze the modular structure of the polymyxin synthetase and determine the substrate specificity of the A domains (Figure 3.9.2.1). Both of these programs identified the presence of 10 A domains in polymyxin synthetase. Six of the A domains were identified as activating ornithine (orn) by one program http://nrps.igs.umaryland.edu/nrps/ and were unable to be classified by the other program. By comparing the sequence homology of *pmxA,B,C,D,E*$_{PKB1}$ with *pmxA,B,C,D,E*$_{E681}$ it was deduced that these A domains must activate the non-proteinogenic amino acid diamino butyric acid (DAB). Ornithine and DAB share a very similar chemical structure and only differ by one carbon in length (Figure 3.9.2.2). Furthermore, the order and arrangement of the A domains suggested that the A domain for the first amino acid (DAB) of polymyxin is present in the *pmxE*$_{PKB1}$ which also contain three

other A domains incorporating DAB molecules in position 3, 4 and 5 of the final peptide product. The A domains for the first and third position DAB residues share exactly the same specificity codon. The other two A domains for the fourth and fifth position DAB residues also share a common specificity codon but it differs in two positions compared to the specificity codon for the one and three position DAB residues of the polymyxin molecule. Two A domains for the DAB residues in position eight and nine are present in the *pmxA*$_{PKB1}$ gene and share the common specificity codon with the A domain for DAB in positions four and five. Although the module that incorporates DAB in position three of the polymyxin molecule eventually incorporates a D-DAB, there is no difference observed in the specificity codon of this A domain compared to all the rest which incorporate L-DAB residues.

The A domain for the second position amino acid was postulated to activate L-Thr and share the same specificity codon with the L-Thr specifying A domain of bacillomycin.

The A domain for Phe in position six is present in *pmxA*$_{PKB1}$ and shares the same specificity codon with the Phe specifying A domain of gramicidin except for the 331 position amino acid of the substrate binding pocket.

The A domain for the amino acid in position six is present in *pmxA*$_{PKB1}$ and was predicted as activating Leu or Ile by the two different algorithms. The specificity codons for Leu and Ile show remarkable variation in positions 279, 299 and 331 of the substrate binding pocket. By comparing the specificity codons with different Leu and Ile incorporating A domains it was not possible to

determine whether the A domain in position six in polymyxin would activate Leu or Ile, but since Leu is more commonly found in polymyxins, Leu was chosen.

The A domain for the last amino acid of polymyxin is present in *pmxB*$_{PKB1}$ and activates Thr. This A domain shares the same specificity codon as that of the A domain for Thr in position two of polymyxin. In this last module of the NRPS for polymyxin biosynthesis a thioesterase (TE) domain was also observed, which indicates the termination and release of the mature peptide from the synthetase. There is no specific information available in the gene sequence to indicate how the TE domain catalyzes the cyclization reaction between the tenth and fourth amino acid. Since the chemical structure is known for many polymyxins, the structure of the polymyxin produced by *P. polymyxa* PKB1 was deduced from the substrate amino acids predicted by the A domains (Figure 3.9.2.3). Although A domains for the DAB residues have substrate binding pocket sequences that are different from those activating other amino acid substrates, the core motifs of all of the A domains for DAB align very well with each other and with the consensus core motif of an A domain, as predicted by comparing A domains from many NRPS (Figure 3.9.2.4) (Marahiel et al. 1997). The core motifs for the other A domains present in the NRPS for polymyxins showed comparatively poor homology with the standard core motif sequences (Figure 3.9.2.5).

Although the *pmx*$_{PKB1}$ genes generally showed very high similarity to the corresponding *pmx*$_{E681}$ genes, there were two notable regions in *pmxA* where this similarity was much lower. The regions of decreased similarity corresponded to the specificity conferring regions of adenylation domains 6 and 7. Similarity in

these regions was as low as 60% whereas most regions showed similarities >90%. These differences in sequence were reflected in the fact that the first two domains of *pmxA*$_{E681}$ specify different amino acids (D-leu and L-thr) than those specified by *pmxA*$_{PKB1}$ (D-phe and L-leu). Otherwise, the two forms of polymyxin contained the same amino acids, including D-DAB in position 3.

3.9.3 Thiolation (T) domains

There are ten thiolation (T) domains present in the polymyxin synthetase gene cluster. Five T domains are present in *pmxE*$_{PKB1}$, four in *pmxA*$_{PKB1}$ and one in *pmxB*$_{PKB1}$. Six out of 10 carry the unusual amino acid derivative DAB during synthesis of polymyxin. There are two T domains which are upstream of epimerization domains rather than condensation domains. The presence of an E domain after the T domain has been reported for many NRPS systems and these E domains are involved in the biosynthesis of peptides or lipopeptides containing D-amino acids. Each E domain is responsible for the stereochemical conformational changes of the substrate amino acid activated by its related A domain and the epimerization takes place after activation but before the formation of the amide bond by the C domain. Although, the presence of an E domain is the most usual way of incorporating D-amino acid, there have been a few NRPSs reported which do not contain E domains but instead contain A domains that can activate D amino acids directly. The fusaricidin synthetase in *P. polymyxin* PKB1 has one A domain in the sixth module which activates D-Ala without having a E domain in the module (Li and Jensen 2008).

The T domains upstream of E domains share a common core motif LGG**D**SIK which is different from the core motifs of T domains that are not adjacent to E domains. The other six T domains, which do not have any E domains downstream, have the core motif LGG**H**SLR, which is common for T domains without downstream E domains (Marahiel et al. 1997). It was suggested that the aspartate residue (D) immediately before the serine (S) residue is required for proper interaction with the E domain and for the stereochemical conversion of the L-amino acid to its D form (Linne et al. 2001). In addition to the H/D and L/I difference in the core motif, a gap is also shown in the amino acid sequence of T_E domain motifs compared to T domain motif sequences. The same observation also has been reported in the phylogenetic analysis of T_E domains of different NRPS systems from many cyanobacteria (Rounge et al. 2008). Downstream of the last T domain, there is a TE domain, which is involved in the cyclization and cleavage of the polymyxin molecules. This N terminal T domain (T_N) has the motif LGG**N**SLR and it is proposed that the asparagine residue before the serine residue may interact with the TE domain and be involved in the cyclization of the final product. In addition to the difference in core motif, the amino acid sequence of the T_N domain motif is noticeably different compared to other T domains (Figure 3.9.3.1).

3.9.4 Epimerization domain

Non-ribosomally synthesized peptides may have more than 300 different building blocks, including D-amino acids. Usually an L-amino acid is activated

by the A domain, which is then incorporated directly, or less often converted to its D form by an epimerization domain (E) domain. There are two E domains present in the polymyxin synthetase encoding gene cluster of *P. polymyxa* PKB1. The first E domain is present in the $pmxE_{PKB1}$ and involves stereochemical change of the DAB residue at position three. The second E domain is in the $pmxA_{PKB1}$ and involves stereochemical change of the Phe residue at position six. A second mechanism of incorporating D-amino acids has been reported for many NRPS systems, in which L-amino acids can be converted to D-amino acids during incorporation without the presence of any E domain. In this instance the C domain catalyzes the condensation as well as the epimerization reaction (Yin and Zabriskie 2006; Balibar et al. 2005; Roongsawang et al. 2003). None of the well known types of polymyxins have D-DAB in position three, but in this study an E domain was observed for the DAB residue in position three. There has been only one other gene cluster encoding a NRPS for polymyxins reported to date and it also contains an E domain in the third module of the NRPS suggesting a D-DAB in the position three. It seems very unusual that polymyxins have been identified and analyzed from many different bacterial cultures over the years and many different variants of polymyxins have been discovered, but none, except for the very rare polymyxin A, has a D-DAB residue in position three or in any other position. It is possible that other polymyxin producers have gene clusters encoding NRPS for polymyxin that do not contain an E domain at position three or if they do have an E domain, that it is not active. The amino acid sequence analysis of the E domains of polymyxin synthetase in PKB1 revealed that both of

the E domains shared the same conserved motifs (Figure 3.9.4.1). The deduced structure of the most common form of polymyxin, polymyxin B, suggests that it has a D-Phe residue in position six which is consistent with the E domain present in module six of polymyxin synthetase in PKB1, but that is the only D-amino acid residue present in the molecule. Therefore the polymyxins produced by PKB1, which have a D-DAB residue in position three, represent novel forms of polymyxin.

3.9.5 Thioesterase domain

A thioesterase (TE) domain was found in the last module of the NRPS for polymyxin biosynthesis, present in the *pmxB*$_{PKB1}$ gene. TE releases the final peptide product from the NRPS system and sometimes is also involved in the macrocyclization of the peptide product. The TE domain in PKB1 showed high structural similarity with the TE domain of the NRPS for surfactin in which the TE domain catalyzes the formation of a loop of seven amino acids.

3.10 Sequence analysis of the transporter-like genes

The transporter-like genes *pmxC*$_{PKB1}$ and *pmxD*$_{PKB1}$ in *P. polymyxa* PKB1 are similar in size but *pmxC*$_{PKB1}$ is slightly larger; *pmxC*$_{PKB1}$ is 1824 bp in size and *pmxD*$_{PKB1}$ is 1731 bp. These two genes are overlapping by 2 bp, but the nucleotide sequence homology between them is very poor. BLAST analysis using GenBank revealed that the *pmxC*$_{PKB1}$ and *pmxD*$_{PKB1}$ genes were highly homologous (98% at the nucleotide level) to their counterparts, *pmxC*$_{E681}$ and *pmxD* $_{E681}$ from *P.*

polymyxa E681 respectively. Further analysis of the sequence and comparison of the translated amino acid sequences with GenBank showed that both of them are members of the ABC membrane super-family. ABC transporter genes have been reported for many NRPS gene clusters and they have been found to play a role in transport of the corresponding peptide products. The multidomain properties of these two proteins, PmxC and PmxD, are very similar to the multidomain properties of the MdlB protein, which is an ABC-type multidrug transport system containing ATPase and permease components, and which acts as a defense mechanism against antibiotics. The N terminal domain of these two proteins also showed high homology to the MsbA protein, which is an essential ABC transporter involved in lipid-A export in *E. coli* and is closely related to multi-drug resistance proteins of eukaryotic origin. The secondary structure of PmxC and PmxD proteins was determined by using the SOSUI tool (Hirokawa et al. 1998; Mitaku and Hirokawa 1999; Mitaku et al. 2002) and six membrane helices for pmxD and five membrane helices of pmxC were predicted. The ATP binding domain was also observed for both of them at the C terminal region (Figure 3.10.1).

3.11 Mutational analysis

Mutational analysis of the genes present in the gene cluster encoding the NRPS for polymyxin biosynthesis is an approach designed to show their true involvement in polymyxin production. The two transporter-like genes, $pmxC_{PKB1}$ and $pmxD_{PKB1}$ are present in the middle of the NRPS gene cluster for

polymyxin. In an earlier study on *P. polymyxa* PKB1, no ABC transporter-like genes were found in the gene cluster for fusaricidin biosynthesis. Since the $pmxC_{PKB1}$ and $pmxD_{PKB1}$ genes resemble ABC transporter genes, they might have a role in the transport of polymyxin out of the cell into the growth medium. In the cyanobacterium *Microcystis aeruginosa*, inactivation of the ABC transporter gene *mcyH*, which is present in the middle of the microcystin gene cluster resulted in loss of production of the toxin microcystin (Pearson et al. 2004). In *Amycolatopsis balhimycina* inactivation of *tba*, an ABC transporter gene, resulted in accumulation of the glycopeptide antibiotic balhimycin inside the cells to a concentration 10 fold higher than normal without affecting the growth of the organism (Menges et al. 2007). Therefore, to investigate the possible association of the *pmxC* and *pmxD* genes with polymyxin transport or biosynthesis, PKB1 mutants were generated by single and/or double deletion of these two genes. PKB1 mutants were also generated by disrupting the $pmxE_{PKB1}$ to prove the involvement of the *pmx* gene cluster in polymyxin production.

Since the complete sequence of Col-8 was known, it was initially chosen for mutational analysis, but due to unidentified reasons, Col-8 behaved strangely when it was transformed into the *E. coli* BW 25113 strain used for PCR targeted mutagenesis. No colonies were ever obtained after transforming the *E. coli* BW 25113 strain containing the cosmid clone Col-8 with the linear PCR product for PCR-targeting, despite repeated trials. It seemed that for unknown reasons, the cells containing Col-8 lysed extensively upon electroporation in the presence of the linear PCR products. Since $pmxC_{PKB1}$, $pmxD_{PKB1}$ and part of $pmxE_{PKB1}$ are all

contained within the 14.1- kb *Eco*R1 fragment, and since this fragment is known to be present in cosmids Col-9, Col-10, Col-15 and Col-20, these cosmids were checked as alternative templates to carry out the PCR-targeting gene disruption and all were found to be usefull. Eventually Col-9 was chosen to carry out the mutational analysis.

3.11.1 Preparation of the Apra^{R}Cm^R*oriT* cassette for mutational analysis

An Apra^{R}Cm^R*oriT* cassette was constructed during the previous study of fusaricidin biosynthesis in *P. polymyxa* PKB1 and the plasmid carrying the cassette was named pJL9 (Li 2007). This antibiotic resistance cassette was constructed by combining *acc(3)IV*, and, *oriT*$_{RK2}$ from the plasmid, pIJ773 and *cat* from the plasmid, pC194, to give a cassette that allows for independent selection of apramycin resistance (Apra^R, encoded by *acc(3)IV*) in *E. coli* and chloramphenicol resistance (Cm^R, encoded by *cat*) in *P. polymyxa* PKB1. The cassette is also flanked by FRT sites to allow eventual removal by FLP mediated recombination, if desired. To use the Apra^{R}Cm^R*oriT* cassette for PCR targeted mutational analysis it was gel purified from pJL9 as a 2.5-kb *Bgl*II fragment and then amplified by PCR with a Redirect primer pair specific to the sites of the gene to be disrupted (Figure 3.11.1.1). The Redirect forward primer was a 59-bp single stranded oligonucleotide containing 39 bp specific to the 5' site of the target gene followed by 20 bp specific to the 5' end of the cassette. The Redirect reverse primer is a 58 bp oligonucleotide containing 39 bp specific to the 3' site of the target gene followed by 19 bp specific to the 3' end of the cassette. Thus the PCR

amplified disruption cassette will have the antibiotic resistance cassette flanked with DNA sequence homologous to the two ends of the gene to be disrupted.

3.12 Generation of mutants and their phenotypes

3.12.1 Generation of a *pmxD*$_{PKB1}$ mutant

In order to assess the involvement of these transporter-like proteins, the *pmxD*$_{PKB1}$ gene was deleted by the PCR targeting gene disruption method (Gust et al. 2003). The overall schematic diagram for generating a *P. polymyxa* PKB1 mutant by PCR targeted gene disruption is given in Figure 3.12.1.1 using a Δ*pmxC+D*$_{PKB1}$ mutant as an example. The same principle was used for generating the other PCR targeted gene disruption mutants.

Two PCR primers were designed targeting the whole *pmxD*$_{PKB1}$ gene (start codon to stop codon). Using primers pSHA15 and pSHA16, the Apra^{R}Cm^R*oriT* cassette was amplified to produce a PCR product about 2.5 kb in size, which was subsequently transformed into *E. coli* BW 25113. This *E. coli* BW25113 strain carrying the plasmid pIJ790 with the λ RED recombination functions had already been transformed previously with the cosmid clone Col-9 carrying the chromosomal fragment of PKB1 containing *pmxC*$_{PKB1}$ and *pmxD*$_{PKB1}$. Homologous recombination mediated by the λ RED genes between the linear PCR product and Col-9, generated the mutant form of Col-9 (Δ*pmxD*$_{PKB1}$ Col-9) by replacing the *pmxD*$_{PKB1}$ with the Apra^{R}Cm^R*oriT* cassette. Colonies carrying the mutant cosmid were selected on LB agar plates containing appropriate antibiotics. Six of them were chosen, grown in LB medium with antibiotics and cosmid DNA

was prepared. The mutant cosmid from one of these was transformed into *E. coli* ET12567 in preparation for transfer of the mutation from *E. coli* to *P. polymyxa* PKB1 by intergeneric conjugation. The intergeneric conjugation was mediated by $oriT_{RK2}$ which was present on the $Apra^{R}Cm^{R}$*oriT* cassette, and a non-transmissible plasmid pUZ8002, already present in the *E.coli* ET12567 strain. The resulting ex-conjugants were screened on a GB agar plate containing 25 µg/ ml polymyxin and 5 µg/ ml chloramphenicol and were further reconfirmed by patching the colonies on a GB agar plate containing double the concentrations of the same antibiotics. Finally, six *P. polymyxa* PKB1 Δ$pmxD_{PKB1}$ mutants (D1-D6) were grown in BHI without antibiotics, and genomic DNA was isolated. PCR was used to check for the absence of $pmxD_{PKB1}$, the presence of the antibiotic resistance marker, and also the location of the antibiotic resistance marker in the chromosome. All of the PCR conditions were verified using WT PKB1 genomic DNA as a control (Figure 3.12.1.2). The desired deletion mutants resulting from a double crossover event should give rise to a 382-bp fragment when their genomic DNA is amplified with primer pair pSHA27 and RED-SEQ-UP, a 590-bp fragment with primer pair pSHA28 and RED-SEQ-DWN, and no product with primer pair pSHA29 and pSHA30 (1,084-bp fragment in the wild type). The double crossover homologous recombination was found to have occurred in three out of six mutants, and mutant D4 was chosen for further study.

3.12.2 Phenotype of the $pmxD_{PKB1}$ mutant

The D4 *P. polymyxa* PKB1 Δ*pmxD*$_{PKB1}$ mutant and the WT *P. polymyxa* PKB1 were grown in GSC medium for 3 d at 28°C on a rotary shaker at 250 rpm. The culture was harvested by centrifugation at 3000 x g for 10 min at 4°C and the culture supernatant was mixed with methanol to 20%. The culture supernatant methanol mixture was kept at 21°C for 30 min and then centrifuged at 3000 x g for 10 min. Antimicrobial activity was determined by agar diffusion bioassay by applying 100 µl of the culture supernatant mixture to a well cut into an LB agar plate inoculated to 0.1% (v/v) with a 16 h culture of *E. coli* DH5α. The zone of inhibition was measured after 16 h of incubation at 37°C and was found to be 21 mm for WT and 18 mm for Δ*pmxD*$_{PKB1}$ mutant. Based on the reduced zone size for the Δ*pmxD*$_{PKB1}$ mutant, it appears that *pmxD*$_{PKB1}$ is involved in polymyxin biosynthesis, but is not essential.

3.12.3 Generation of the *pmxC*$_{PKB1}$ mutant

The transporter-like genes *pmxC*$_{PKB1}$ and *pmxD*$_{PKB1}$ are both present in the middle of the polymyxin synthetic gene cluster and deletion of *pmxD*$_{PKB1}$ resulted in reduced antibacterial compound production. Therefore, the other transporter like gene, *pmxC*$_{PKB1,}$ was also deleted to check whether this has a similar effect on antibiotic production. The *pmxC*$_{PKB1}$ gene was deleted using the same PCR targeting gene disruption method. Two new PCR primers were designed targeting the entire *pmxC*$_{PKB1}$ gene (start codon to stop codon) and using these two primers pSHA22 and pSHA23, the same A pra^{R}Cm^{R}*oriT* cassette was amplified to produce a PCR product of about 2.5 kb in size. This PCR amplified antibiotic

resistant cassette was subsequently transformed into the *E. coli* BW25113 strain carrying the plasmid pIJ790 with the λ RED recombination functions and cosmid clone Col-9 carrying the chromosomal fragment of PKB1 containing $pmxC_{PKB1}$ and $pmxD_{PKB1}$. Homologous recombination mediated by the λ RED gene between the linear PCR product and Col-9 generated the mutant form of Col-9 ($\Delta pmxC_{PKB1}$ Col-9) by replacing the $pmxC_{PKB1}$ gene with the A $pra^{R}Cm^{R}oriT$ cassette. Colonies carrying the mutant cosmid clone were selected on LB agar plates containing appropriate antibiotics. Six of them were chosen, grown in LB medium with the appropriate antibiotics and cosmid DNA was prepared. The mutant cosmid from one of these was transformed into *E. coli* ET12567/pUZ8002 and subsequently transferred into *P. polymyxa* PKB1 by intergeneric conjugation. The resulting ex-conjugants were screened on a GB agar plate containing 25 μg/ml polymyxin and 5 μg/ml choramphenicol and were further confirmed by patching the colonies on a GB agar plate containing double the concentrations of the same antibiotics. Finally six *P. polymyxa* PKB1 $\Delta pmxC_{PKB1}$ mutants (C1-C6) were grown in BHI without antibiotics and genomic DNA was made. PCR was used to check for the absence of $pmxC_{PKB1,}$ the presence of the antibiotic resistance marker, and also the location of the antibiotic resistance marker in the chromosome. All of the PCR conditions were verified using WT PKB1 genomic DNA as a control (Figure 3.12.3.1). The desired deletion mutants resulting from a double crossover event should give rise to a 457-bp fragment when their genomic DNA is amplified with primer pair pSHA32 and RED-SEQ-UP, a 406-bp fragment with primer pair pSHA33 and RED-SEQ-DWN, and no product with

primer pair pSHA34 and pSHA35 (995-bp fragment in the wild type). The double crossover homologous recombination was found to have occurred in five out of six mutants and C6 was chosen for further study.

3.12.4 Phenotype of the *pmxC*$_{PKB1}$ mutant

The C6 *P. polymyxa* PKB1 Δ*pmxC*$_{PKB1}$ mutant and WT *P. polymyxa* PKB1 were grown in GSC medium and the culture supernatants bioassayed for polymyxin as described in 3.12.2. The zone of inhibition was measured after 16 h of incubation at 37 ^{0}C and was found to be 21 mm for WT and 17 mm for Δ*pmxC*$_{PKB1}$ mutant. Therefore, it appears that *pmxC*$_{PKB1}$ is involved in polymyxin biosynthesis, but like *pmxD*$_{PKB1,}$ is not essential.

3.12.5 Generation of the *pmxC+D*$_{PKB1}$ mutant

Deletion of *pmxC*$_{PKB1}$ and *pmxD*$_{PKB1}$ individually each reduce the ability of *P. polymyxa* PKB1 to produce antibiotic by at least twofold as determined by an agar diffusion bioassay. Therefore they appear to have individual or overlapping roles in the production or release of the antibiotic to the media. To determine whether deletion of both of the transporter-like genes has a more severe effect on polymyxin biosynthesis, both of them were deleted using the PCR targeting gene disruption method. In this case the forward primer, pSHA22 targeting the beginning of *pmxC*$_{PKB1}$ and the reverse primer, pSHA16 targeting the end of *pmxD*$_{PKB1,}$ were used to amplify the Apra^{R}Cm^R*oriT* cassette and the linear PCR product was subsequently transformed into *E. coli* BW 25113 carrying

the plasmid pIJ790 with the λ RED recombination functions and the cosmid clone Col-9 carrying the chromosomal fragment of PKB1 containing *pmxC*$_{PKB1}$ and *pmxD*$_{PKB1}$ and flanking sequences. Homologous recombination mediated by the λ RED genes between the linear PCR product and Col-9 generated the mutant form of Col-9 (Δ*pmxC+D*$_{PKB1}$ Col-9) by replacing the *pmxC*$_{PKB1}$ and *pmxD*$_{PKB1}$ genes with the Apra^{R}Cm^{R}*oriT* cassette. Transformants carrying the mutant cosmid clone were selected on LB agar plates containing appropriate antibiotics. The mutant cosmid was isolated from the *E. coli* BW 25113 strain and transformed into *E. coli* ET12567-pUZ8002 and from there introduced into *P. polymyxa* PKB1 by intergeneric conjugation. The resulting ex-conjugants were screened on a GB agar plate containing 25 μg/ml polymyxin and 5 μg/ml choramphenicol and were further confirmed by patching the colonies on GB agar plates containing double the concentrations of the same antibiotics. Finally, eight *P. polymyxa* PKB1 Δ*pmxC+D*$_{PKB1}$ mutants (CD1-CD8) were grown in BHI without antibiotics and genomic DNA was isolated. PCR was used to check for the absence of *pmxC*$_{PKB1}$ and *pmxD*$_{PKB1,}$ the presence of the antibiotic resistance marker, and also the location of the antibiotic resistance marker in the chromosome (Figure 3.12.5.1 and 3.12.5.2). All of the PCR conditions were verified using WT genomic DNA as a control. True double crossover mutants should not give a product with either of primer pairs pSHA34 and pSHA35, and pSHA29 and pSHA30, whereas the wild type should give 995-bp and 1,084-bp products respectively, and true mutants should give 457-bp and 590-bp products with primer pairs pSHA32 and RED-SEQ-UP and pSHA28 and RED-SEQ-DWN, respectively. Double

crossover homologous recombination was found to have occurred in four out of eight mutants, and mutant CD6 was chosen for further study.

3.12.6 Phenotype of the *pmxC+D*$_{PKB1}$ mutant

The CD6 *P. polymyxa* PKB1 Δ*pmxC+D*$_{PKB1}$ mutant and WT *P. polymyxa* PKB1 were grown in GSC medium and the culture supernatants bioassayed for polymyxin as described in 3.12.2. The zone of inhibition was measured after 16 h of incubation at 37 °C and was found to be 21 mm for WT and 17 mm for Δ*pmxC+D*$_{PKB1}$ mutant (Figure 3.12.6.1). Therefore, it appears that *pmxC*$_{PKB1}$ and *pmxD*$_{PKB1}$ are not essential for polymyxin biosynthesis, but they are involved in polymyxin production and that the combined *pmxC+D* mutation does not have any more severe effect on polymyxin production than the *pmxC* mutation alone.

3.12.7 Generation of the *pmxE*$_{PKB1}$ mutant

To verify the role of the NRPS genes *pmxE*$_{PKB1}$ was disrupted by using the PCR targeting gene disruption method. A new set of primers, pSHA36 and pSHA37 was designed targeting about the first 2 kb of the *pmxE*$_{PKB1}$ gene and the Apra^{R}Cm^R*oriT* cassette was amplified with these primers. The PCR product was transformed into *E. coli* BW 25113 carrying the plasmid pIJ790 with the λ RED recombination functions, as well as the cosmid clone, Col-9. Homologous recombination mediated by the λ RED genes, and taking place between the linear PCR product and Col-9, generated the mutant form of Col-9 (Δ*pmxE*$_{PKB1}$ Col-9) by disrupting the *pmxE*$_{PKB1}$ gene with the Apra^{R}Cm^R*oriT* cassette. Transformants

carrying the mutant cosmid were selected on LB agar plates containing appropriate antibiotics. Six of them were chosen, and grown in LB medium with antibiotics, and then cosmid DNA was prepared. The mutant cosmid DNA was transformed into *E. coli* ET12567-pUZ8002 and subsequently was introduced into *P. polymyxa* PKB1 by intergeneric conjugation. The resulting ex-conjugants were screened on a GB agar plates containing 25 μg/ml polymyxin and 5 μg/ml choramphenicol and were further reconfirmed by patching the colonies on GB agar plates containing double concentration of the same antibiotics. Finally eight *P. polymyxa* PKB1 ($\Delta pmxE_{PKB1}$) mutants (E1-E8) were grown in BHI without antibiotics and genomic DNA was made. PCR was used to check for the absence of the replaced part of $pmxE_{PKB1}$ the presence of the antibiotic resistance cassette and also the location of the antibiotic resistance cassette in the chromosome (Figure 3.12.7.1). Primer pairs pSHA40 and RED-SEQ-UP, pSHA39 and RED-SEQ-DWN, and pSHA41 and pSHA42, were expected to give a 358-bp band, a 415-bp band, and no band, respectively, for true mutants, whereas the wild type should give a 797-bp band with primer pair pSHA41 and pSHA42 only. All of the PCR conditions were verified using WT genomic DNA as a control. The double crossover homologous recombination was found to have occurred in all eight of the mutants and E3 and E7 were chosen for further study.

3.12.8 Phenotype of the $pmxE_{PKB1}$ mutant

The E3 and E7 *P. polymyxa* PKB1 $\Delta pmxE_{PKB1}$ mutants and WT *P. polymyxa* PKB1 were grown in GSC medium and the culture supernatants

bioassayed for polymyxin as described in 3.12.2. The zone of inhibition was measured after 16 h of incubation at 37 ^{0}C and was found to be 21 mm for WT and no zone was observed for the Δ*pmxE*$_{PKB1}$ mutant (Figure 3.12.6.1). Based on these results it can be concluded that the *pmx*$_{PKB1}$ gene cluster is definitely involved in polymyxin biosynthesis and under the studied conditions no other antibiotic was produced that inhibited the growth of *E. coli* DH5α. This also implies that the decrease in zone of inhibition sizes seen in *pmxC*$_{PKB1}$ and *pmxD*$_{PKB1}$ mutants must be associated with residual polymyxin production since no other antibiotics are detected under these conditions.

3.12.9 Generation of an in-frame *pmxC+D*$_{PKB1}$ mutant

Deletion of both *pmxC*$_{PKB1}$ and *pmxD*$_{PKB1}$ resulted in a reduction of antibiotic production and from the bioassay, it could be concluded that there was no additive or synergistic effect of deletion of both the transporter-like genes on antibiotic production compared to deletion of *pmxC*$_{PKB1}$alone. However, since *pmxC*$_{PKB1}$ and *pmxD*$_{PKB1}$ are immediately upstream of *pmxE*$_{PKB1}$, the question of possible polar effects can be raised. No transcriptional information is available for this gene cluster, but large multi-cistronic operons are often involved in NRPS production. Since *pmxE*$_{PKB1}$ was shown to be an essential gene for polymyxin production, conceivably, deletion of *pmxC*$_{PKB1}$ and *pmxD*$_{PKB1}$ might have no effect related to the loss of the products of these genes. Rather, the insertion of the Apra^{R}Cm^R*oriT* cassette might be decreasing expression of the downstream *pmxE*$_{PKB1}$ gene leading to decreased antibiotic production. To eliminate the

question of a polar effect, an in-frame Δ*pmxC*+D_{PKB1} *P. polymyxa* PKB1 mutant was made. The protocol described by Gust et. al. (Gust et al. 2003) for making in-frame deletion mutants free of selectable markers was followed with some modification. The schematic diagram for generating Δ*pmxC*+D_{PKB1}-IF-*P. polymyxa* PKB1 mutants is given in Figure 3.12.9.1.

The procedure began with the introduction of the Δ*pmxC*+D_{PKB1} Col-9 mutant cosmid into *E. coli* BT340, an *E.coli* DH5α derivative strain carrying a plasmid that expresses FLP recombinase. FLP recombinase recognizes the FRT sites flanking the antibiotic resistance (Apra^{R}Cm^R*oriT*) cassette and catalyzes FRT site specific recombination to remove the intervening sequence between the FRT sites leaving a 81-bp "scar". Since the number of nucleotides in the scar is a multiple of three, the scar should keep the downstream sequence in-frame, thus eliminating any question of polarity in the downstream region of the gene cluster. The FLP recombinase plasmid is temperature sensitive and once the FLP recombinase mediated recombination had taken place, the plasmid was removed by growing the cells at 42 ^{0}C. Colonies were isolated and checked for loss of apramycin and chloramphenicol resistance indicating that the cassette had been flipped out of the mutant Col-9 cosmid. Because the Cm^R gene is the only suitable antibiotic resistance marker known for *P. polymyxa* PKB1, this created added complications. The resulting Col-9 mutant cosmids now have an 81-bp in-frame scar in place of $pmxC_{PKB1}$ and $pmxD_{PKB1}$, rather than Apra^{R}Cm^R*oriT* cassette, but no selectable marker useful in *P. polymyxa* PKB1 remains. Therefore, a new selectable marker was required before transforming the mutant cosmid with the

'scar' in to *E. coli* ET12567 t o p roceed w ith t he i ntergeneric c onjugation. Streptomycin was ini tially believed to be a s uitable selectable a ntibiotic when used at a c oncentration, 20 µg/ml to i nhibit the gr owth of w ild t ype P KB1. Therefore, a st reptomycin resistance ge ne ca ssette w as i ntroduced into the Δ*pmxC*+*D*$_{PKB1}$Col-9 in-frame mutant in place of the kanamycin resistance marker of SuperCos1 by again following the PCR targeting gene disruption method. Two primers: ' kan r eplace f wd' a nd ' kan r eplace r ev' w ere de signed t o amplify t he streptomycin resistant cassette, isolated by gel pur ification as a *Eco*R1/*Hin*dIII fragment from pI J778, w ith e nds t argeting t he kanamycin r esistance gene. A λ-RED mediated r ecombination be tween t he l inear P CR pr oduct a nd t he Δ*pmxC*+*D*$_{PKB1}$-IF (in-frame) Col-9 mutant took place in *E. coli* BW25113 and the streptomycin resistant f orm of the C ol-9 Δ*pmxC*+*D*$_{PKB1}$-IF mutant cos mid w as isolated from *E. coli* BW25113 a nd t ransformed i nto *E. coli* ET12567. Unfortunately, how ever a fter s everal a ttempts a t c onjugation be tween *E. coli* ET12567 carrying t he ne w Δ*pmxC*+*D*$_{PKB1}$-IF-Col-9-St^{R} mutant c osmid (streptomycin resistant) and either Δ*pmxC*+*D*$_{PKB1}$ *P. polymyxa* PKB1 or WT PKB1, no ex-conjugants with the desired in-frame 'scar' were generated. Further analysis revealed t hat the high de nsities of *P. polymyxa* PKB1 cells n eeded for intergeneric conjugation gave rise only to false positive streptomycin resistant ex-conjugants.

As a n alternative approach to select the ex-conjugants m aking us e of t he chloramphenicol r esistant m arker, t he Δ*pmxC*+*D*$_{PKB1}$-IF Col -9 m utant c osmid clone was further modified. The Apra^{R}Cm^{R}*oriT* cassette was reintroduced into the

Δ*pmxC*+D_{PKB1}-IF Col-9 mutant cosmid clone by replacing the resident kanamycin resistance marker (*kan*) of SuperCos1 with the Apra^{R}Cm^R*oriT* cassette by another round of PCR targeting gene disruption. In this case the forward primer, 'kan replace fwd' targeting the 5' end of the *kan* gene of SuperCos1, and the reverse primer, 'kan replace rev' targeting the 3' end of *kan* were used to amplify the Apra^{R}Cm^R*oriT* cassette. The linear PCR product was subsequently transformed into *E. coli* BW 25113 carrying the plasmid pIJ790 with the λ RED recombination functions, and Δ*pmxC*+D_{PKB1}-IF Col-9 mutant cosmid. Homologous recombination mediated by the λ RED genes took place between the linear PCR product and Col-9 to generated the mutant Col-9 by replacing *kan* with the Apra^{R}Cm^R*oriT* cassette. Transformants carrying the mutant cosmid were selected on LB agar plate containing appropriate antibiotics. The mutant cosmid Δ*pmxC*+D_{PKB1}-IF-Apra^{R}Cm^R*oriT*, was transformed into *E. coli* ET12567/pUZ8002 to transfer the mutations from *E. coli* to wild type *P. polymyxa* PKB1 by intergeneric conjugation. With this mutant cosmid, chloramphenicol resistance could only arise in *P. polymyxa* PKB1, if the cosmid recombined with the PKB1 chromosome by single crossover in the region of the polymyxin synthetase gene cluster. The resulting ex-conjugants were screened on GB agar plates containing 25 μg/ml polymyxin and 5 μg/ml choramphenicol. To remove the SuperCos1 vector carrying the Apra^{R}Cm^R*oriT* cassette from the genome of PKB1, the ex-conjugants were grown without any antibiotic selection for three generations and then grown in liquid culture to form spores. The sporulated culture was boiled for 15 min and re-isolated on LB agar medium

without selection. Resulting colonies were checked for chloramphenicol sensitivity and sensitive strains were grown in BHI medium without selection at 37^0C for 16 h. Genomic DNA was prepared from 10 of these isolates and checked by PCR for the presence of the in-frame 'scar' (Figure 3.12.9.2). Strains in which the transfer of the mutation was unsuccessful should give a 4.6 kb-band, the same size as the wild type, whereas strains in which the *pmxC+D* genes were replaced by an in-frame scar should give a 1,128-bp band. Two of the chloramphenicol sensitive isolates (IF9 and IF10) were found to contain the 81-bp in-frame 'scar'. Therefore in these new mutants the expression of $pmxE_{PKB1}$ should not be altered by any polar effect due to the deletion of $pmxC+D_{PKB1}$.

3.12.10 Phenotype of the Δ*pmxC+D*$_{PKB1}$-IF mutant

The IF9 and IF10 *P. polymyxa* PKB1 Δ$pmxC+D_{PKB1}$-IF, Δ$pmxC+D_{PKB1}$ mutants and the WT were grown in GSC medium and the culture supernatants bioassayed for polymyxin as described in 3.12.2. The zone of inhibition was measured after 16 h of incubation at 37^0C and was found to be 21 mm for WT and 17 mm for the Δ$pmxC+D_{PKB1}$ mutant and also for the Δ$pmxC+D_{PKB1}$-IF mutant. Therefore, the decrease in antibiotic production seen in strains carrying a deletion of $pmxC_{PKB1}$ and $pmxD_{PKB1}$ does not result from a polar effect on $pmxE_{PKB1}$ expression. Rather, the reduction of antimicrobial activity is due to the deletion of the transporter-like genes Δ*pmxC* and $pmxD_{PKB1}$.

3.13 Concentrating the antibiotics

Since mutation of the transporter-like genes caused only a partial loss of antibiotic production, it was important to ensure that the bioactivity was due to polymyxin only, and that the transporter mutations did not cause loss of one of a mixture of antibiotics. To identify the compound that was responsible for the reduced amount or complete loss of antimicrobial activity of the Δ*pmxC+D*$_{PKB1}$ and Δ*pmxE*$_{PKB1}$ mutants respectively, HPLC was chosen. To simplify the purification of the compound, the culture supernatant was concentrated using a Sep-Pak C_{18} cartridge (Waters Corporation). Five millilitres of culture supernatant containing 20% methanol was passed through a Sep-Pak C_{18} cartridge and all of the antibiotic activity was bound to the packing material of the cartridge. The cartridge was washed with 5 ml of water and eluted with 5 ml portions of 20%, 40%, 60%, 80% and 100% methanol, and 100µl of the water wash and of each eluant was assayed against *E. coli* DH5α. No antimicrobial activity was obtained with the water wash or 20% methanol, very little with 40% and most of the activity eluted with 60% and onward. Therefore, it was decided to use 80% followed by 100% methanol as eluant to extract the antibiotic from the Sep-Pak C_{18} cartridge. Before analysis by HPLC, the eluant was dried completely by flushing with air in an ice bath. The dried sample was dissolved in mobile phase for the HPLC analysis.

3.14 Amount of polymyxin produced by the WT and mutants

Since the phenotype of the WT and mutants were compared with respect to the amount of antibiotic produced as determined by agar diffusion bioassay, the relative amounts of antibiotic produced were determined by comparing the sizes of the zones of inhibition to those produced by different concentrations of a standard polymyxin B sulphate solution. A standard curve was plotted using the data obtained with different concentrations of polymyxin B sulphate solution and the correlation coefficient was calculated (Figure 3.14.1). Good correlation of the standard antibiotic concentration with the sizes of the zone of inhibition was obtained with a R value 0.99, although the slope of the standard curve was very shallow. Based on the standard curve, the antibiotic concentrations in culture supernatants were expressed as mg/ml polymyxin B sulphate equivalent. The amount of the antibiotic produced by the WT was estimated to be 0.5 mg/ml of polymyxin B sulphate equivalent, and Δ*pmxC*+D_{PKB1} mutant produced 0.25 mg/ml of polymyxin B sulphate equivalent.

3.15 Effect of mutations on fusaricidin production

In an earlier study on *P. polymyxa* PKB1 there were no transporter-like genes observed in gene cluster for fusaricidin biosynthesis (Li et al. 2007), and it was speculated that transporter genes from some other gene cluster involved in production of NRP type peptide might be responsible for transport of fusaricidin into the growth medium. Therefore, the three day old culture supernatants of the various *pmx* mutants and WT were assayed for fusaricidin by agar diffusion

bioassay against *Leptosphaeria maculans* as indicator organism. The agar diffusion bioassay showed that the WT and the Δ*pmxE*$_{PKB1}$ mutant produced the same amount of fusaricidin but the Δ*pmxC*$_{PKB1,}$ Δ*pmxD*$_{PKB1}$ and Δ*pmxC+D*$_{PKB1}$ mutants produced reduced amounts of fusaricidin (Figure 3.15.1). A Δ*fusA* mutant of PKB1 (Li 2007; Li et al. 2007) was also grown under the same conditions and the 3 d old culture was checked by agar diffusion bioassay against *L. maculans* as a negative control. Therefore, it was evident that the deletion of ABC transporter-like genes of the polymyxin gene cluster also had an effect on the transport of fusaricidin. It can be speculated from this observation that these ABC transporters might be involved in exporting cationic lipopeptide type molecules from inside the cells. However, in the current study it could not be proven conclusively because the complementation of *pmxC+D*$_{PKB1}$ could not be accomplished. However, there is no indication that the fusaricidin and polymyxin gene clusters are located near one another and so there was no reason to suspect that disruption of *pmxC+D*$_{PKB1}$ would have any polar effects on fusaricidin production.

3.16 HPLC analysis of mutants

The bioassay results suggested that all of the mutants that were developed in this study showed reduced or no antibiotic production compared to WT. The antibacterial compounds that were responsible for the bioactivity against *E. coli* DH5α were subjected to HPLC analysis to try to separate and identify the bioactive components. Different solvent systems and HPLC conditions were tried, and eventually 0.1M Na_2HPO_4 adjusted to pH 3.0 with phosphoric acid:

acetonitrile (77:23) was found to be the best solvent system with a Phenomenex® Bondclone C_{18} column at 25 ^{0}C, a flow rate of 2.0 ml/min, and detection at 212 nm. Within the chromatographic profile two peaks with retention times of 8.0 min and 15.3 min were resolved and found to be present in the WT but absent in the Δ*pmxE*$_{PKB1}$ mutant and reduced in size in the Δ*pmxC+D*$_{PKB1}$ mutants (Figure 3.16.1). The HPLC fractions were collected for the WT and Δ*pmxE*$_{PKB1}$ mutant and 100 µl of each fraction was checked by agar diffusion bioassay against *E. coli* DH5α. Bioactivity was observed for these two peaks from the WT culture supernatant and no activity was observed for the corresponding fractions from the Δ*pmxE*$_{PKB1}$ mutant (Figure 3.16.2). Therefore, it can be inferred that all of the bioactivity was due to these two peaks. It can be seen from the chromatogram that the peak heights of both of the peaks in the Δ*pmxC+D*$_{PKB1}$ mutant are much smaller than those of the corresponding peaks in WT suggesting that the amount of polymyxin produced by Δ*pmxC+D*$_{PKB1}$ mutant is much reduced compared to WT, but the agar diffusion bioassay suggested only a 50% reduction in the amount of antibiotic produced by the Δ*pmxC+D*$_{PKB1}$ mutant compared to the WT. The same results were obtained with the IF mutants, which strengthens the interpretation that there were no polar effects on the downstream gene due to the deletion of *pmxC+D*$_{PKB1}$ (Figure 3.16.3). To identify the compounds represented by the peaks in the WT chromatogram, the profile was compared to the chromatogram of standard polymyxin B sulphate (Sigma-Aldrich) solution, but the major peaks in the standard chromatogram eluted with different retention times compared to the peaks present in the WT, suggesting that the peaks in the

standard chromatogram are not due to the same compounds present in the WT samples (Figure 3.16.4). To confirm this finding, the standard polymyxin B sulphate solution was mixed with the culture supernatant of the WT and a portion of the mixture was analyzed under the same chromatographic conditions and the same results were obtained. This confirms that the bioactive compounds present in the WT culture supernatant are not authentic polymyxin B_1 and B_2.

3.17 Mass spectrometric analysis

HPLC analysis and the agar diffusion bioassay suggested that the peaks isolated in the HPLC analysis have antimicrobial properties, but that the peaks do not represent any known form of polymyxin B. Therefore, to obtain additional information about the compounds and whether they are polymyxins or not, mass spectrometric analysis was carried out. The fractions corresponding to the bioactive peaks collected from the HPLC system were analysed by matrix assisted laser desorption/ionization-time of flight (MALDI-TOF) mass spectrometry, but the results were inconclusive because of the presence of salt in the samples (The MALDI-TOF analysis was done in the Mass spectrometry facility of the Department of Chemistry, U of A). Therefore, the Sep-Pak C_{18} cartridge purified samples of WT, Δ*pmxC*+*D*$_{PKB1}$ and Δ*pmxE*$_{PKB1}$ mutant culture supernatants were analyzed by LC-MS. For LC-MS analysis the HPLC system coupled with single quadrupole mass spectrometer (Micromass ZMD-2, Waters Corporation) was used. Although the conditions used for HPLC analysis as described above were able to resolve the bioactive peaks, the solvent system and the column used in the

HPLC analysis were not suitable for LC-MS analysis because of the high flow rate of solvent and the presence of salt in the mobile phase. Therefore, different solvents containing only volatile components with isocratic and gradient mobile phase systems were tested with an XTerra®MS C_{18} column (150 mm x 2.1 mm, 5 µm) to identify a system that would give good separation of the antibiotic compounds. A mobile phase consisting of 0.1% formic acid (adjusted to a pH 3.0 with ammonium hydroxide) and acetonitrile (77:23) was found to be the best to resolve the antibiotic compounds. In the UV chromatogram at 212 nm, two peaks with retention times of 4.48 and 6.10 min were found and the corresponding peaks were not observed for the Δ*pmxE*$_{PKB1}$ mutant (Figure 3.17.1). In the mass chromatogram the relative abundance of ions was plotted over time and there were two peaks observed for the WT samples with retention times of 4.58 and 6.21 min (representing the two peaks with retention times of 4.48 and 6.10 min respectively found in the UV chromatogram) where the abundance of ions was very high. In the mass chromatogram for the Δ*pmxE*$_{PKB1}$ sample, no peaks were found near the corresponding area and this finding correlates with the HPLC analysis. The mass spectra of these two peaks revealed that the peak with retention time 4.58 min has ions with m/z 595.68 $[M+2H]^{2+}$ and 1189.54 $[M+H]^{+}$ and the peak with retention time 6.21 min has ions with m/z 602.7 $[M+2H]^{2+}$ and 1203.57 $[M+H]^{+}$ (Figure 3.17.2). This observation is consistant with the protonated molecular mass of polymyxin B_2 (Theoretical mass 1188) and B_1 (Theoretical mass 1202), respectively. However, the modular arrangement of the NRPS domains found in *P. polymyxa* PKB1 suggested that the PKB1 polymyxins

have D-DAB in position three in the molecule instead of L-DAB as is found in the polymyxin B standard. Since the molecular masses of the PKB1 polymyxins are the same as the two main forms of polymyxin B found in the standard, but the stereochemistry is different and the PKB1 polymyxins do not co-chromatograph with authentic polymyxin B, these two PKB1 polymyxins can be considered as novel forms of polymyxin B.

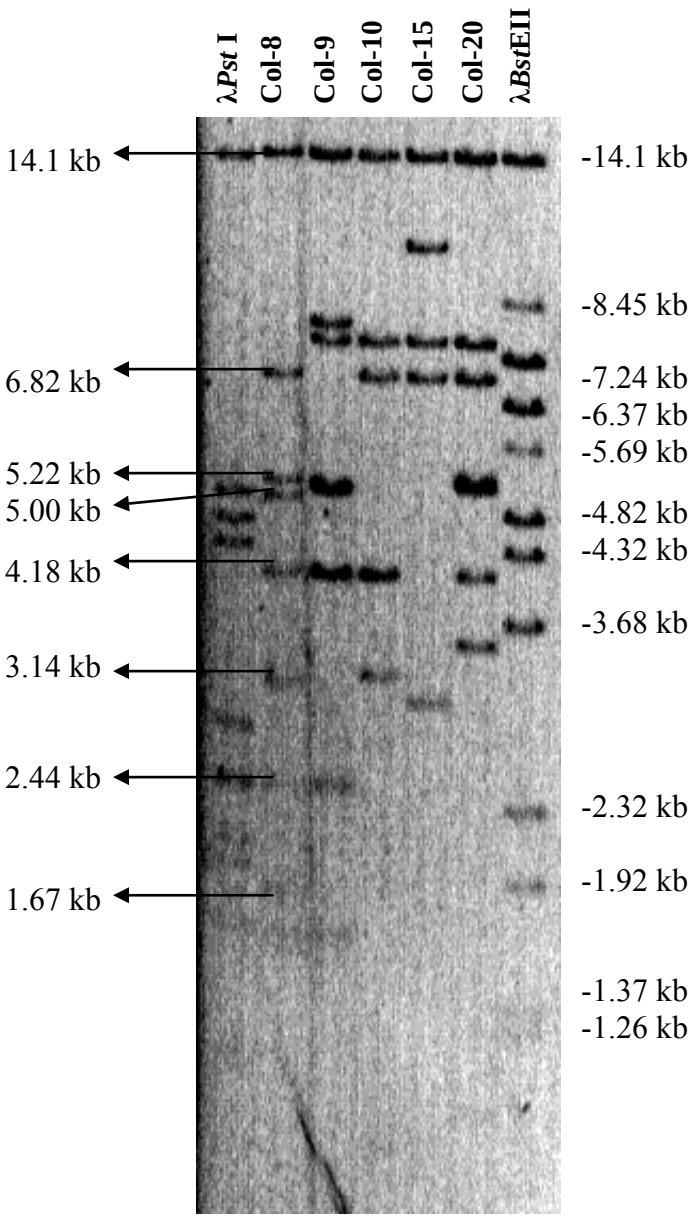


Figure: 3.2.1: *Eco*R1 restriction pattern of five cosmid clones. The cosmid clones were digested with *Eco*R1 at 37 °C for 16 h and were fractionated in a 0.5% agarose gel followed by staining with ethidium bromide. The numbers on the left show the fragment sizes of Col-8, the numbers on the right show the fragment sizes of the λ*Bst*EII digest.

Eco RI (545)
Eco RI (1571)
Eco RI (8255)
Eco RI (10778)
Eco RI (14849)
Eco RI (18074)
Eco RI (28877)
Eco RI (36362)
Eco RI (40196)
5000
10000
15000
20000
25000
30000
35000
40000
7.5 kb
pmxA
pmxB
pmxC
pmxD
pmxE
Eco RI (3164)
Eco RI (4774)
Eco RI (6572)
Eco RI (6733)
Eco RI (11710)
Eco RI (13380)
Eco RI (15822)
Eco RI (19908)
Eco RI (33936)
5.1 kb

Figure: 3.2.2: Comparison of the *pmx* sequence of *P. polymyxa* E681 (Top) with the sequence of Col-8 (Bottom). The in silico *Eco*R1 restriction map for both the *pmx* sequence and Col-8 sequence are shown to compare the fragments by sizes. This was the initial comparison made with *pmx* sequence when the *pmxC* and *pmxD* sequence was not known for *P. polymyxa* E681. The dotted lines indicate the homologous area between the Col-8 sequence and the *pmx* gene cluster. The double lines on the map indicate truncated ORFs.

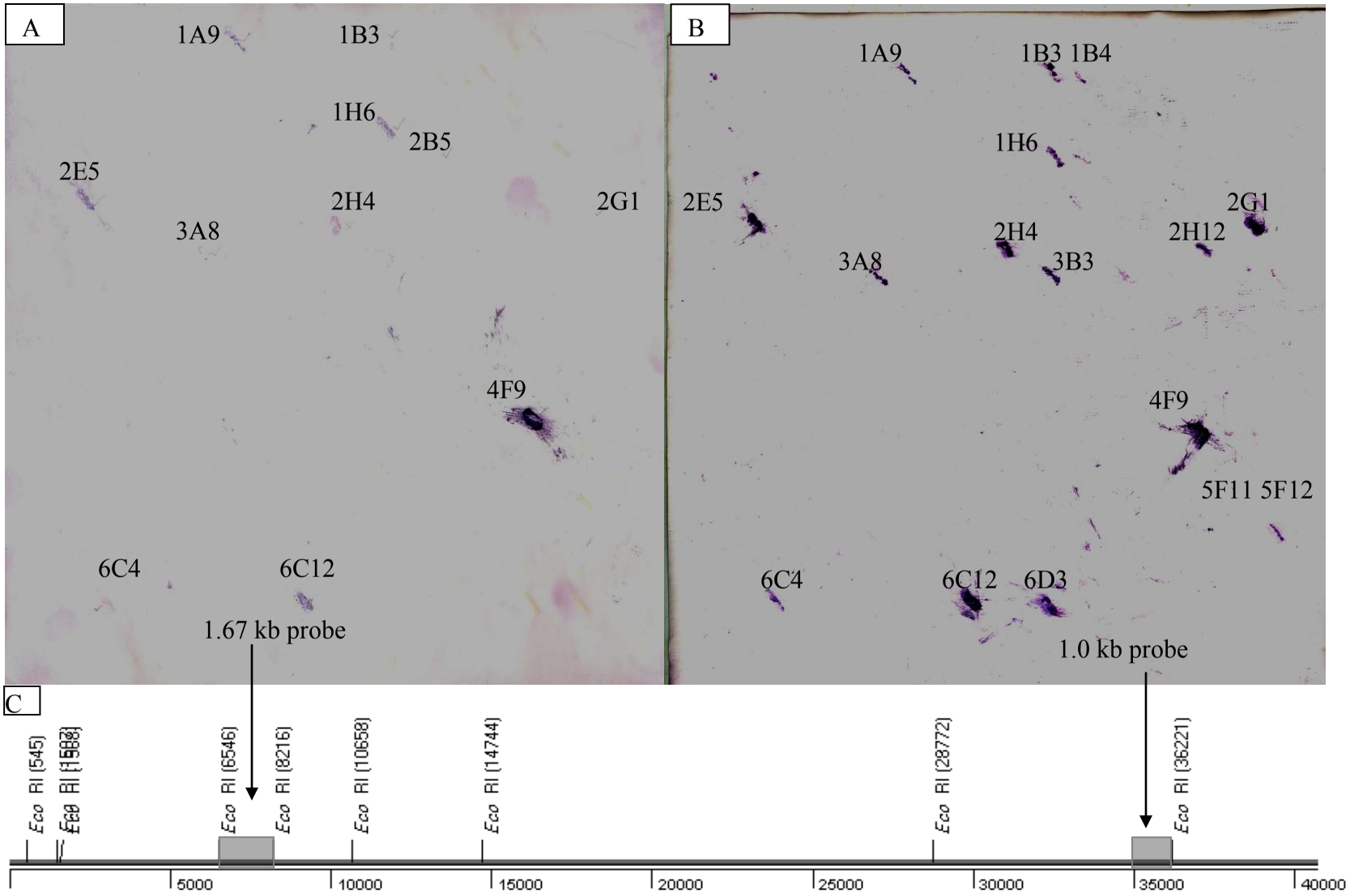

A
1A9
1B3
1H6
2B5
2E5
2H4
2G1
3A8
4F9
6C4
6C12
1.67 kb probe
B
1A9
1B3 1B4
1H6
2E5
2G1
2H4
2H12
3A8
3B3
4F9
5F11 5F12
6C4
6C12
6D3
1.0 kb probe
C
Eco RI (545)
Eco RI (6546)
Eco RI (8216)
Eco RI (10658)
Eco RI (14744)
Eco RI (28772)
Eco RI (36221)
5000
10000
15000
20000
25000
30000
35000
40000

Figure 3.4.1 : Screening of the *P. polymyxa* PKB1 fosmid library by Southern analysis with non-radioactively labelled DNA probes. The first blot (A) containing the cloned DNA was probed with a non-radioactively labeled 1.67 kb *Eco*R1 fragment from Col-8. The second blot (B) was probed with a non-radioactively labeled 1.1 kb PCR product generated with the primers, pSHA09 and pSHA10 and Col-15 as template. The fosmid clones that hybridized with the probes are indicated on the blot. (C)The gray shaded area shows the position of the probes on a pmx_{PKB1} *Eco*R1 restriction map.

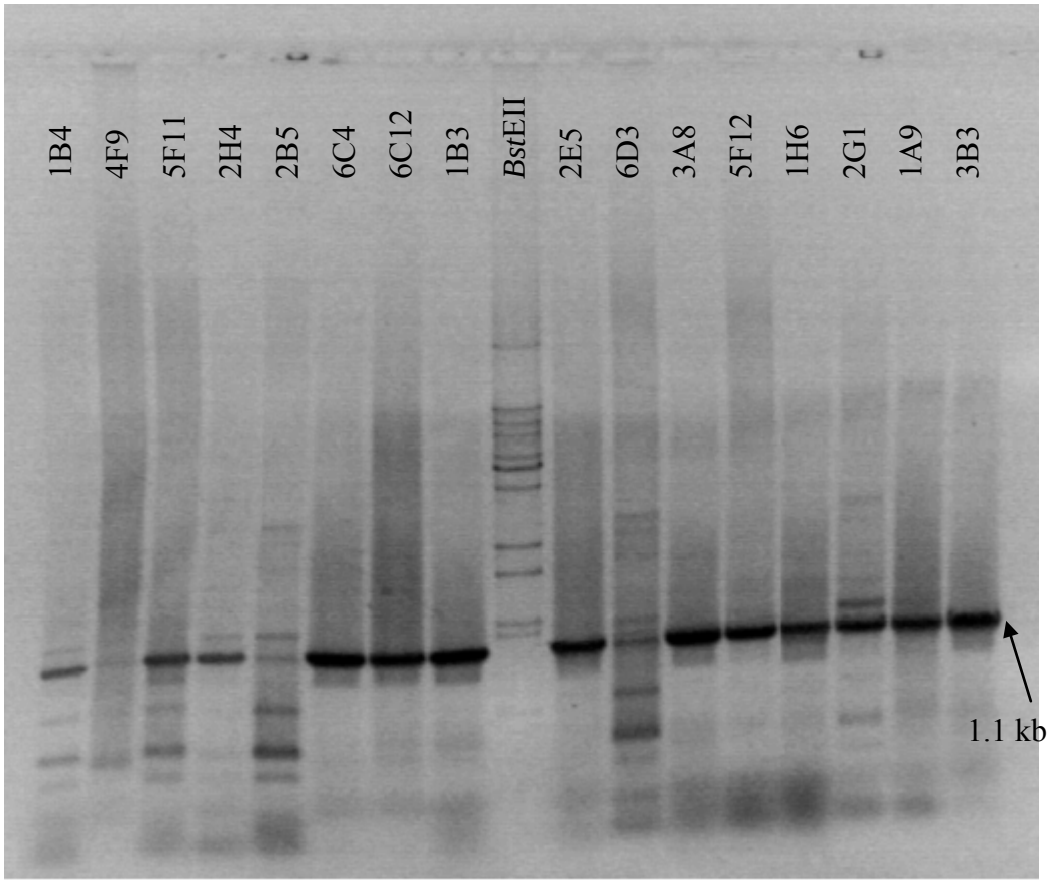


Figure 3.4.2 : PCR confirmation of the fosmid clones that hybridized to probes from each end of the Col-8 insert. The 1.1 kb band is a PCR product generated by the primers pSHA09 and pSHA10 with different fosmid clones as template. The same 1.1 kb PCR product generated by the same primers with Col-15 as template was used as a probe in the Southern analysis. All of the fosmid clones except 4F9, 2B5 and 6D3 produced the 1.1 kb PCR product.

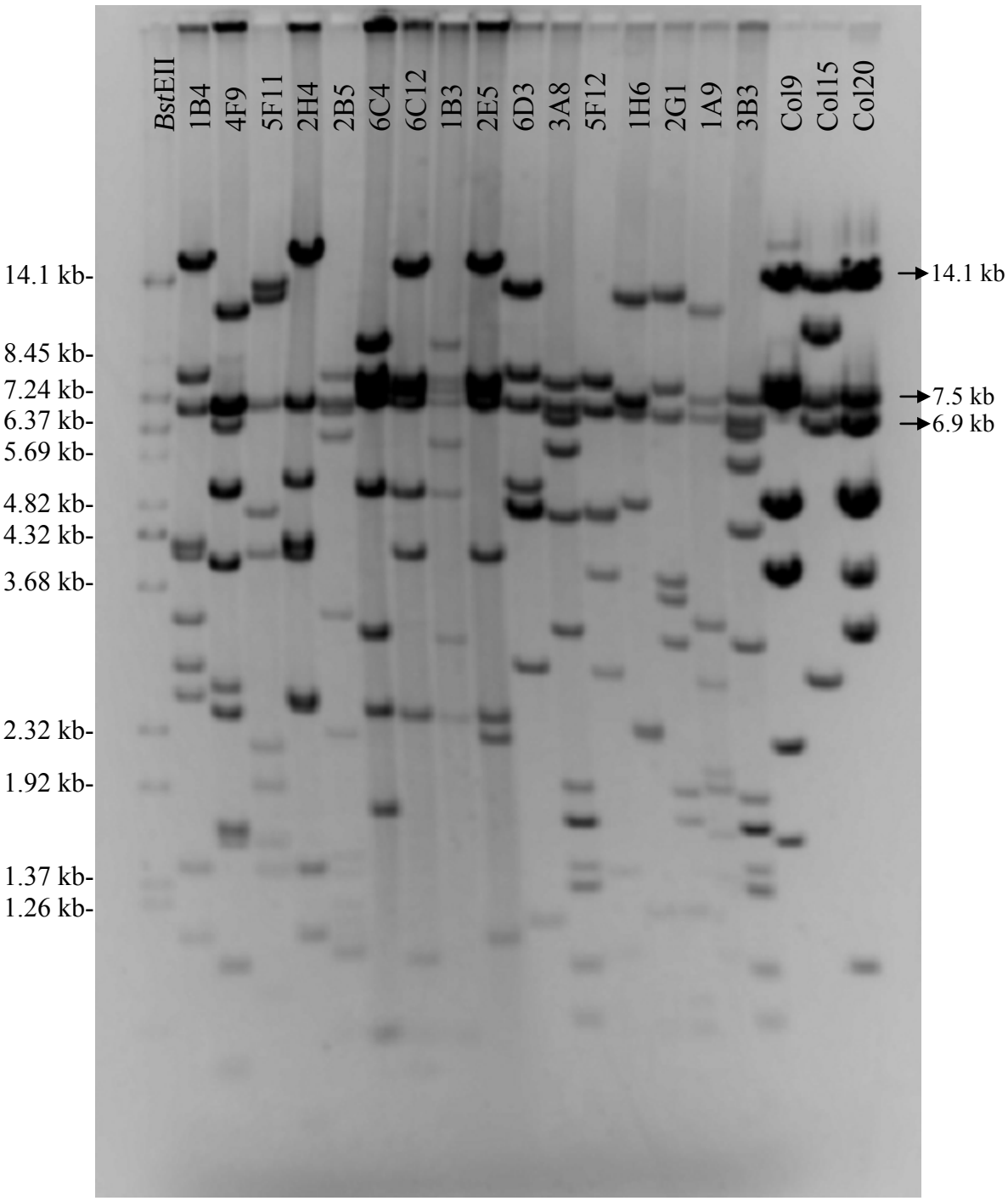

BstEII
1B4
4F9
5F11
2H4
2B5
6C4
6C12
1B3
2E5
6D3
3A8
5F12
1H6
2G1
1A9
3B3
Col9
Col15
Col20
14.1 kb-
8.45 kb-
7.24 kb-
6.37 kb-
5.69 kb-
4.82 kb-
4.32 kb-
3.68 kb-
2.32 kb-
1.92 kb-
1.37 kb-
1.26 kb-
14.1 kb
7.5 kb
6.9 kb

Figure: 3.5.1: The *Eco*R1 restriction digestion pattern of 16 positive fosmid clones and three cosmid clones. One microgram amounts of cosmid and fosmid clones were digested with *Eco*R1 to completion at 37°C for 16 h and fractionated in an 0.8% agarose gel. The numbers on the left show the fragment sizes of λ*Bst*EII digest. The arrows on the right show the 14.1 kb, 7.5 kb and 6.9 kb fragment sizes. The 6.9-kb fragment is considered to be the linear cosmid or fosmid vector.

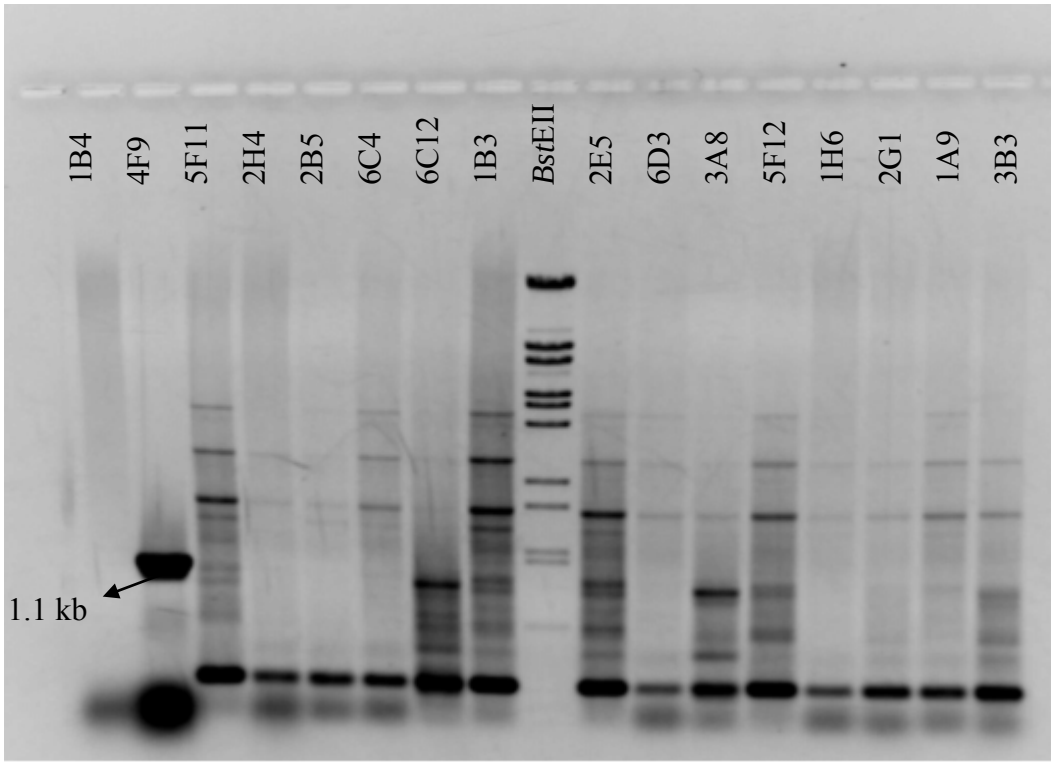


Figure 3.7.1.1 PCR identification of fosmid clones representing the 5' end of the gene cluster encoding NRPS for polymyxin. PCR was carried out with the primers pSHA11 and 12 with the positive fosmid clones (clones that hybridized with the probes in Figure 3.5.2) as template. pSHA11 and pSHA12 were used to sequence the area of the gene cluster 1.1 kb downstream from the beginning of the cluster from the fosmid clone, 4F9. Thus 4F9 was the positive control in this experiment.

Positive control
Speculated fragments
BstEII
1B3
4F9
5F11
2H4
2B5
6C4
6C12
1B3
2E5
6D3
3A8
5F12
1H6
2G1
1A9
3B3
Col-9
Col-15
Col-20

BstEII
1B3
4F9
5F11
2H4
2B5
6C4
6C12
1B3
2E5
6D3
3A8
5F12
1H6
2G1
1A9
3B3
Col-9
Col-15
Col-20

Figure 3.7.2.1: Southern analysis of the *Eco*R1 digested fosmid clones and cosmid clones. Cosmid and fosmid DNA was digested to completion with *Eco*R1 at 37°C for 16 h and fractionated in an 0.8% agarose gel. The *Eco*R1 fragments of the clones were transferred to a nylon membrane followed by hybridization with a non-radioactively labelled 1.1-kb PCR product generated with the primers, pSHA13 and pSHA 14 using 9A as a template (9A is a sub-clone of fosmid clone 3A8). (A) The ethidium bromide stained agarose gel (also shown previously in Figure 3.5.1). (B) Corresponding nylon blot hybridized with the probe. The 6.3-kb band shown in a circle in 3A 8 was the positive control and the 5.1 kb bands shown in circles in 2E5 and 6C12 were proposed to be true representatives of the genomic DNA for this area.

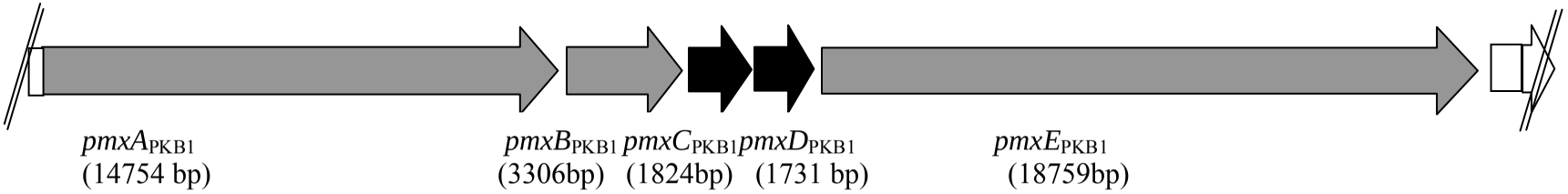


Figure 3.8.1: Arrangement of ORFs in the gene cluster encoding the NRPS for polymyxin in *P. polymyxa* PKB1. The grey arrows represent the NRPS genes whereas the black arrows represent the ABC transporter-like genes and the direction of the arrows indicates the direction of transcription. The numbers refer to the length of each ORF and the figure is drawn to scale. The white bars indicate adjacent areas of the gene cluster not associated with polymyxin biosynthesis and the white arrow on the right indicates the beginning of an ORF encoding a protein homologous to a hypothetical protein of *Geobacillus* spp.

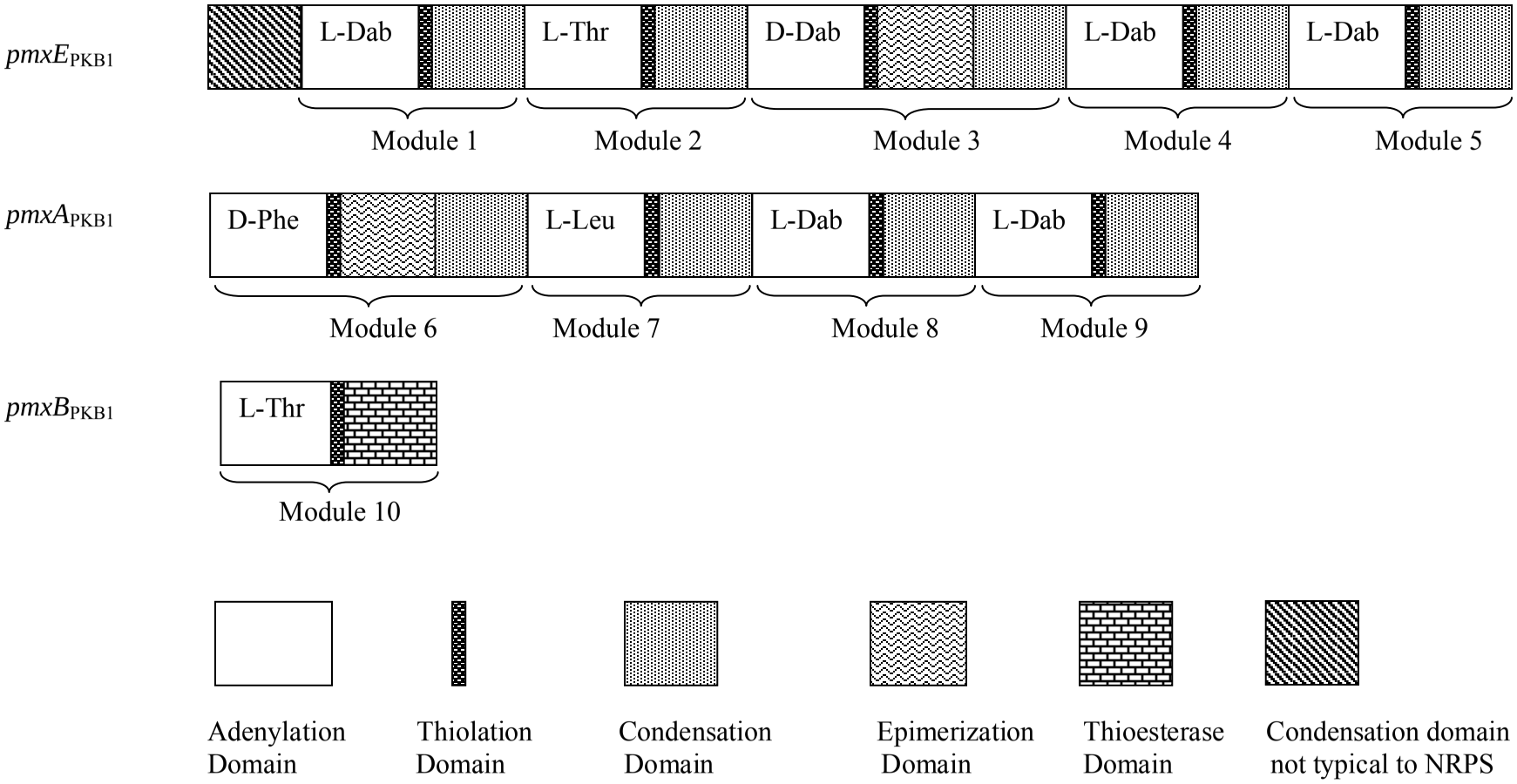

pmxE PKB1
L-Dab
L-Thr
D-Dab
L-Dab
L-Dab
Module 1
Module 2
Module 3
Module 4
Module 5
pmxA PKB1
D-Phe
L-Leu
L-Dab
L-Dab
Module 6
Module 7
Module 8
Module 9
pmxB PKB1
L-Thr
Module 10
Adenylation Domain
Thiolation Domain
Condensation Domain
Epimerization Domain
Thioesterase Domain
Condensation domain not typical to NRPS

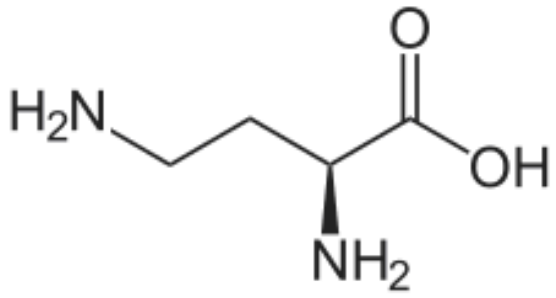
Figure 3.9. 1: Module and domain arrangement of the NRPS for polymyxin biosynthesis. The distribution of domains in different ORFs is shown by blocks with different patterns. Both *pmxA*$_{PKB1}$ and *pmxE*$_{PKB1}$ contain one E domain which is responsible for the stereochemical conversion of substrate L-amino acid. The unusual feature of the gene cluster is that polymyxin biosynthesis begins from the last ORF (*pmxE*$_{PKB1}$) which incorporates the first five amino acids of polymyxin molecule.

	C1 LW SxAQxR xL MY	C2 RHExLRTxF	C3 W MHHxISDG S V	C4 F YxD AVW Y	C5 I QC GxFVNT xR V LA	C6 H Y QD PFE N V	C7 RDxSRNPL
C domain for DAB_4 ($pmxE_{PKB1}$)	SSAQKRLYIL	RHETLRTGF	THHIVSDGVS	YKDYAVW	IGMFVNTLAIR	HQNYPFE	RDLSRNPL
C domain for DAB_8 ($pmxA_{PKB1}$)	SFAQKRLFIL	RHEMLRTGF	MHHIISDGVS	YKDYTVW	IGIFLNTLAIR	HQNYPFE	RDLSRNPL
C domain for DAB_1 ($pmxE_{PKB1}$)	SSEQKRLYVL	RHETLRTGI	MHHIVSDGIS	YKDYAVW	VGMFINTLAIR	HQNYLFE	RATGRNPL
C domain for DAB_9 ($pmxA_{PKB1}$)	SSAQKRLFIQ	RHESLRTGF	MHHMVSDGVS	YKDYAVW	IGMFVNTLAIR	NQDYPFE	REPGRFPL
C domain for DAB_5 ($pmxE_{PKB1}$)	SSVQKRLYIQ	RHEILRTGF	IHHIVTDGMS	YKEFAAW	IGMFVNTLAIR	HQDYPFE	RDQSRNPV
C domain for DAB_3 ($pmxE_{PKB1}$)	TPLQKGMLFH	RNEALRTNF	HHHILMDGWC	PVTSYSR	IGLFINTIPVR	YETFPLF	RDLINHIM
C domain for Thr_2 ($pmxE_{PKB1}$)	SSAQKRLYVL	RHAMLRTGF	IHHIGSDGLS	YKDYAVW	IGMFVNTLALR	HQDYPFE	RDLSRNPL
C domain for Phe_6 ($pmxA_{PKB1}$)	SSAQKRLFIL	RHETLRTGF	MHHIVSDGVS	YKDYAVW	IGMFVNTLALR	HQNYPFE	RDLSRNPL
C domain for leu_7 ($pmxA_{PKB1}$)	TPMQKGMWFH	RHLVLRANF	FQHILMDGWC	GSDYGAY	IGLFINTVPVR	YDYYPLY	QNLINHII

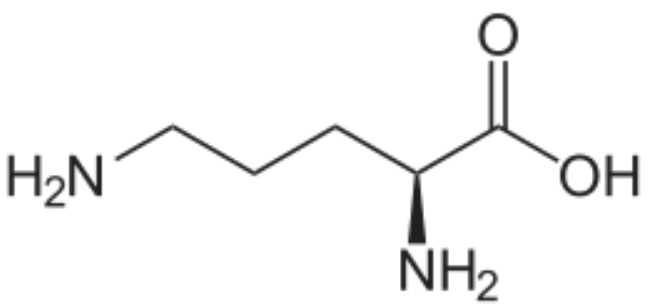
Figure 3.9.1 .1: C omparison of t he amino a cid s equences of c ore m otifs of t he c ondensation domains pr esent in t he N RPS f or polymyxin biosynthesis. Conserved core motif sequences found in the C domains of the NRPS for polymyxins are shown shaded in yellow. C domains that are involved in the formation of peptide bonds between L and D amino acids showed greater similarity to each other and are shown shaded in cyan (Marahiel et al. 1997).

Adenylation domain	*orf*	Amino acid residues in the substrate binding pocket										**Predicted substrate**
		235	**236**	**239**	**278**	**299**	**301**	**322**	**330**	**331**	**517**	
A_1	*pmxE*$_{PKB1}$	**D**	**V**	**G**	**E**	**I**	**S**	**S**	**I**	**D**	**K**	Orn/Dab
A_2	*pmxE*$_{PKB1}$	**D**	**F**	**W**	**N**	**I**	**G**	**M**	**V**	**H**	**K**	Thr
A_3	*pmxE*$_{PKB1}$	**D**	**V**	**G**	**E**	**I**	**S**	**S**	**I**	**D**	**K**	Orn/Dab
A_4	*pmxE*$_{PKB1}$	**D**	**V**	**G**	**E**	**I**	**S**	**A**	**I**	**D**	**K**	Orn/Dab
A_5	*pmxE*$_{PKB1}$	**D**	**V**	**G**	**E**	**I**	**S**	**A**	**I**	**D**	**K**	Orn/Dab
A_6	*pmxA*$_{PKB1}$	**D**	**A**	**W**	**T**	**I**	**A**	**A**	**I**	**A**	**K**	Phe
A_7	*pmxA*$_{PKB1}$	**D**	**G**	**F**	**L**	**L**	**G**	**L**	**V**	**Y**	**K**	Ile/Leu
A_8	*pmxA*$_{PKB1}$	**D**	**V**	**G**	**E**	**I**	**S**	**A**	**I**	**D**	**K**	Orn/Dab
A_9	*pmxA*$_{PKB1}$	**D**	**V**	**G**	**E**	**I**	**S**	**A**	**I**	**D**	**K**	Orn/Dab
A_{10}	*pmxB*$_{PKB1}$	**D**	**F**	**W**	**N**	**I**	**G**	**M**	**V**	**H**	**K**	Thr

Figure 3.9.2.1: Predicted substrate specificity of the adenylation (A) domains of NRPS for polymyxin biosynthesis identified by the online database http://www.nii.res.in/nrps-pks.html based on the selectivity conferring code of NRPS. Ornithine and DAB share a very similar structure and due to that, ornithine is identified as substrate by the online algorithm for six A domains which actually recognize DAB.



Diaminobutyrate



Ornithine

Figure 3.9.2.2: Structure of diaminobutyric (Dab) acid and ornithine.

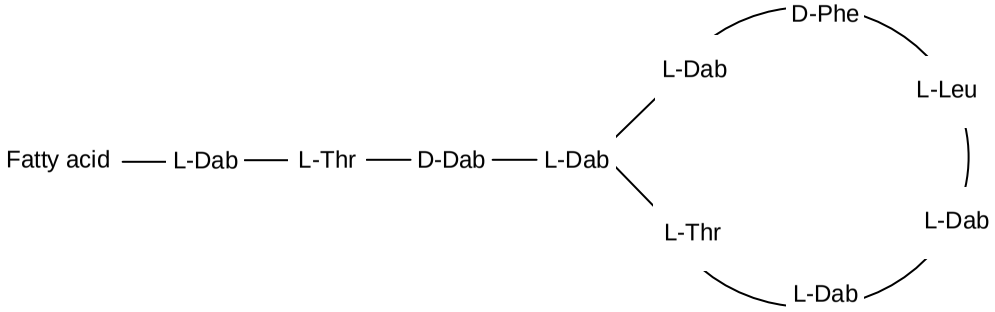


Figure 3.9. 2.3: Structure of polymyxin based on the amino acids predicted by the online algorithm for NRPS http://www.nii.res.in/nrps-pks.html. Polymyxins usually contain a heptapeptide ring attached to the fatty acid by a tripeptide side chain. The ten amino acid building blocks of polymyxin were incorporated by the NRPS encoded by three genes of the gene cluster in *P. polymyxa* PKB1: $pmxA_{PKB1,}$ $pmxB_{PKB1}$ and $pmxE_{PKB1.}$

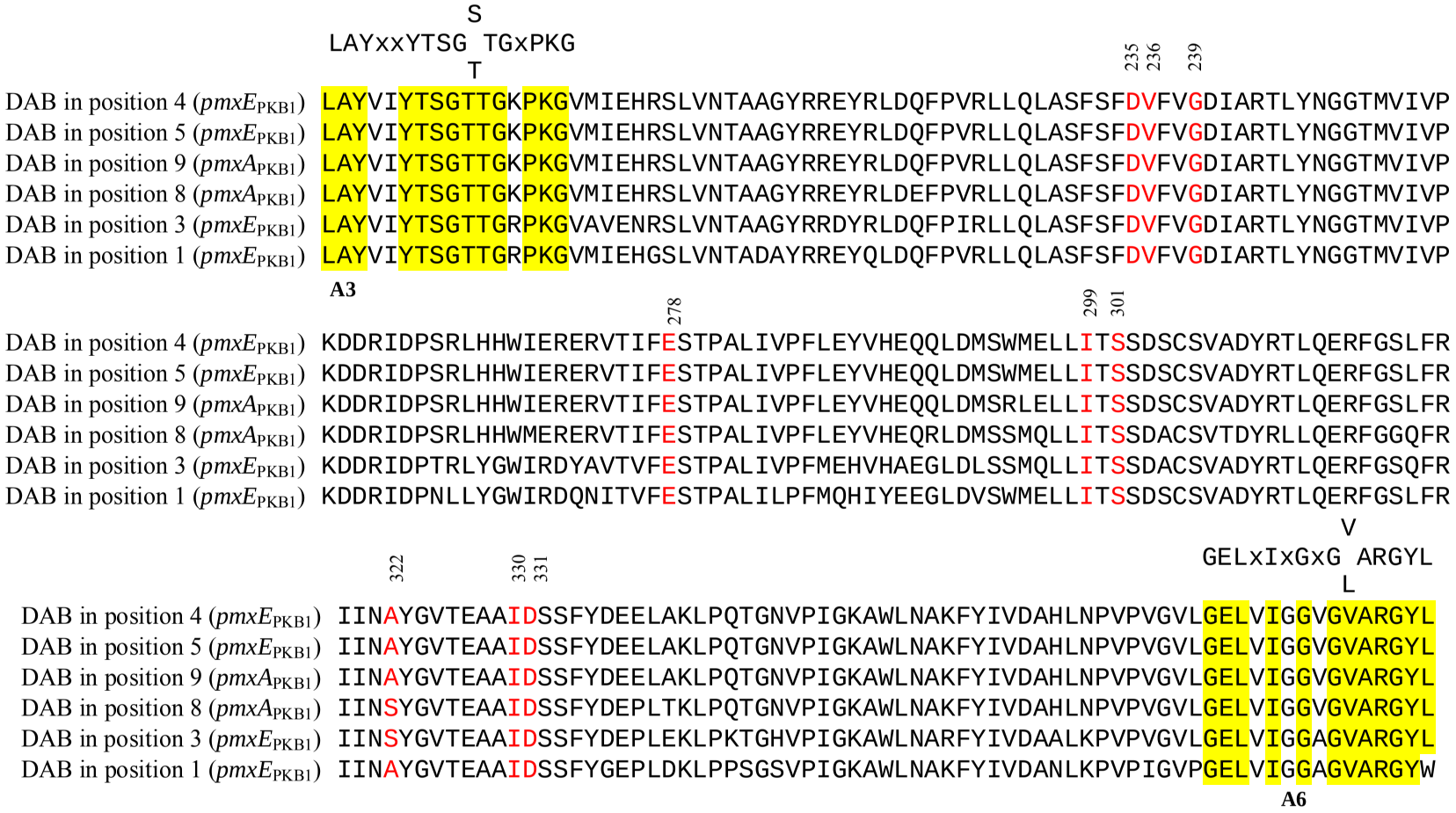

S
LAYxxYTSG TGxPKG
T
235
236
239
DAB in position 4 (pmxE PKB1) LAYVIYTSGTTGKPKGVMIEHRSLVNTAAGYRREYRLDQFPVRLLQLASFSFDVFVGDIARTLYNGGTMVIVP
DAB in position 5 (pmxE PKB1) LAYVIYTSGTTGKPKGVMIEHRSLVNTAAGYRREYRLDQFPVRLLQLASFSFDVFVGDIARTLYNGGTMVIVP
DAB in position 9 (pmxA PKB1) LAYVIYTSGTTGKPKGVMIEHRSLVNTAAGYRREYRLDQFPVRLLQLASFSFDVFVGDIARTLYNGGTMVIVP
DAB in position 8 (pmxA PKB1) LAYVIYTSGTTGKPKGVMIEHRSLVNTAAGYRREYRLDEFPVRLLQLASFSFDVFVGDIARTLYNGGTMVIVP
DAB in position 3 (pmxE PKB1) LAYVIYTSGTTGRPKGVAVENRSLVNTAAGYRRDYRLDQFPIRLLQLASFSFDVFVGDIARTLYNGGTMVIVP
DAB in position 1 (pmxE PKB1) LAYVIYTSGTTGRPKGVMIEHGSLVNTADAYRREYQLDQFPVRLLQLASFSFDVFVGDIARTLYNGGTMVIVP
A3
278
299
301
DAB in position 4 (pmxE PKB1) KDDRIDPSRLHHWIERERVTIFESTPALIVPFLEYVHEQQLDMSWMELLITSSDSCSVADYRTLQERFGSLFR
DAB in position 5 (pmxE PKB1) KDDRIDPSRLHHWIERERVTIFESTPALIVPFLEYVHEQQLDMSWMELLITSSDSCSVADYRTLQERFGSLFR
DAB in position 9 (pmxA PKB1) KDDRIDPSRLHHWIERERVTIFESTPALIVPFLEYVHEQQLDMSRLELLITSSDSCSVADYRTLQERFGSLFR
DAB in position 8 (pmxA PKB1) KDDRIDPSRLHHWMERERVTIFESTPALIVPFLEYVHEQRLDMSSMQLLITSSDACSVTDYRLLQERFGGQFR
DAB in position 3 (pmxE PKB1) KDDRIDPTRLYGWIRDYAVTVFESTPALIVPFMEHVHAEGLDLSSMQLLITSSDACSVADYRTLQERFGSQFR
DAB in position 1 (pmxE PKB1) KDDRIDPNLLYGWIRDQNITVFESTPALILPFMQHIYEEGLDVSWMELLITSSDSCSVADYRTLQERFGSLFR
V
GELxIxGxG ARGYL
L
322
330
331
DAB in position 4 (pmxE PKB1) IINAYGVTEAAIDSSFYDEELAKLPQTGNVPIGKAWLNAKFYIVDAHLNPVPVGVLGELVIGGVGVARGYL
DAB in position 5 (pmxE PKB1) IINAYGVTEAAIDSSFYDEELAKLPQTGNVPIGKAWLNAKFYIVDAHLNPVPVGVLGELVIGGVGVARGYL
DAB in position 9 (pmxA PKB1) IINAYGVTEAAIDSSFYDEELAKLPQTGNVPIGKAWLNAKFYIVDAHLNPVPVGVLGELVIGGVGVARGYL
DAB in position 8 (pmxA PKB1) IINSYGVTEAAIDSSFYDEPLTKLPQTGNVPIGKAWLNAKFYIVDAHLNPVPVGVLGELVIGGVGVARGYL
DAB in position 3 (pmxE PKB1) IINSYGVTEAAIDSSFYDEPLEKLPKTGHVPIGKAWLNARFYIVDAALKPVPVGVLGELVIGGAGVARGYL
DAB in position 1 (pmxE PKB1) IINAYGVTEAAIDSSFYGEPLDKLPPSGSVPIGKAWLNAKFYIVDANLKPVPIGVPGELVIGGAGVARGYW
A6

Figure 3.9.2.4: Alignment of the amino acid sequences of core motifs A3 to A6 of the A domains for the six DAB residues of polymyxin. The amino acid positions of the substrate binding pocket that confer the substrate specificity are shown in red. Nine of the substrate specificity conferring amino acids are present in the A3 to A6 region and the tenth amino acid is at position 517, which is always a lysine (K). The core motifs A3 and A6 are shown shaded in yellow (Marahiel et al. 1997).

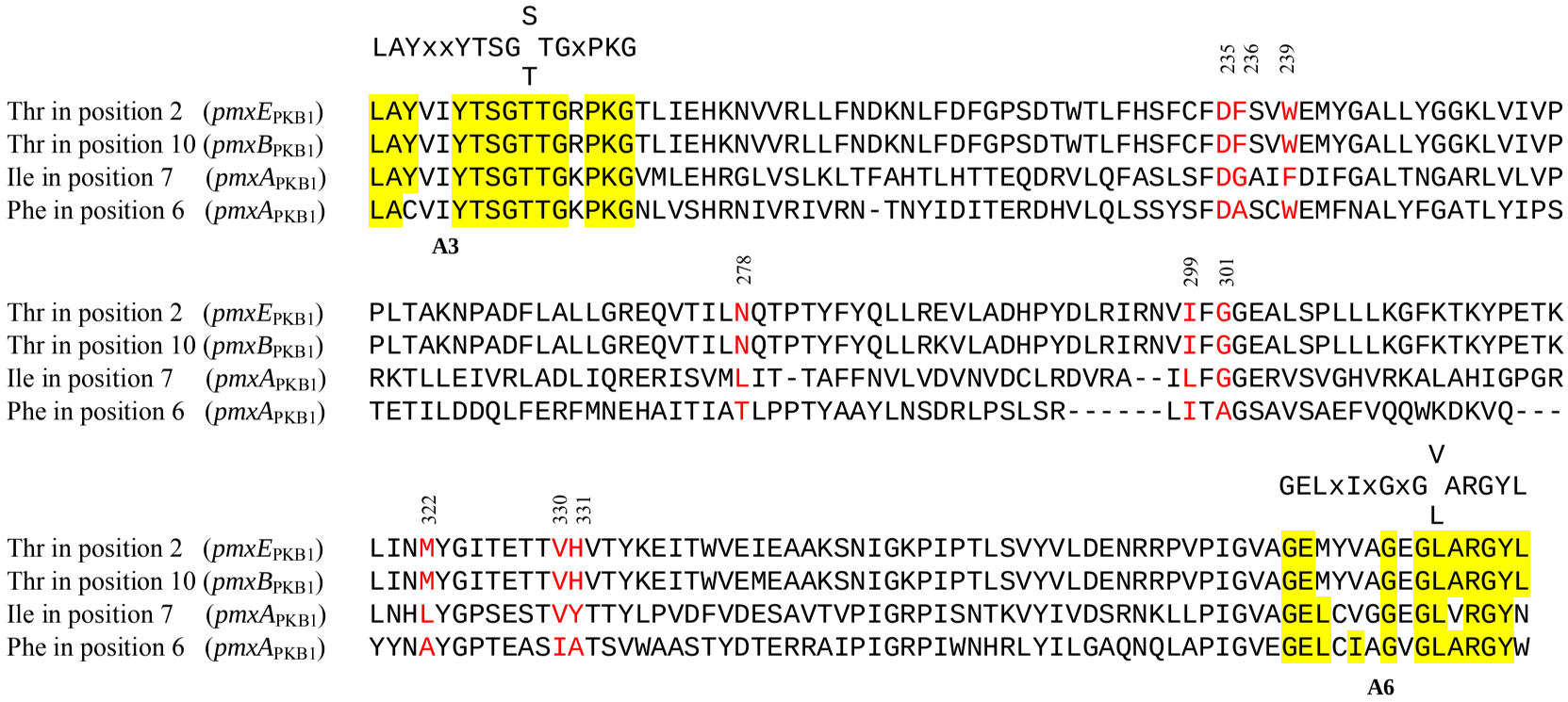


Figure 3.9.2.5: Alignment of the amino acid sequences of motifs A3 to A6 of A domains for residues other than DAB in polymyxin. The amino acid positions of the substrate binding pocket that confer the substrate specificity are shown in red. Nine of the substrate specificity conferring amino acids are present in the A3 to A6 region, and the tenth amino acid position is 517, which is always a lysine (K). The core motifs A3 and A6 are shown shaded in yellow (Marahiel et al. 1997).

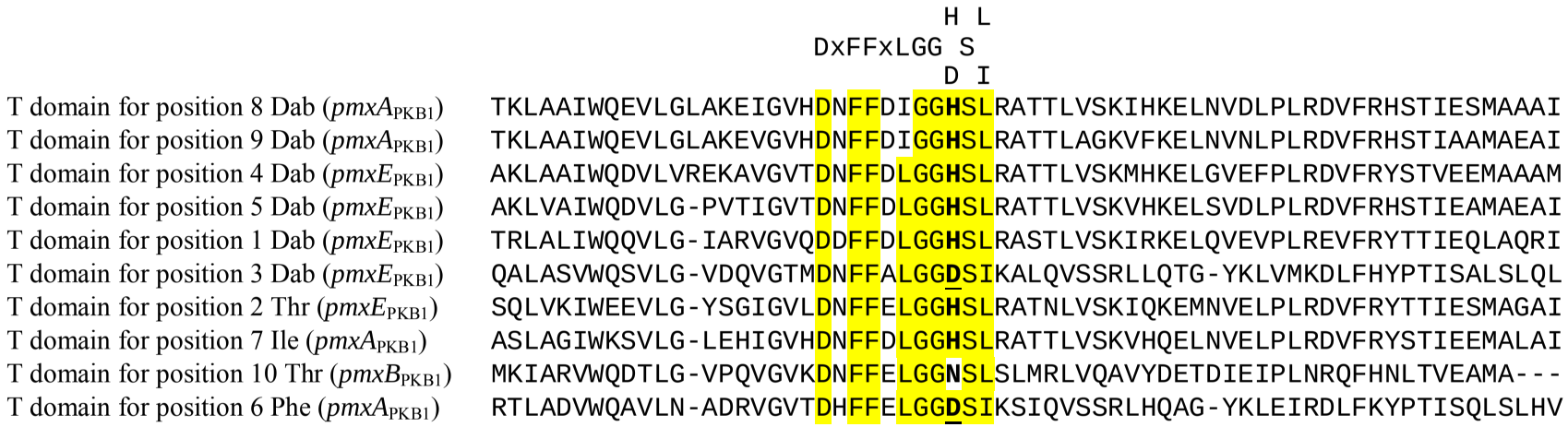


Figure 3.9.3.1: Alignment of the amino acid sequences of thiolation domains of all ten amino acid of polymyxin. Conserved core motifs are highlighted in yellow. The amino acid before the serine residue (S) in the conserved motif is shown in bold. An aspartate (D) residue before the serine residue (S) is required for the interaction with epimerization domain (E) to change the stereochemistry of substrate amino acid. Thus the third amino acid of polymyxin should be a D-Dab and the sixth amino acid should be a D-Phe. All of the T domains for L-amino acids have histidine (H) before the serine residue (S) except in the last T domain for Thr which has an asparagine residue (Marahiel et al. 1997).

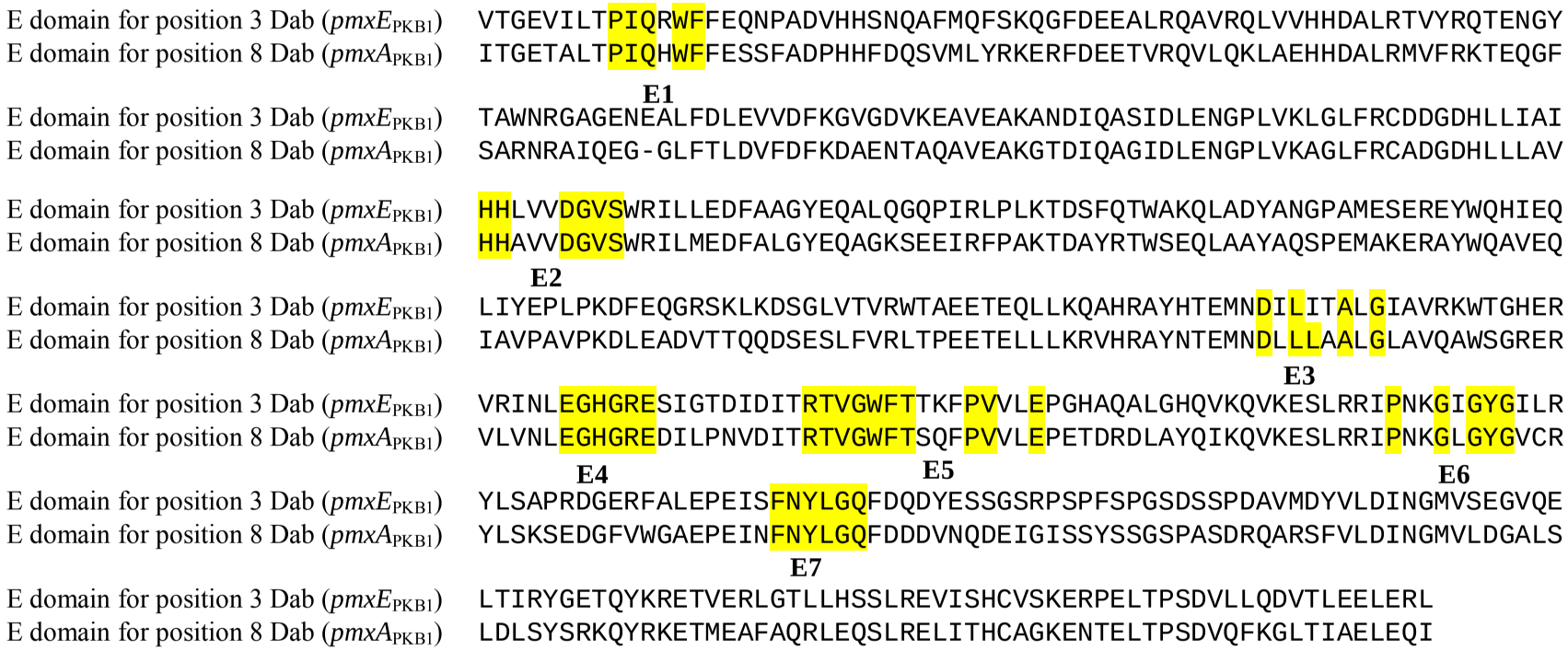


Figure 3.9.4.1: Alignment of the amino acid sequences of epimerization domains present in the NRPS for polymyxin biosynthesis. Conserved core motifs are highlighted in yellow (Marahiel et al. 1997). Both of the epimerization domains share the same motif sequence. It can be assumed that both of them are active and responsible for incorporating the D isomer of Dab and Phe in position three and six of the polymyxin respectively.

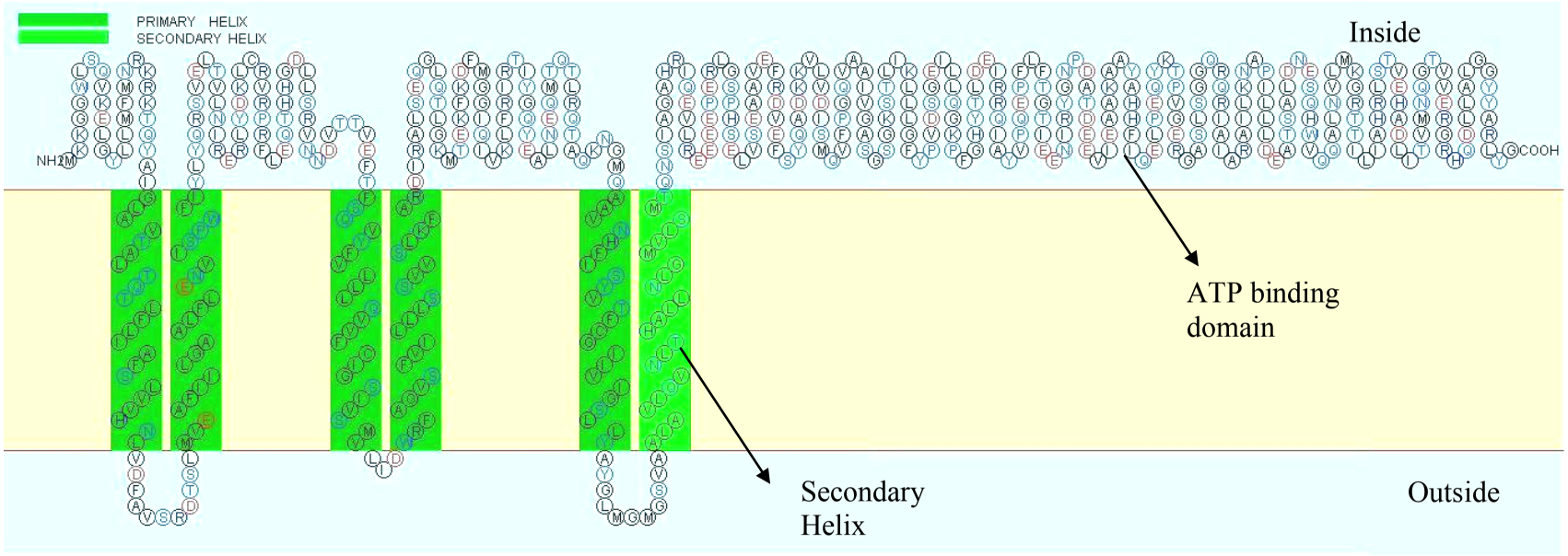


Figure 3.10.1: Secondary structure of PmxD as determined by SOSUI. Five primary membrane helices and one secondary membrane helix were determined and the ATP binding domain was on the C-terminal region of the protein (Hirokawa et al. 1998; Mitaku and Hirokawa 1999; Mitaku et al. 2002).

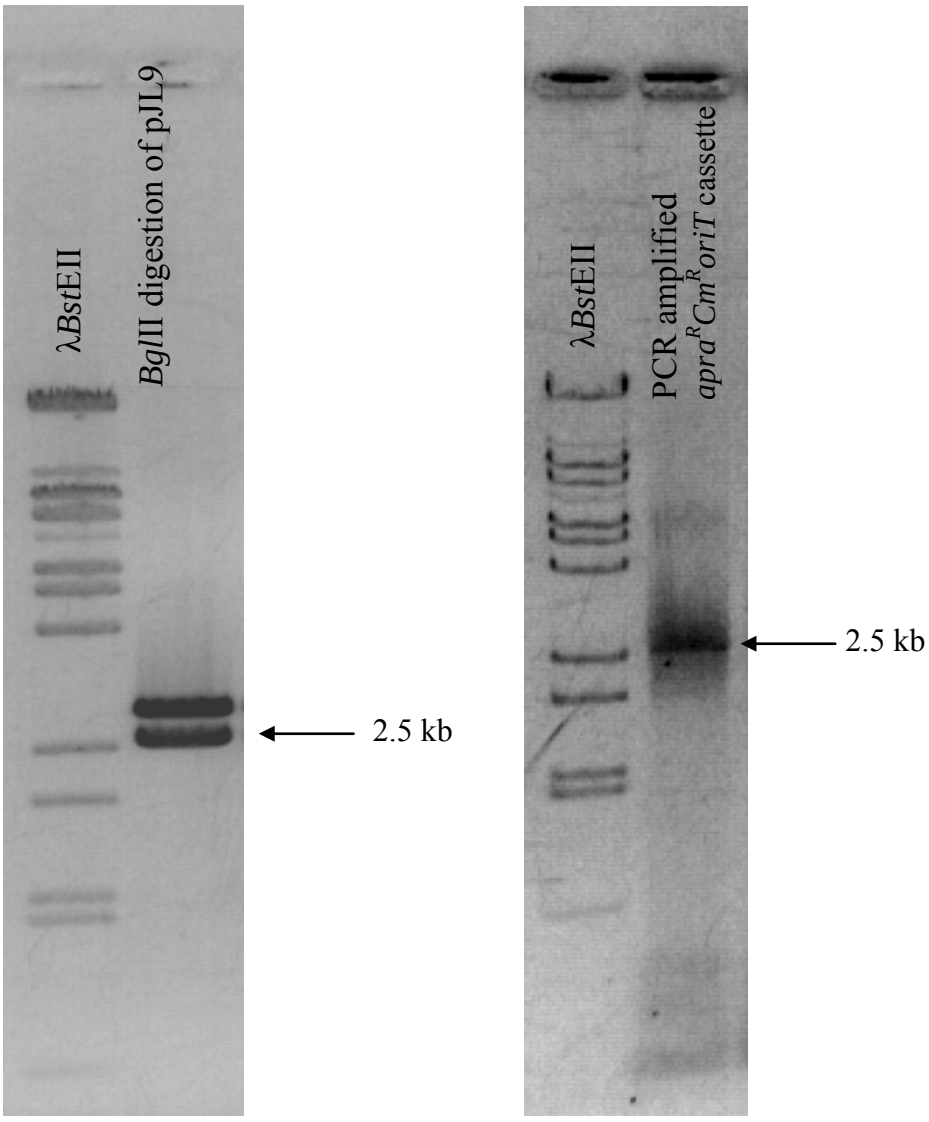


Figure 3.11. 1.1: P reparation and P CR a mplification of the *apra*^{R}*Cm*^{R}*oriT* cassette f or PCR-targeted mutational analysis. (A) Isolation of the *apra*^{R}*Cm*^{R}*oriT* cassette as a *Bgl*II fragment from pJL9. The 2.5 kb (lower band) is the *apra*^{R}*Cm*^{R}*oriT* cassette, which was gel pur ified a nd us ed as a t emplate f or P CR a mplification. (B) T he P CR a mplified *apra*^{R}*Cm*^{R}*oriT* cassette generated by using the Redirect primers pSHA15 and pSHA16 to disrupt the *pmxD*$_{PKB1}$.

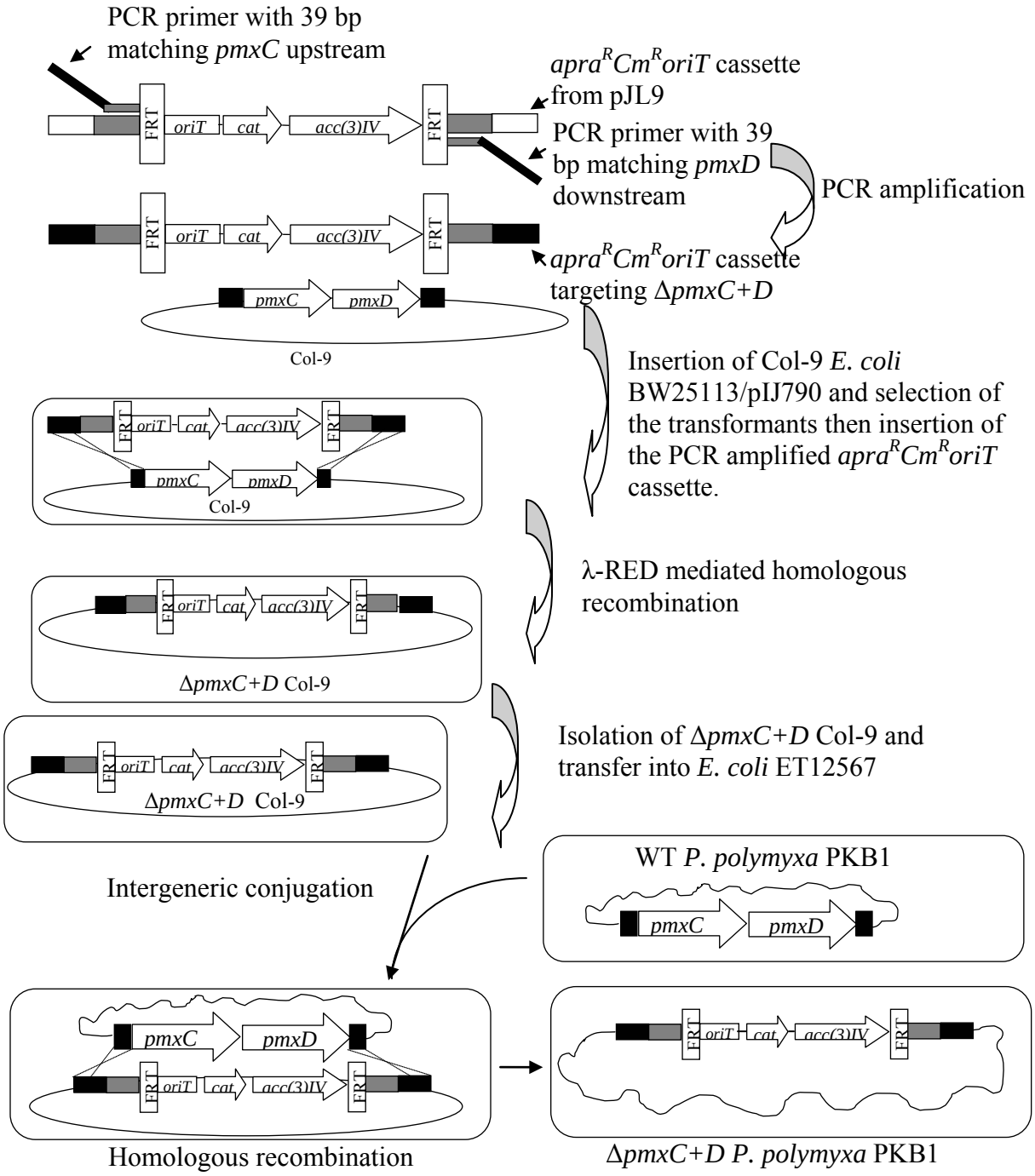


Figure 3.12.1.1: Schematic representation of PCR targeted mutagenesis to generate *P. polymyxa* Δ*pmxC+D*$_{PKB1}$ mutants.

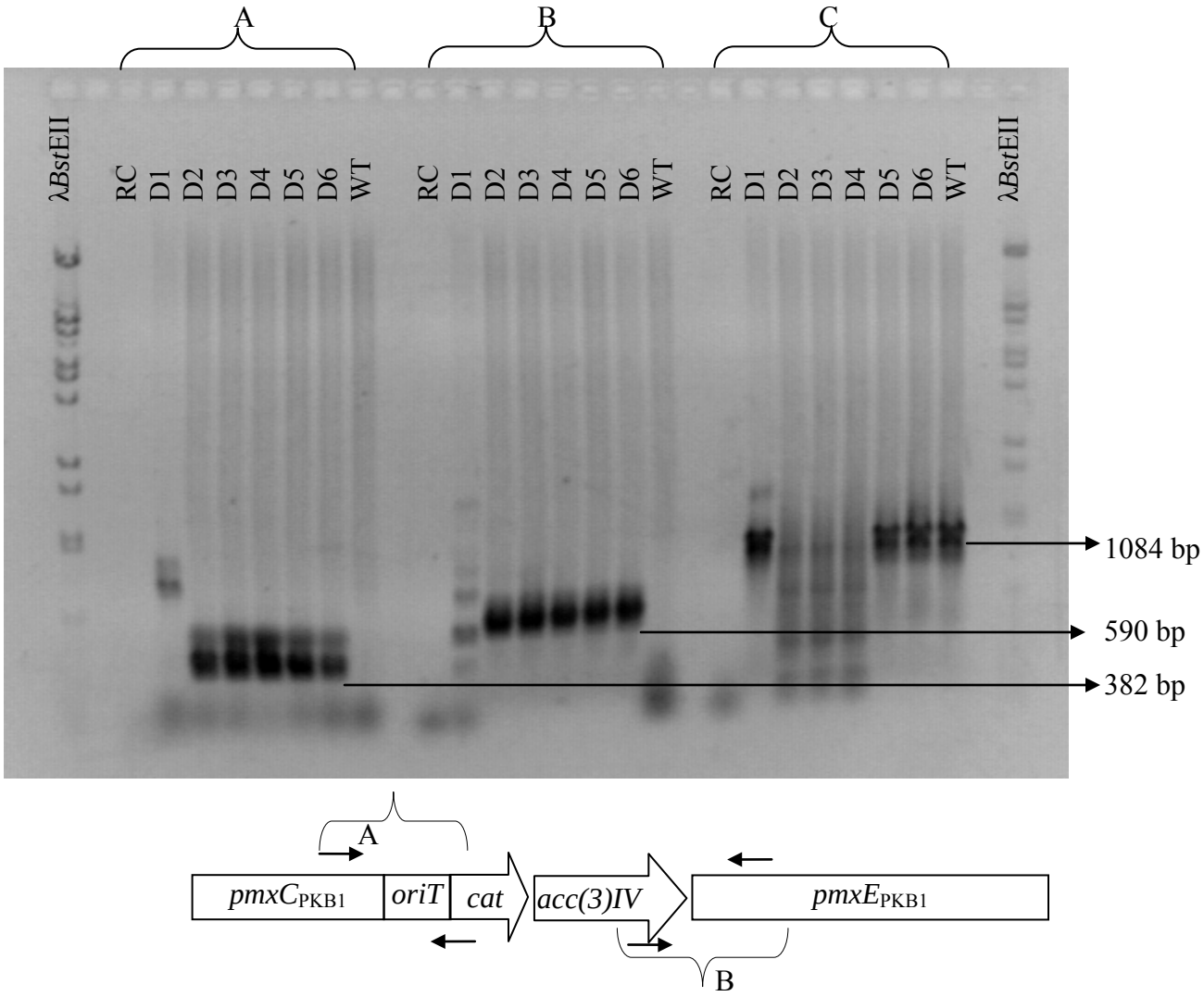


Figure 3.12.1.2: Confirmation of *P. polymyxa ΔpmxD*$_{PKB1}$ mutants by PCR. PCR was carried out with 3 sets of primers with the genomic DNA of six ex-conjugant (D1-D6) candidate *P. polymyxa ΔpmxD*$_{PKB1}$ mutants and WT *P. polymyxa*. (A) Two sets of primers (pSHA27 and RED-SEQ-UP) were used to determine the location of the antibiotic cassette at the 5' end. (B) Two primers pSHA28 and RED-SEQ-DWN were used to determine the location of the antibiotic cassette at the 3' end. (C) Two primers pSHA29 and pSHA30 were used to detect the presence of *pmxD*$_{PKB1}$ (WT should produce a PCR product of 1084 bp).

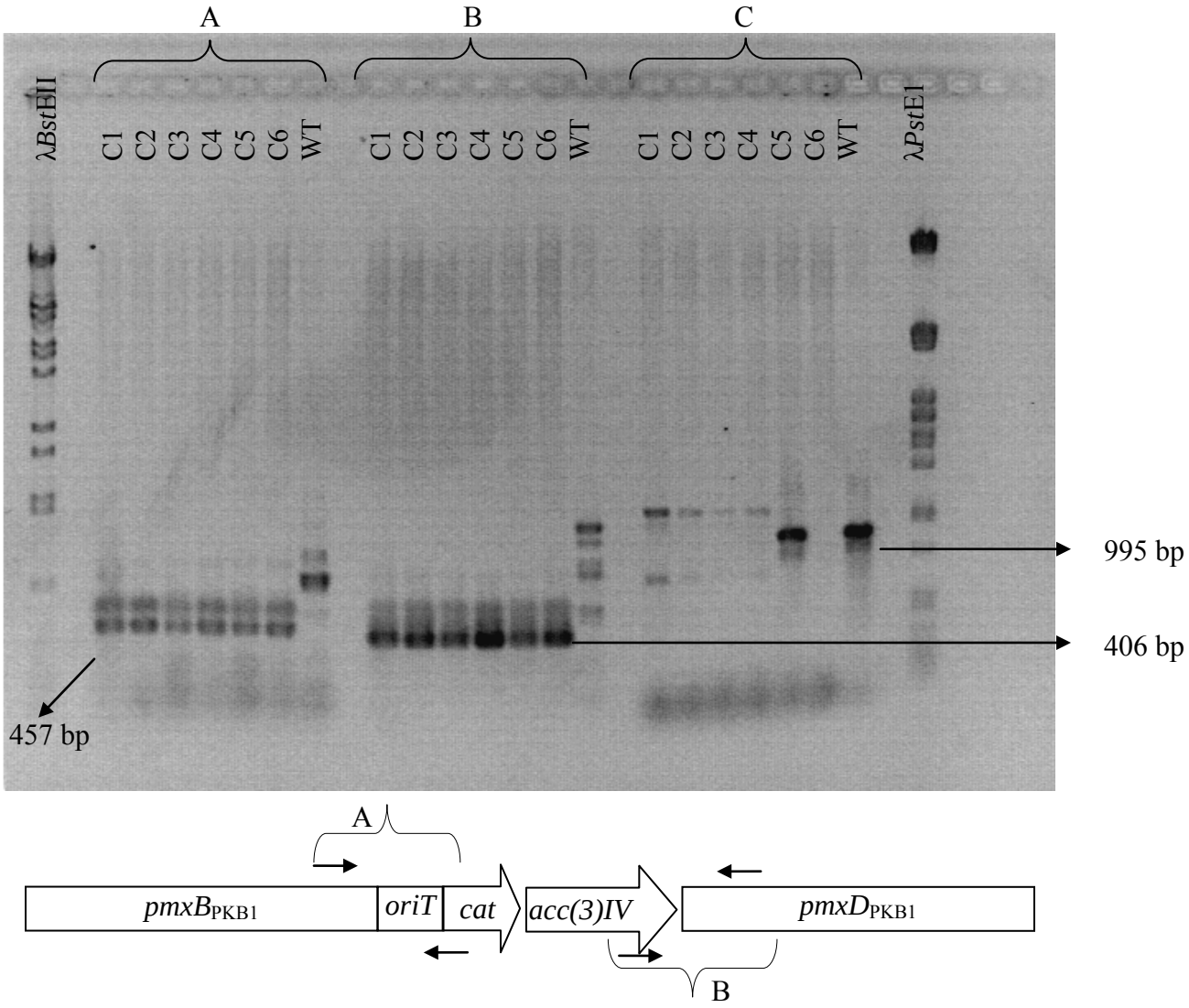


Figure 3.12.3.1: Confirmation of *P. polymyxa ΔpmxC*$_{PKB1}$ mutants by PCR. PCR was carried out with three sets of primers with the genomic DNA of six ex-conjugant (C1-C6) candidate *P. polymyxa ΔpmxC*$_{PKB1}$ mutants and WT *P. polymyxa*. (A) Two sets of primers (pSHA32 and RED-SEQ-UP) were used to determine the location of the antibiotic cassette at the 5' end. (B) Two primers pSHA33 and RED-SEQ-DWN were used to determine the location of the antibiotic cassette at the 3' end. (C) Two primers pSHA34 and pSHA35 were used to detect the presence of *pmxC*$_{PKB1}$ (WT should produce a PCR product of 995 bp).

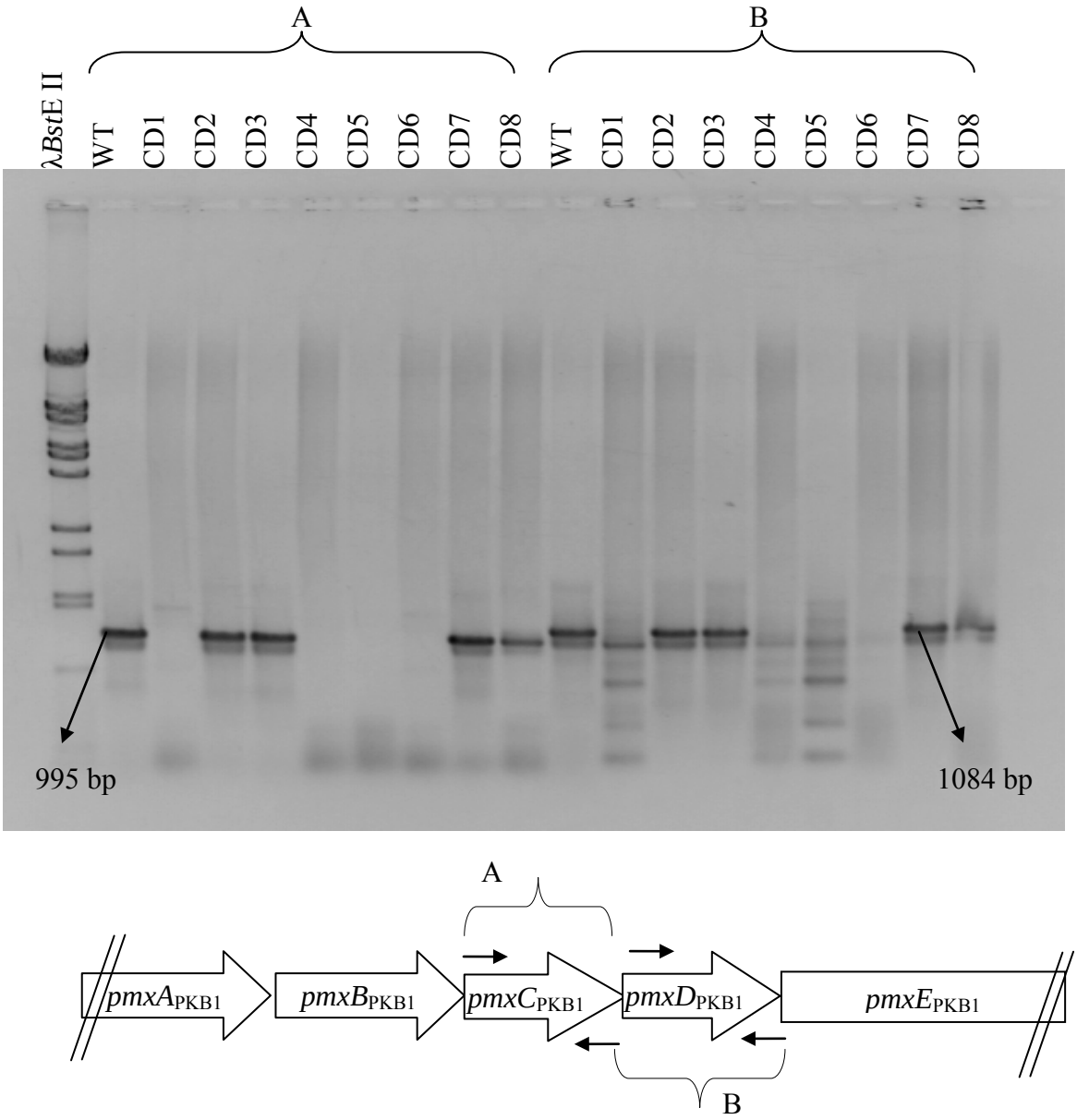


Figure 3.12.5.1: Confirmation of loss of *pmxC+D* in *P. polymyxa ΔpmxC+D*$_{PKB1}$ mutants by PCR. PCR was carried out with two sets of primers with the genomic DNA of eight ex-conjugant (CD1-CD8) candi date *P. polymyxa ΔpmxC+D*$_{PKB1}$ mutants a nd W T *P. polymyxa*. (A) T wo s ets of pr imers (pSHA34 a nd pS HA35) w ere us ed t o de tect t he presence of *pmxC*$_{PKB1}$. (B) T wo pr imers pS HA29 a nd pS HA30 w ere us ed t o de tect t he presence of *pmxD*$_{PKB1}$.

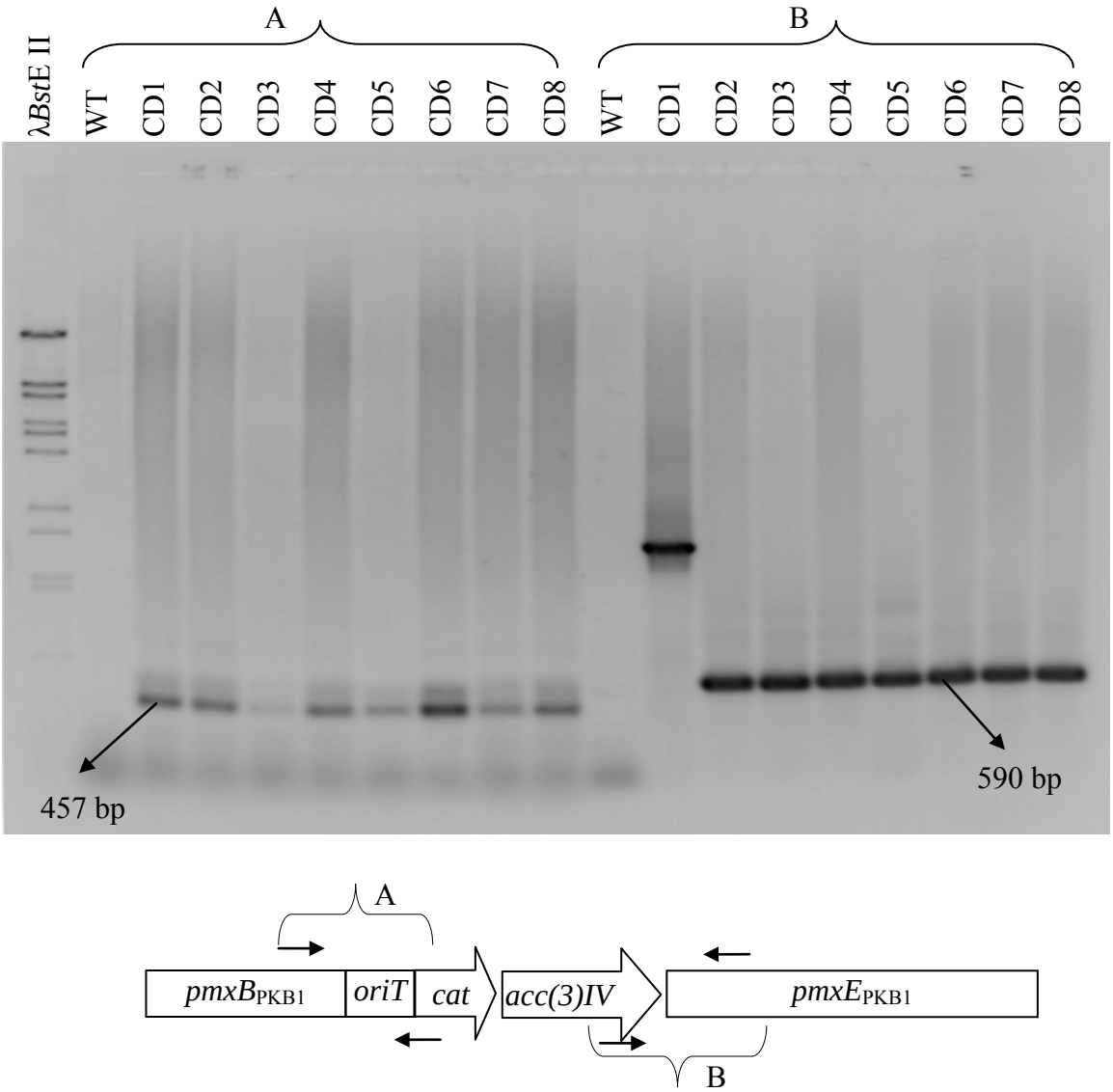


Figure 3.12 .5.2: C onfirmation of t he pr esence of the Apra^{R}Cm^{R}*oriT* cassette in *P. polymyxa ΔpmxC+D*$_{PKB1}$ mutants by PCR. PCR was carried out with two sets of primers with t he g enomic D NA of e ight e x-conjugant (CD1-CD8) candi date *P. polymyxa ΔpmxC+D*$_{PKB1}$ mutants a nd W T *P. polymyxa*. (A) T wo s ets of pr imers (pSHA32 a nd RED-SEQ-UP) were used to determine the orientation of the antibiotic cassette at the 5' end. (B) T wo pr imers pS HA28 a nd RED-SEQ-DWN were us ed t o determine the orientation of the antibiotic cassette at the 3' end.

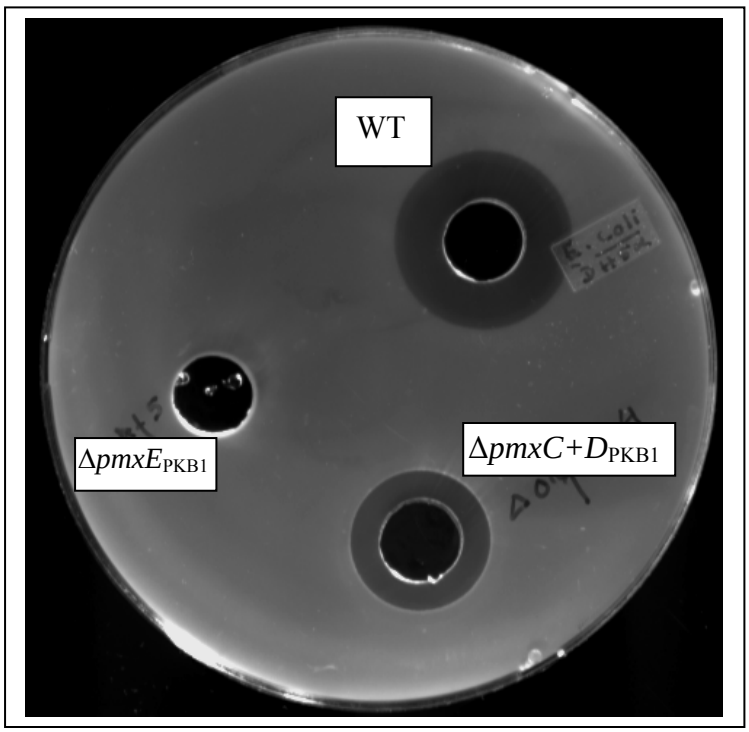


Figure 3.12. 6.1: Agar di ffusion bi oassay of s upernatant of 3 d ol d c ultures of W T, Δ*pmxC+D*$_{PKB1}$ and Δ*pmxE*$_{PKB1}$ *P. polymyxa* PKB1. The bacterial strains were grown in GSC medium at 30°C for 3 da ys in a rotary shaker at 250 rpm. The culture supernatant was harvested by centrifugation and mixed with methanol to a final concentration of 20% methanol. The mixture was centrifuged again and the clear supernatant was bioassayed.

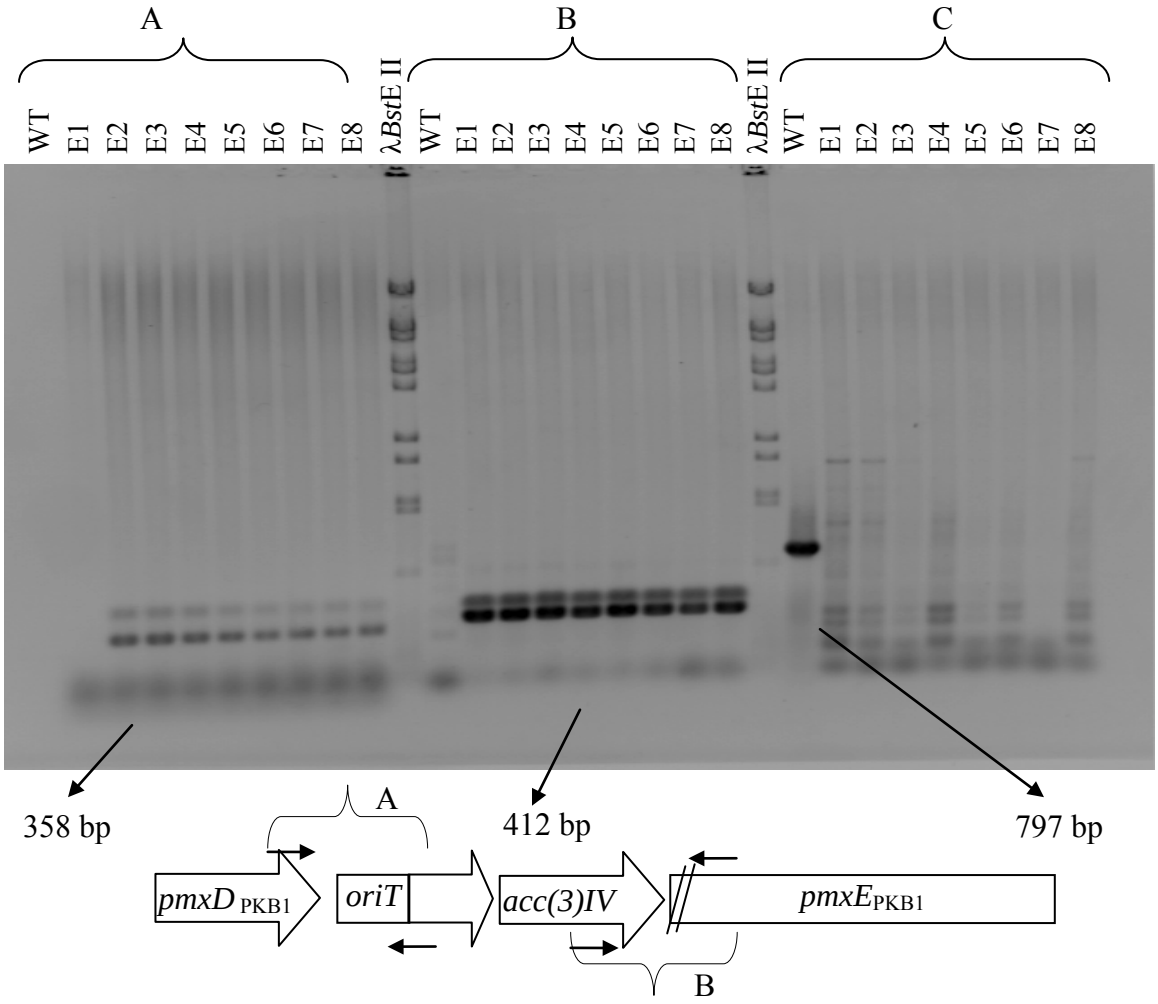


Figure 3.12.7.1: Confirmation of *P. polymyxa* $\Delta pmxE_{PKB1}$ mutants by PCR. PCR was performed with three sets of primers with the genomic DNA of eight ex-conjugant (E1-E8) candidate *P. polymyxa* $\Delta pmxE_{PKB1}$ mutants and WT *P. polymyxa*. (A) Two sets of primers (pSHA40 and RED-SEQ-UP) were used to determine the location of the antibiotic cassette at the 5' end. (B) Two primers pSHA39 and RED-SEQ-DWN were used to determine the location of the antibiotic cassette at the 3' end. (C) Two primers pSHA41 and pSHA42 were used to detect the absence of the deleted portion of $pmxE_{PKB1}$. The double line indicates that the 5' end of $pmxE_{PKB1}$ was removed by the mutation.

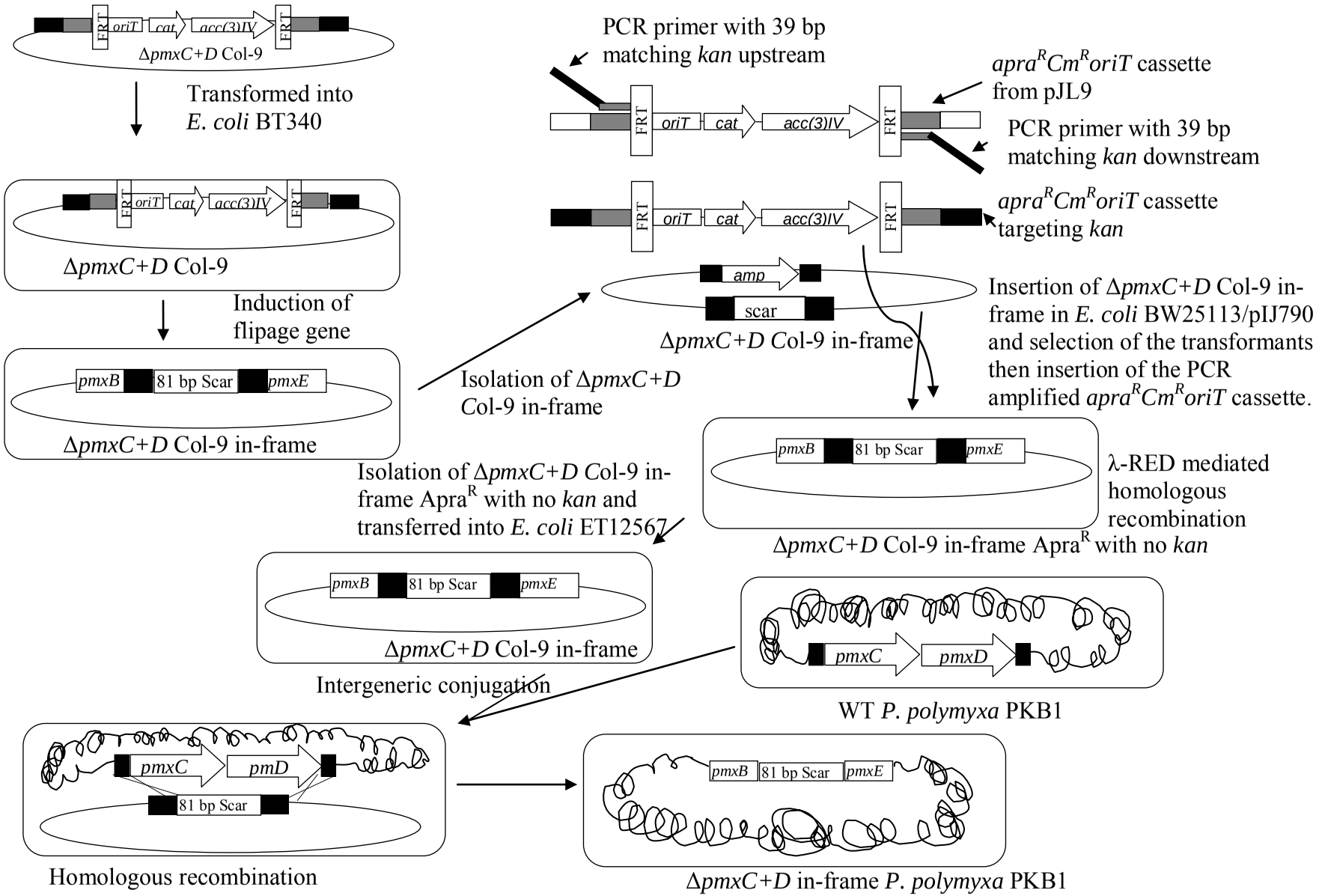

FRT
oriT
cat
acc(3)IV
FRT
ΔpmxC+D Col-9
Transformed into
E. coli BT340
ΔpmxC+D Col-9
Induction of
flipage gene
pmxB
81 bp Scar
pmxE
ΔpmxC+D Col-9 in-frame
Isolation of ΔpmxC+D
Col-9 in-frame
PCR primer with 39 bp
matching kan upstream
apraRCmRoriT cassette
from pJL9
PCR primer with 39 bp
matching kan downstream
apraRCmRoriT cassette
targeting kan
amp
scar
ΔpmxC+D Col-9 in-frame
Insertion of ΔpmxC+D Col-9 in-
frame in E. coli BW25113/pIJ790
and selection of the transformants
then insertion of the PCR
amplified apraRCmRoriT cassette.
λ-RED mediated
homologous
recombination
ΔpmxC+D Col-9 in-frame ApraR with no kan
Isolation of ΔpmxC+D Col-9 in-
frame ApraR with no kan and
transferred into E. coli ET12567
ΔpmxC+D Col-9 in-frame
Intergeneric conjugation
pmxC
pmxD
WT P. polymyxa PKB1
pmxC
pmD
81 bp Scar
Homologous recombination
ΔpmxC+D in-frame P. polymyxa PKB1

Figure 3.12 .9.1: Schematic representation of PCR targeted mutagenesis to generate *P. polymyxa* Δ*pmxC*+D_{PKB1} in-frame mutants.

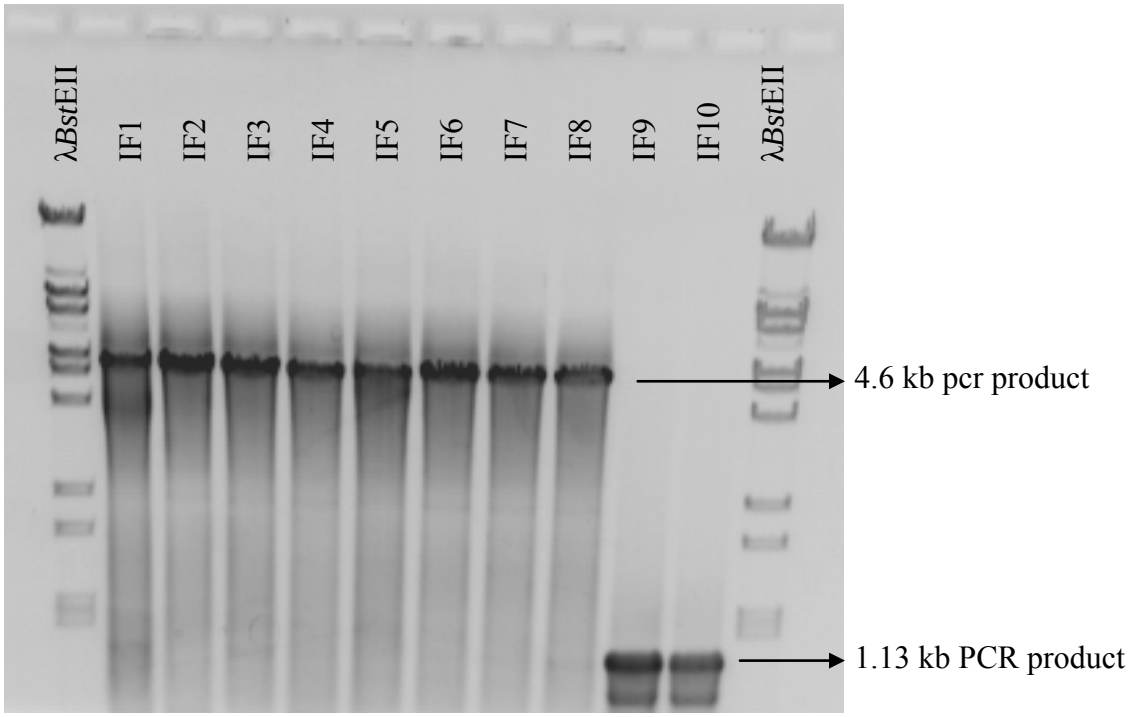


Figure 3.12.9.2: Confirmation of the in-frame (IF) mutant of Δ*pmxC+D*$_{PKB1}$ by PCR. Two primers, pSHA32 and pSHA28, were used to detect for the presence of the 81-bp 'scar'. These two primers amplify the entire *pmxC+D* genes in the wild type strain. The numbers on the right indicate the size of the PCR product for WT (4.6 kb) and for the in-frame Δ*pmxC+D*$_{PKB1}$ mutant (1.13 kb containing the 81 bp 'scar').

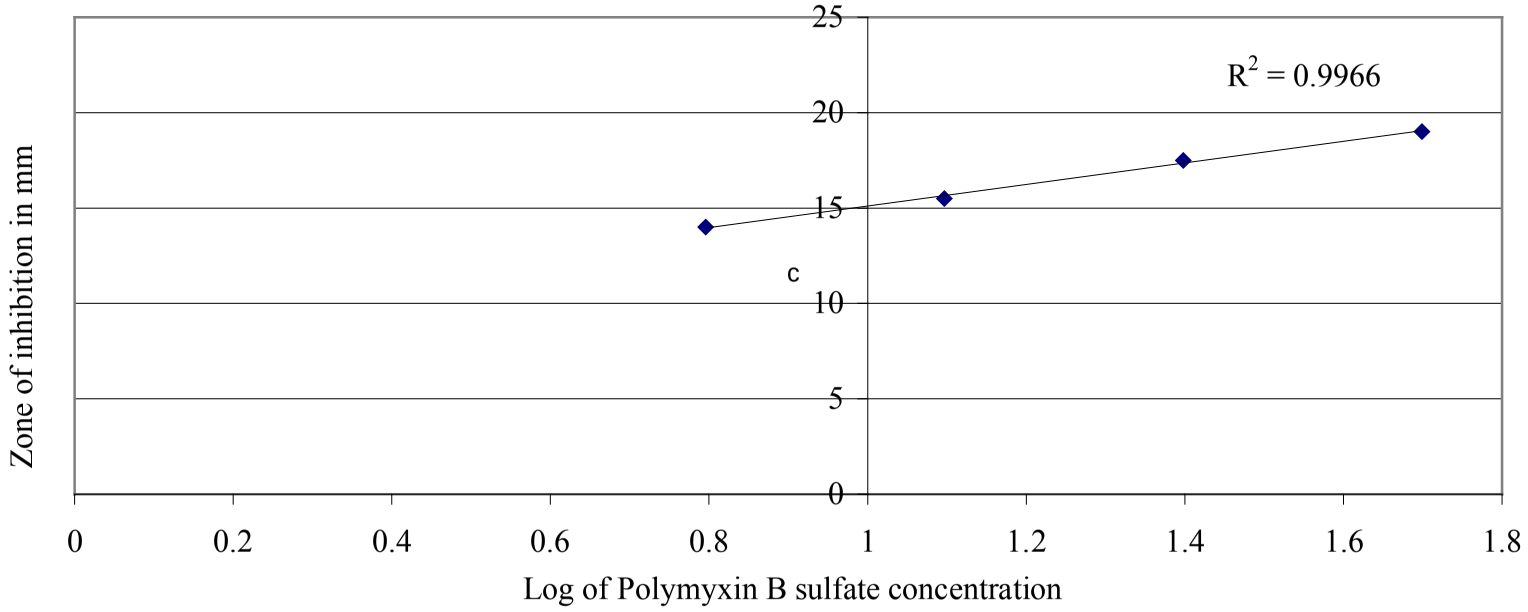


Figure 3.14.1: Standard curve relating zone of inhibition to amount of polymyxin B standard as determined by agar diffusion bioassay. The standard curve was generated by applying different amounts of standard polymyxin B sulfate (Sigma) in a constant volume of 100 µl to wells cut into LB agar plates inoculated with *E. coli* DH5α. Zones of inhibition were measured after incubation at 37°C for 16 h. The calculated amount of polymyxins produced in WT culture supernatant = 0.5 mg/ml.

Strains	Zone of inhibition (mm)
WT *P. polymyxa* PKB1	21
Δ*pmxC P. polymyxa* PKB1	18
Δ*pmxD P. polymyxa* PKB1	16
Δ*pmxC+D P. polymyxa* PKB1	16
Δ*pmxE P. polymyxa* PKB1	21
Δ*fusA P. polymyxa* PKB1 (Negative control)	None

Figure 3.15.1: Fusaricidin bioassay of culture supernatants of WT and different mutants of *P. polymyxa* PKB1. The agar diffusion bioassay was carried out in PDA medium against *L. maculans* as an indicator organism. The size of the well was 8 mm.

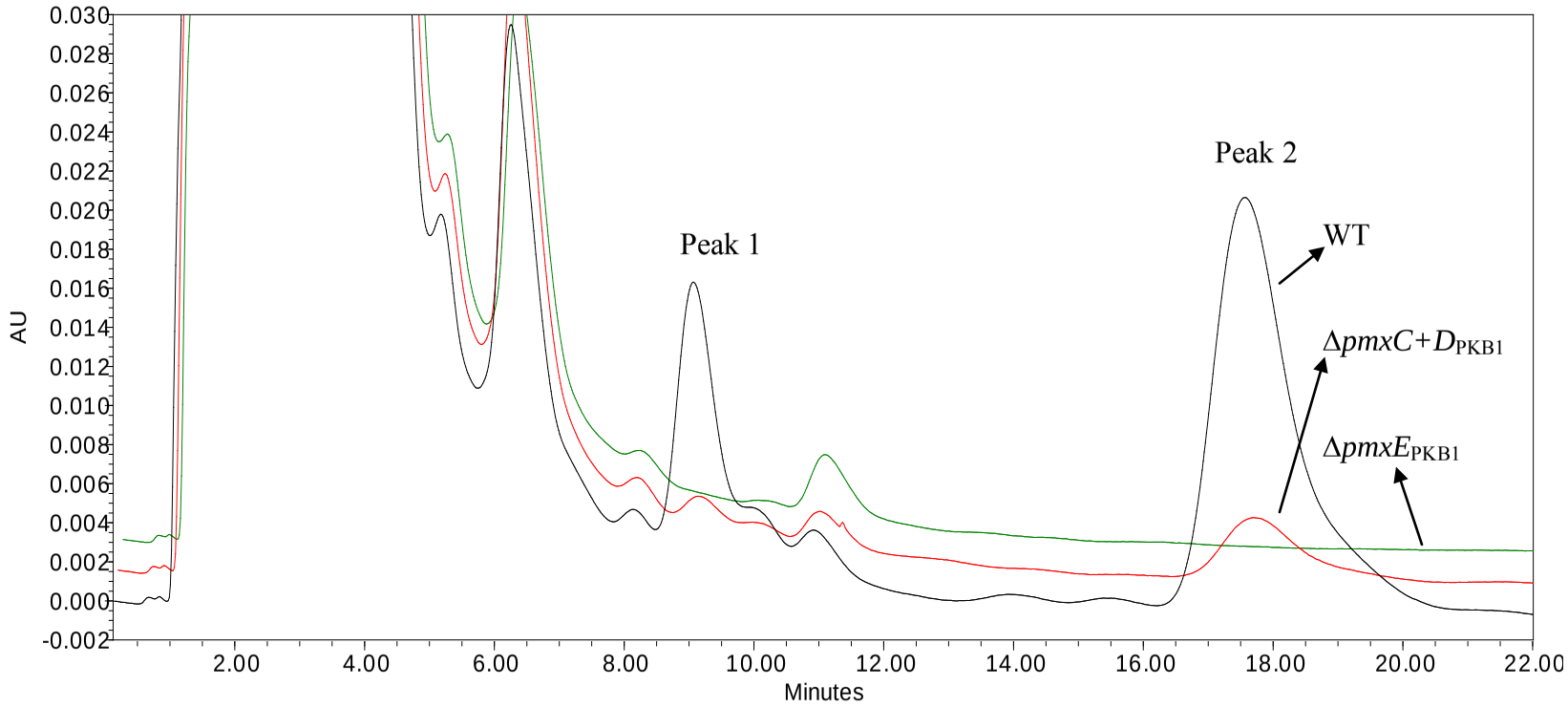


Figure 3.16.1: HPLC analysis of culture supernatants of WT *P. polymyxa* PKB1 and different mutant strains and comparison of the amount of two compounds produced by them. One hundred microlitre amounts of samples were analyzed on a column with an isocratic solvent system containing 0.1M Na_2HPO_4 (pH:3.0) and acetonitrile in a ratio of 77:23 for 22 min.

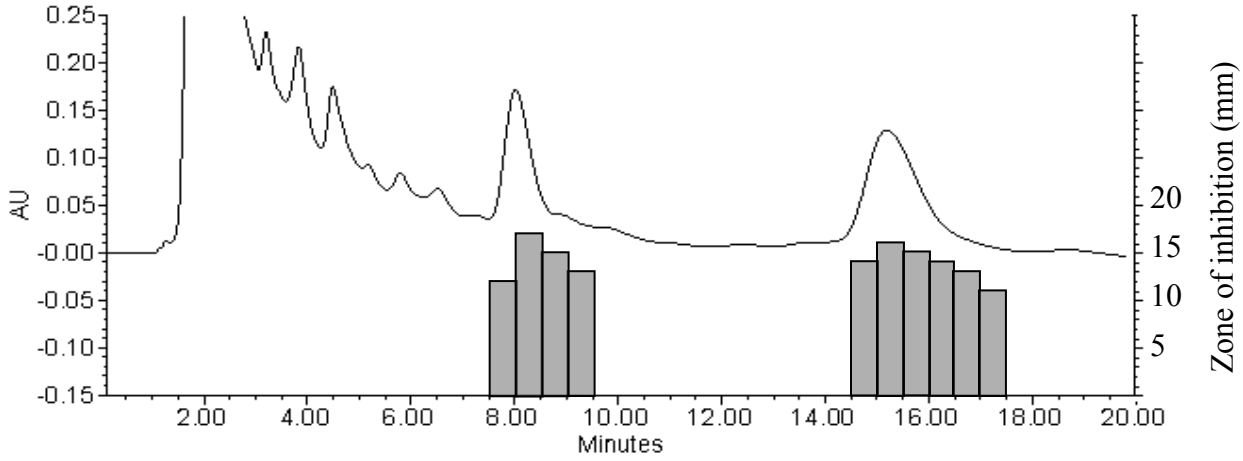


Figure 3.16.2: Antibacterial activity profile of the HPLC fractions collected from WT culture supernatant analysis. In the figure the UV chromatograph at 212 nm for the WT sample is shown. Thirty-second fractions were collected for the entire run and each fraction was checked for bioactivity by agar diffusion bioassay against *E. coli* DH5α. Bioactivity was only observed with the fractions corresponding to the peaks 1 and 2. The zone of inhibition was measured in mm and is shown in the right axis.

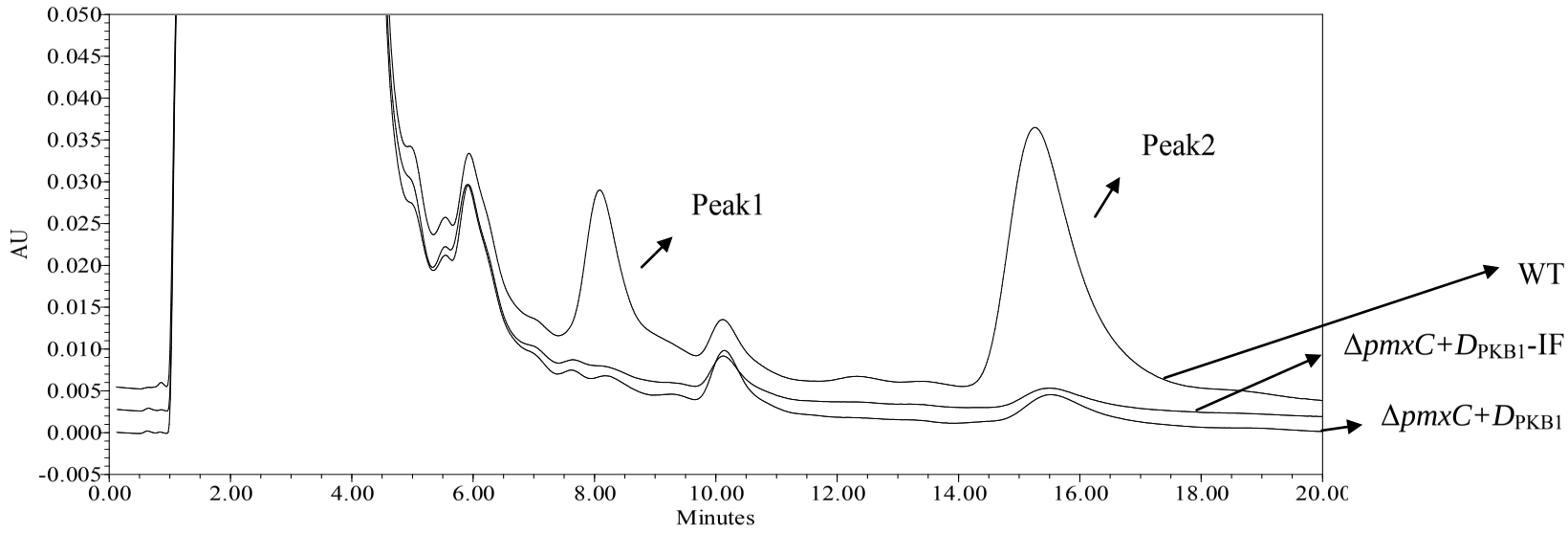


Figure 3.16.3: Comparison of the amount of antibacterial compounds produced by WT, Δ*pmxC*+D_{PKB1}deletion mutant and in-frame (IF) mutant. 100 µl of 3 d old culture supernatants of WT and mutants were analyzed by HPLC. The samples were analyzed on a Bondclone C_{18} column with an isocratic solvent system containing 0.1M Na_2HPO_4 (pH:3.0) and acetonitrile in a ratio of 77:23 for 20 min. In the figure the UV chromatograph at 212 nm for the sample is shown.

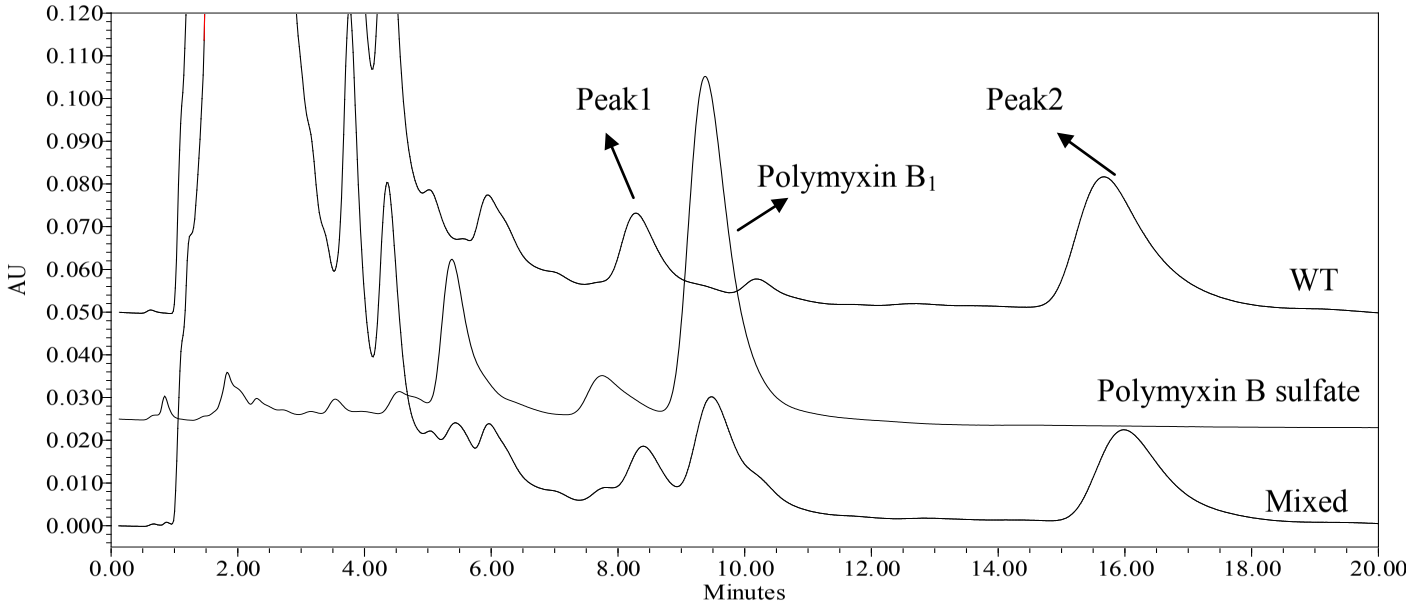


Figure 3.16.4: HPLC profile of WT culture supernatant and the standard polymyxin B sulfate solution. The samples were analyzed on a Bondclone C_{18} column with an isocratic solvent system containing 0.1M Na_2HPO_4 (pH:3.0) and acetonitrile in a ratio of 77:23 for 20 min. In the figure the UV chromatograph at 212 nm for the sample is shown.

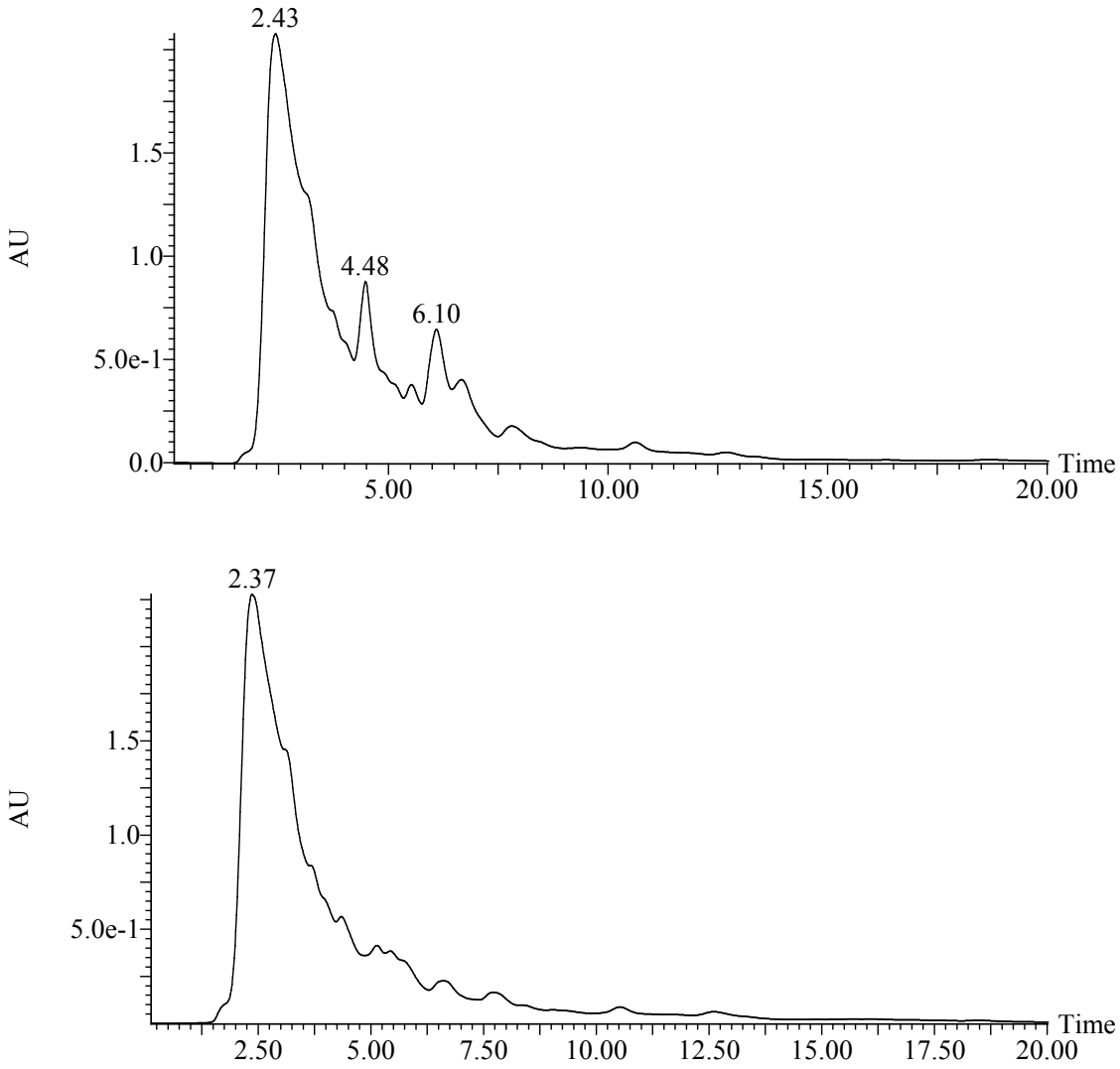


Figure 3.17 .1: T he U V-absorbance c hromatograph a t 2 12 n m from L C-MS analysis of WT and Δ*pmxE*$_{PKB1}$ culture supernatants. The peaks at retention times 4.48 min and 6.10 min observed in the WT chromatograph (A) and missing in the Δ*pmxE*$_{PKB1}$ chromatograph (B) were speculated to be peaks for polymyxins.

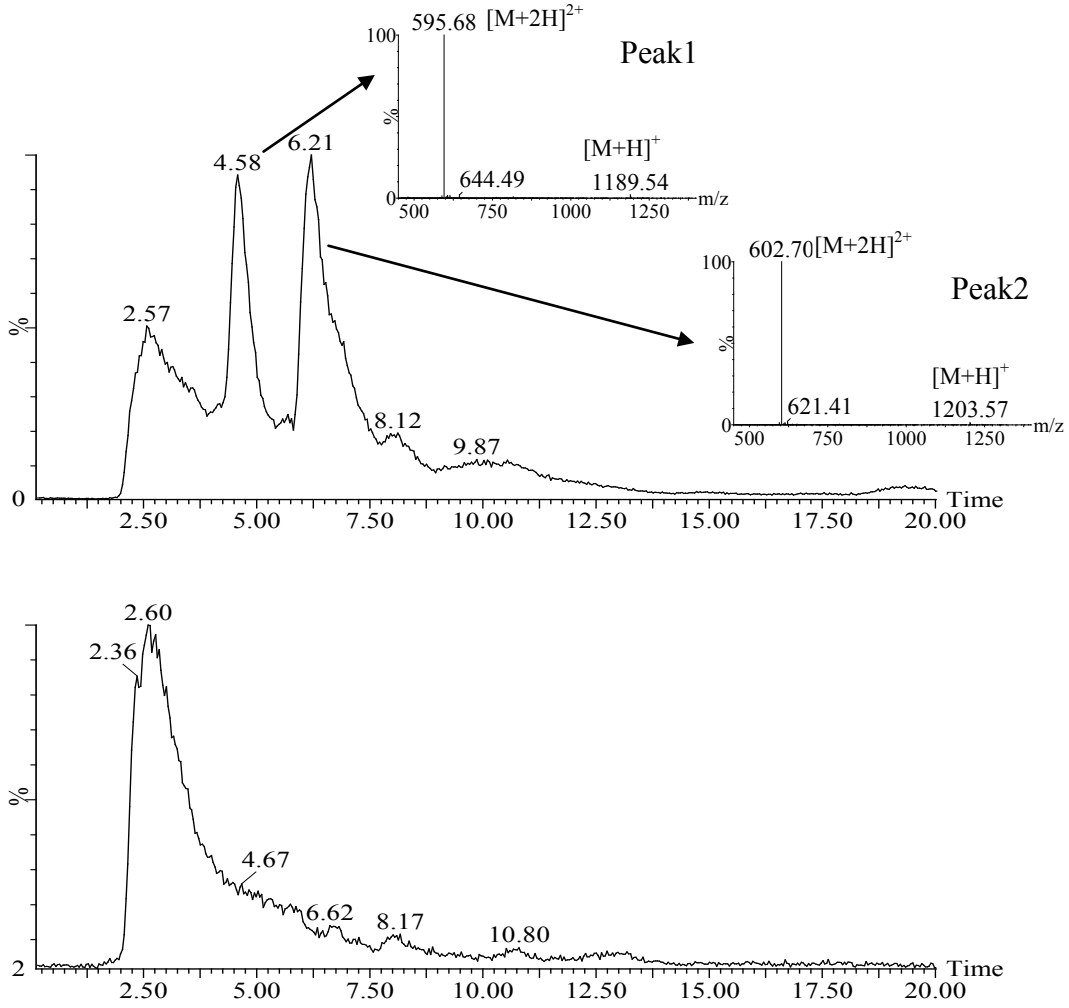


Figure 3.17 .2: T otal ion current c hromatogram f rom the mass s pectrometric analysis of *P. polymyxa* PKB1 WT and Δ*pmxE*$_{PKB1}$ culture s upernatants. T he peaks at retention times 4.58 and 6.21 min observed in the WT chromatogram (A) and missing in the Δ*pmxE*$_{PKB1}$ chromatogram (B) were shown to have molecular masses of 1188.54 and 1202.57 respectively.

4. Conclusion and significance

Polymyxins are clinically important lipopeptide antibiotics produced by various species of *Paenibacillus*. Identification and characterization of the genes responsible for polymyxin production could aid in understanding the detailed mechanisms of polymyxin biosynthesis.

The first characterization of a gene cluster involved in polymyxin production was reported for *P. polymyxa* E681 in 2009 and was an outcome of a complete genome sequencing project (Choi et al. 2009). In this present study, the gene cluster encoding the NRPS for polymyxin production in *P. polymyxa* PKB1 has been identified and characterized. The gene cluster for polymxyin in *P. polymyxa* PKB1 was compared with that in *P. polymyxa* E681 and a high homology at the nucleotide level was observed. In both organisms the gene clusters consist of five ORFs: three encode the NRPS genes, two encode two ABC type transporter genes and all five ORFs are transcriptionally oriented in the same direction. The genes are also arranged in the same unusual order with the gene encoding the first five modules of the NRPS coming at the end of the cluster. Significant information about the genes flanking the polymyxin gene cluster in *P. polymyxa* PKB1 was not obtained in this study due to a lack of suitable clones. However the adjacent areas of the polymyxin gene cluster of *P. polymyxa* E681 were studied, but no genes for biosynthesis of the fatty acid moiety of polymyxins, or genes with regulatory functions were identified.

Mutational studies were carried out to verify the involvement of the genes of the polymyxin gene cluster in polymyxin biosynthesis in *P. polymyxa* PKB1.

The PCR targeting gene disruption method developed for *Streptomyces* was adapted to make different *P. polymyxa* PKB1 mutants with the Apra^{R}Cm^R*oriT* cassette, previously used in the fusaricidin biosynthetic gene cluster study in the same organism. No trace of polymyxins was detected by agar diffusion bioassay against *E. coli* as well as by HPLC analysis in the culture supernatant of a *pmxE*-disrupted *P. polymyxa* PKB1 mutant. This observation confirmed the involvement of the $pmxE_{PKB1}$ biosynthetic gene of the polymyxin gene cluster in polymyxin production. Mutants were also generated by deleting $pmxC_{PKB1}$ and $pmxD_{PKB1}$ individually and by deleting both together to determine their involvement in polymyxin production. All of these mutations reduced the production of polymyxin in the culture supernatants thereby proving their association with polymyxin production. An in-frame mutation of these transporter genes was also made to eliminate the possibility of a polar effect on the downstream synthetic genes.

Deletion of the ABC transporter-like genes also reduced the production of fusaricidin in the culture supernatants. A possible explanation for this observation is that these ABC transporters might be involved in exporting both types of lipopeptide metabolites from inside the cells.

The chemical structure of the polymyxins was identified shortly after they were discovered. Polymyxins are decapeptides and for the most common varieties of polymyxins, six out of ten of the amino acids are L-DAB. The differences among the different polymyxins are due to the variations in the component amino acids in position six and seven and also due to different fatty acid moieties.

However, the gene cluster for polymyxin biosynthesis in *P. polymyxa* PKB1 as well as in *P. polymyxa* E681 suggested that the DAB residue present in the position three is a D-Dab, as the NRPS module that is responsible for the incorporation of the amino acid in the third position has an E domain. The HPLC and LC-MS analysis suggested that under the studied conditions *P. polymyxa* PKB1 produces at least two forms of polymyxins and their molecular weights match the molecular weights of polymyxin B_1 and B_2. No forms of polymyxin B have been reported so far containing a D-DAB in position three. However the information obtained from the NRPS specificity codon and domain arrangement suggested a D-DAB in position three in the newly identified polymyxin molecules. This finding was strengthened by the observation that the commercially available polymyxin B did not co-elute upon HPLC with the polymyxins produced in the culture supernatant of *P. polymyxa* PKB1. No polymyxins other than B_1 and B_2 are known to have the same molecular weight as the polymyxins produced by *P. polymyxa* PKB1, and yet they are not polymyxin B_1 and B_2. Therefore, these two polymyxins are considered as new forms of polymyxin B, and hence are referred to as D-DAB-polymyxin B_1 and D-DAB-polymyxin B_2.

5. Future study

This study identified the gene cluster encoding the NRPS for polymyxin biosynthesis in *P. polymyxa* PKB1. This genetic information provided insights into how the peptide parts of the polymyxin molecules are synthesized and also indicated possible mechanisms of their export outside the cells. Despite considerable effort, the adjacent areas of the polymyxin gene cluster could not be explored fully due to a lack of appropriate clones. Therefore, the complete picture for polymyxin biosynthesis is still not available. However, the complete genome sequence of *Paenibacillus polymyxa* E681, a polymyxin A producing bacterium, has become available recently (GenBank Accession No.: CP000154). Exploring the adjacent area of its polymyxin gene cluster did not reveal any significant information about the regulation of the polymyxin genes or the biosynthesis of the fatty acid moiety of the polymyxin molecules. The fatty acid moiety of the polymyxin molecules is of great importance for their antibacterial activity as it supports additional interaction with the Gram negative bacterial cell envelope. Therefore, studies aimed at understanding the biosynthesis of the fatty acid component, its mechanism of transfer to become bound to the peptide and the factors that determine the type of fatty acid attached to the peptide part are still needed.

Since the amino acid composition and arrangement in the polymyxin molecule are very important determinants of biological properties like toxicity and antimicrobial activity, in-depth analysis is still required to establish the relationships between amino acid composition and position in the polymyxin

molecules and their therapeutic properties. Non-proteinogenic amino acids such as D-amino acids are generally believed to make the peptide more antibacterial but also may make it more toxic. Optimal combinations of these unusual amino acids might give polymyxins that are more antibacterial but have less toxicity to the human host. The third position D-DAB found in the *P. polymyxa* PKB1 polymyxins is unusual; therefore more complete validation of the identity of that residue is required.

An evolutionary study of polymyxin gene clusters is also of interest. The polymyxin gene cluster in *P. polymyxa* E681 shows very high homology at the nucleotide level to that in *P. polymyxa* PKB1. However, *P. polymyxa* E681 has been reported to produce polymyxin A whereas *P. polymyxa* PKB1 produces novel forms of polymyxin B. These variations appear to be due to differences encoded in their polymyxin synthetase genes rather than merely due to amino acid substitutions based on composition of the growth medium. Therefore, these gene clusters for polymyxin biosynthesis must have been changed during their evolutionary journey by swapping domains or modules from other NRPS genes obtained by lateral or horizontal gene transfer. As sequences for more polymyxin gene clusters become available their relationships will become more clear.

6. References

Arnold, T.M., Forrest, G.N., and Messmer, K.J. 2007. Polymyxin antibiotics for Gram-negative infections. Am. J. Health. Syst. Pharm. **64**(8): 819-826.

Ash, C., Priest, F.G., and Collins, M.D. 1993. Molecular identification of rRNA group 3 bacilli (Ash, Farrow, Wallbanks and Collins) using a PCR probe test. Proposal for the creation of a new genus *Paenibacillus*. Antonie Van Leeuwenhoek **64**(3-4): 253-260.

Bairoch, A., and Boeckmann, B. 1991. The SWISS-PROT protein sequence data bank. Nucleic Acids Res. **19 Suppl**: 2247-2249.

Balibar, C.J., Vaillancourt, F.H., and Walsh, C.T. 2005. Generation of D amino acid residues in assembly of arthrofactin by dual condensation/epimerization domains. Chem. Biol. **12**(11): 1189-1200.

Beatty, P.H., and Jensen, S.E. 2002. *Paenibacillus polymyxa* produces fusaricidin-type antifungal antibiotics active against *Leptosphaeria maculans*, the causative agent of blackleg disease of canola. Can. J. Microbiol. **48**(2): 159-169.

Bengoechea, J.A., and Skurnik, M. 2000. Temperature-regulated efflux pump/potassium antiporter system mediates resistance to cationic antimicrobial peptides in *Yersinia*. Mol. Microbiol. **37**(1): 67-80.

Birnboim, H.C., and Doly, J. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**(6): 1513-1523.

Bonomo, R.A., and Szabo, D. 2006. Mechanisms of multidrug resistance in *Acinetobacter* species and *Pseudomonas aeruginosa*. Clin. Infect. Dis. **43 Suppl 2**: S49-56.

Bruner, S.D., Weber, T., Kohli, R.M., Schwarzer, D., Marahiel, M.A., Walsh, C.T., and Stubbs, M.T. 2002. Structural basis for the cyclization of the lipopeptide antibiotic surfactin by the thioesterase domain SrfTE. Structure **10**(3): 301-310.

Bucklin, S.E., Lake, P., Logdberg, L., and Morrison, D.C. 1995. Therapeutic efficacy of a polymyxin B-dextran 70 conjugate in experimental model of endotoxemia. Antimicrob. Agents Chemother. **39**(7): 1462-1466.

Campos, M.A., Vargas, M.A., Regueiro, V., Llompart, C.M., Alberti, S., and Bengoechea, J.A. 2004. Capsule polysaccharide mediates bacterial resistance to antimicrobial peptides. Infect. Immun. **72**(12): 7107-7114.

Challis, G.L., and Naismith, J.H. 2004. Structural aspects of non-ribosomal peptide biosynthesis. Curr. Opin. Struct. Biol. **14**(6): 748-756.

Challis, G.L., Ravel, J., and Townsend, C.A. 2000. Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains. Chem. Biol. **7**(3): 211-224.

Choi, S., Park, S., Kim, R., Kim, S., Lee, C., Kim, J.F., and Park, S. 2009. Identification of a polymyxin synthetase gene cluster of *Paenibacillus polymyxa* and heterologous expression of the gene in *Bacillus subtilis*. J. Bacteriol. **191**(10): 3350-3358.

Clausell, A., Garcia-Subirats, M., Pujol, M., Busquets, M.A., Rabanal, F., and Cajal, Y. 2007. Gram-negative outer and inner membrane models: insertion of cyclic cationic lipopeptides. J Phys Chem B **111**(3): 551-563.

Conti, E., Stachelhaus, T., Marahiel, M.A., and Brick, P. 1997. Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S. EMBO J. **16**(14): 4174-4183.

DeCrescenzo Henriksen, E., Phillips, D.R., and Peterson, J.B. 2007. Polymyxin E production by *P. amylolyticus*. Lett. Appl. Microbiol. **45**(5): 491-496.

Delgado, M.A., Mouslim, C., and Groisman, E.A. 2006. The PmrA/PmrB and RcsC/YojN/RcsB systems control expression of the *Salmonella* O-antigen chain length determinant. Mol. Microbiol. **60**(1): 39-50.

Dieckmann, R., Lee, Y.O., van Liempt, H., von Dohren, H., and Kleinkauf, H. 1995. Expression of an active adenylate-forming domain of peptide synthetases corresponding to acyl-CoA-synthetases. FEBS Lett. **357**(2): 212-216.

Du, L., He, Y., and Luo, Y. 2008. Crystal structure and enantiomer selection by D-alanyl carrier protein ligase DltA from *Bacillus cereus*. Biochemistry **47**(44): 11473-11480.

Du, L., Chen, M., Sanchez, C., and Shen, B. 2000. An oxidation domain in the BlmIII non-ribosomal peptide synthetase probably catalyzing thiazole formation in the biosynthesis of the anti-tumor drug bleomycin in *Streptomyces verticillus* ATCC15003. FEMS Microbiol. Lett. **189**(2): 171-175.

Duitman, E.H., Hamoen, L.W., Rembold, M., Venema, G., Seitz, H., Saenger, W., Bernhard, F., Reinhardt, R., Schmidt, M., Ullrich, C., Stein, T., Leenders, F., and Vater, J. 1999. The mycosubtilin synthetase of *Bacillus subtilis*

ATCC6633: a multifunctional hybrid between a peptide synthetase, an amino transferase, and a fatty acid synthase. Proc. Natl. Acad. Sci. U. S. A. **96**(23): 13294-13299.

Ehmann, D.E., Trauger, J.W., Stachelhaus, T., and Walsh, C.T. 2000. Aminoacyl-SNACs as small-molecule substrates for the condensation domains of nonribosomal peptide synthetases. Chem. Biol. **7**(10): 765-772.

Elemam, A., Rahimian, J., and Doymaz, M. 2010. In vitro evaluation of antibiotic synergy for polymyxin B-resistant carbapenemase-producing *Klebsiella pneumoniae*. J. Clin. Microbiol. **48**(10): 3558-3562.

Evans, M.E., Feola, D.J., and Rapp, R.P. 1999. Polymyxin B sulfate and colistin: old antibiotics for emerging multiresistant Gram-negative bacteria. Ann. Pharmacother. **33**(9): 960-967.

Falagas, M.E., Rafailidis, P.I., and Matthaiou, D.K. 2010. Resistance to polymyxins: Mechanisms, frequency and treatment options. Drug Resist Updat **13**(4-5): 132-138.

Falagas, M.E., and Kasiakou, S.K. 2006. Toxicity of polymyxins: a systematic review of the evidence from old and recent studies. Crit. Care **10**(1): R27.

Fischbach, M.A., and Walsh, C.T. 2006. Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: logic, machinery, and mechanisms. Chem. Rev. **106**(8): 3468-3496.

Fulda, M., Heinz, E., and Wolter, F.P. 1994. The *fadD* gene of *Escherichia coli* K12 is located close to rnd at 39.6 min of the chromosomal map and is a new member of the AMP-binding protein family. Mol. Gen. Genet. **242**(3): 241-249.

Gales, A.C., Jones, R.N., and Sader, H.S. 2006. Global assessment of the antimicrobial activity of polymyxin B against 54 731 clinical isolates of Gram-negative bacilli: report from the SENTRY antimicrobial surveillance programme (2001-2004). Clin. Microbiol. Infect. **12**(4): 315-321.

Go, E.S., Urban, C., Burns, J., Kreiswirth, B., Eisner, W., Mariano, N., Mosinka-Snipas, K., and Rahal, J.J. 1994. Clinical and molecular epidemiology of *Acinetobacter* infections sensitive only to polymyxin B and sulbactam. Lancet **344**(8933): 1329-1332.

Govaerts, C., Orwa, J., Van Schepdael, A., Roets, E., and Hoogmartens, J. 2002a. Liquid chromatography-ion trap tandem mass spectrometry for the

characterization of polypeptide antibiotics of the colistin series in commercial samples. J. Chromatogr. A **976**(1-2): 65-78.

Govaerts, C., Orwa, J., Van Schepdael, A., Roets, E., and Hoogmartens, J. 2002b. Characterization of polypeptide antibiotics of the polymyxin series by liquid chromatography electrospray ionization ion trap tandem mass spectrometry. J. Pept. Sci. **8**(2): 45-55.

Grunewald, J., and Marahiel, M.A. 2006. Chemoenzymatic and template-directed synthesis of bioactive macrocyclic peptides. Microbiol. Mol. Biol. Rev. **70**(1): 121-146.

Grunewald, J., Sieber, S.A., Mahlert, C., Linne, U., and Marahiel, M.A. 2004. Synthesis and derivatization of daptomycin: a chemoenzymatic route to acidic lipopeptide antibiotics. J. Am. Chem. Soc. **126**(51): 17025-17031.

Guillemette, T., Sellam, A., and Simoneau, P. 2004. Analysis of a nonribosomal peptide synthetase gene from *Alternaria brassicae* and flanking genomic sequences. Curr. Genet. **45**(4): 214-224.

Gunn, J.S., Ryan, S.S., Van Velkinburgh, J.C., Ernst, R.K., and Miller, S.I. 2000. Genetic and functional analysis of a PmrA-PmrB-regulated locus necessary for lipopolysaccharide modification, antimicrobial peptide resistance, and oral virulence of *Salmonella enterica* serovar typhimurium. Infect. Immun. **68**(11): 6139-6146.

Gunn, J.S., Lim, K.B., Krueger, J., Kim, K., Guo, L., Hackett, M., and Miller, S.I. 1998. PmrA-PmrB-regulated genes necessary for 4-aminoarabinose lipid A modification and polymyxin resistance. Mol. Microbiol. **27**(6): 1171-1182.

Gust, B., Challis, G.L., Fowler, K., Kieser, T., and Chater, K.F. 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**(4): 1541-1546.

Hancock, R.E.W. 1984. Alterations in outer membrane permeability. Annu. Rev. Microbiol. **38**:237-264

Hancock, R.E. 1997. Antibacterial peptides and the outer membranes of gram-negative bacilli. J. Med. Microbiol. **46**(1): 1-3.

Hancock, R.E., and Lehrer, R. 1998. Cationic peptides: a new source of antibiotics. Trends Biotechnol. **16**(2): 82-88.

Hancock, R.E. 1997. Peptide antibiotics. The Lancet **349**(9049): 418-422.

Hausmann, W., and Craig, L.C. 1954. Polymyxin B_1.[1] Fractionation, molecular weight determination, amino acid and fatty acid composition. J. Am. Chem. Soc. **76**(19): 4892-4896.

Hermsen, E.D., Sullivan, C.J., and Rotschafer, J.C. 2003. Polymyxins: pharmacology, pharmacokinetics, pharmacodynamics, and clinical applications. Infect. Dis. Clin. North Am. **17**(3): 545-562.

Hillson, N.J., and Walsh, C.T. 2003. Dimeric structure of the six-domain VibF subunit of vibriobactin synthetase: mutant domain activity regain and ultracentrifugation studies. Biochemistry **42**(3): 766-775.

Hirokawa, T., Boon-Chieng, S., and Mitaku, S. 1998. SOSUI: classification and secondary structure prediction system for membrane proteins. Bioinformatics **14**(4): 378-379.

Iijima, S., Kato, D., Yabuki, S., Niwa, O., and Mizutani, F. 2010. Enzymatically amplified electrochemical detection for lipopolysaccharide using ferrocene-attached polymyxin B and its analogue. Biosens. Bioelectron.

Jensen, T., Pedersen, S.S., Garne, S., Heilmann, C., Hoiby, N., and Koch, C. 1987. Colistin inhalation therapy in cystic fibrosis patients with chronic *Pseudomonas aeruginosa* lung infection. J. Antimicrob. Chemother. **19**(6): 831-838.

Jurtshuk, R.J., Blick, M., Bresser, J., Fox, G.E., and Jurtshuk, P.,Jr. 1992. Rapid *in situ* hybridization technique using 16S rRNA segments for detecting and differentiating the closely related Gram-positive organisms *Bacillus polymyxa* and *Bacillus macerans*. Appl. Environ. Microbiol. **58**(8): 2571-2578.

Kato, A., Latifi, T., and Groisman, E.A. 2003. Closing the loop: the PmrA/PmrB two-component system negatively controls expression of its posttranscriptional activator PmrD. Proc. Natl. Acad. Sci. U. S. A. **100**(8): 4706-4711.

Keating, T.A., and Walsh, C.T. 1999. Initiation, elongation, and termination strategies in polyketide and polypeptide antibiotic biosynthesis. Curr. Opin. Chem. Biol. **3**(5): 598-606.

Keating, T.A., Marshall, C.G., Walsh, C.T., and Keating, A.E. 2002. The structure of VibH represents nonribosomal peptide synthetase condensation, cyclization and epimerization domains. Nat. Struct. Biol. **9**(7): 522-526.

Kohli, R.M., and Walsh, C.T. 2003. Enzymology of acyl chain macrocyclization in natural product biosynthesis. Chem. Commun. (Camb) **(3)**(3): 297-307.

Kohli, R.M., Walsh, C.T., and Burkart, M.D. 2002. Biomimetic synthesis and optimization of cyclic peptide antibiotics. Nature **418**(6898): 658-661.

Koike, M., Iida, K., and Matsuo, T. 1969. Electron microscopic studies on mode of action of polymyxin. J. Bacteriol. **97**(1): 448-452.

Konz, D., Doekel, S., and Marahiel, M.A. 1999. Molecular and biochemical characterization of the protein template controlling biosynthesis of the lipopeptide lichenysin. J. Bacteriol. **181**(1): 133-140.

Kopp, F., and Marahiel, M.A. 2007. Where chemistry meets biology: the chemoenzymatic synthesis of nonribosomal peptides and polyketides. Curr. Opin. Biotechnol. **18**(6): 513-520.

Lambalot, R.H., Gehring, A.M., Flugel, R.S., Zuber, P., LaCelle, M., Marahiel, M.A., Reid, R., Khosla, C., and Walsh, C.T. 1996. A new enzyme superfamily - the phosphopantetheinyl transferases. Chem. Biol. **3**(11): 923-936.

Lautru, S., and Challis, G.L. 2004. Substrate recognition by nonribosomal peptide synthetase multi-enzymes. Microbiology **150**(Pt 6): 1629-1636.

Lewendon, A., Murray, I.A., Shaw, W.V., Gibbs, M.R., and Leslie, A.G. 1994. Replacement of catalytic histidine-195 of chloramphenicol acetyltransferase: evidence for a general base role for glutamate. Biochemistry **33**(7): 1944-1950.

Li, J. 2007. Cloning and characterization of the peptide synthetase gene cluster involved in the nonribosomal biosynthesis of fusaricidin-type antifungal antibiotics in *Paenibacillus polymyxa* PKB1. Ph.D. Thesis, University of Alberta, Edmonton, Alta.

Li, J., and Jensen, S.E. 2008. Nonribosomal biosynthesis of fusaricidins by *Paenibacillus polymyxa* PKB1 involves direct activation of a D-amino acid. Chem. Biol. **15**(2): 118-127.

Li, J., Beatty, P.K., Shah, S., and Jensen, S.E. 2007. Use of PCR-targeted mutagenesis to disrupt production of fusaricidin-type antifungal antibiotics in *Paenibacillus polymyxa*. Appl. Environ. Microbiol. **73**(11): 3480-3489.

Li, J., Nation, R.L., Milne, R.W., Turnidge, J.D., and Coulthard, K. 2005. Evaluation of colistin as an agent against multi-resistant Gram-negative bacteria. Int. J. Antimicrob. Agents **25**(1): 11-25.

Li, J., Nation, R.L., Turnidge, J.D., Milne, R.W., Coulthard, K., Rayner, C.R., and Paterson, D.L. 2006. Colistin: the re-emerging antibiotic for multidrug-resistant Gram-negative bacterial infections. Lancet Infect. Dis. **6**(9): 589-601.

Linne, U., Doekel, S., and Marahiel, M.A. 2001. Portability of epimerization domain and role of peptidyl carrier protein on epimerization activity in nonribosomal peptide synthetases. Biochemistry **40**(51): 15824-15834.

Lipmann, F. 1973. Non-ribosomal polypeptide synthesis. Acta Endocrinol. Suppl. (Copenh) **180**: 294-300.

Lipmann, F. 1971. Attempts to map a process evolution of peptide biosynthesis. Science **173**(4000): 875-884.

Lipmann, F., Gevers, W., Kleinkauf, H., and Roskoski, R.,Jr. 1971. Polypeptide synthesis on protein templates: the enzymatic synthesis of gramicidin S and tyrocidine. Adv. Enzymol. Relat. Areas Mol. Biol. **35**: 1-34.

Loutet, S.A., Mussen, L.E., Flannagan, R.S., and Valvano, M.A. 2010. A two-tier model of polymyxin B resistance in *Burkholderia cenocepacia*. Environmental Microbiology Reports: no. doi: 10.1111/j.1758-2229.2010.00222.x.

Luo, L., Kohli, R.M., Onishi, M., Linne, U., Marahiel, M.A., and Walsh, C.T. 2002. Timing of epimerization and condensation reactions in nonribosomal peptide assembly lines: kinetic analysis of phenylalanine activating elongation modules of tyrocidine synthetase B. Biochemistry **41**(29): 9184-9196.

Marahiel, M.A. 1997. Protein templates for the biosynthesis of peptide antibiotics. Chem. Biol. **4**(8): 561-567.

Marahiel, M.A. 1992. Multidomain enzymes involved in peptide synthesis. FEBS Lett. **307**(1): 40-43.

Marahiel, M.A., and Essen, L.O. 2009. Chapter 13. Nonribosomal peptide synthetases mechanistic and structural aspects of essential domains. Methods Enzymol. **458**: 337-351.

Marahiel, M.A., Stachelhaus, T., and Mootz, H.D. 1997. Modular peptide synthetases involved in nonribosomal peptide synthesis. Chem. Rev. **97**(7): 2651-2674.

Martin, N.I., Hu, H., Moake, M.M., Churey, J.J., Whittal, R., Worobo, R.W., and Vederas, J.C. 2003. Isolation, structural characterization, and properties of

mattacin (polymyxin M), a cyclic peptide antibiotic produced by *Paenibacillus kobensis* M. J. Biol. Chem. **278**(15): 13124-13132.

Mathur, J., and Waldor, M.K. 2004. The *Vibrio cholerae* ToxR-regulated porin OmpU confers resistance to antimicrobial peptides. Infect. Immun. **72**(6): 3577-3583.

May, J.J., Kessler, N., Marahiel, M.A., and Stubbs, M.T. 2002. Crystal structure of DhbE, an archetype for aryl acid activating domains of modular nonribosomal peptide synthetases. Proc. Natl. Acad. Sci. U. S. A. **99**(19): 12120-12125.

Menges, R., Muth, G., Wohlleben, W., and Stegmann, E. 2007. The ABC transporter Tba of *Amycolatopsis balhimycina* is required for efficient export of the glycopeptide antibiotic balhimycin. Appl. Microbiol. Biotechnol. **77**(1): 125-134.

Mitaku, S., and Hirokawa, T. 1999. Physicochemical factors for discriminating between soluble and membrane proteins: hydrophobicity of helical segments and protein length. Protein Eng. **12**(11): 953-957.

Mitaku, S., Hirokawa, T., and Tsuji, T. 2002. Amphiphilicity index of polar amino acids as an aid in the characterization of amino acid preference at membrane-water interfaces. Bioinformatics **18**(4): 608-616.

Mootz, H.D., and Marahiel, M.A. 1997. Biosynthetic systems for nonribosomal peptide antibiotic assembly. Curr. Opin. Chem. Biol. **1**(4): 543-551.

Mootz, H.D., Schwarzer, D., and Marahiel, M.A. 2002. Ways of assembling complex natural products on modular nonribosomal peptide synthetases. Chembiochem **3**(6): 490-504.

Mootz, H.D., Schwarzer, D., and Marahiel, M.A. 2000. Construction of hybrid peptide synthetases by module and domain fusions. Proc. Natl. Acad. Sci. U. S. A. **97**(11): 5848-5853.

Moyne, A.L., Cleveland, T.E., and Tuzun, S. 2004. Molecular characterization and analysis of the operon encoding the antifungal lipopeptide bacillomycin D. FEMS Microbiol. Lett. **234**(1): 43-49.

Nolan, E.M., and Walsh, C.T. 2009. How nature morphs peptide scaffolds into antibiotics. Chembiochem **10**(1): 34-53.

Orwa, J.A., Govaerts, C., Busson, R., Roets, E., Van Schepdael, A., and Hoogmartens, J. 2001. Isolation and structural characterization of polymyxin B components. J. Chromatogr. A **912**(2): 369-373.

Ouderkirk, J.P., Nord, J.A., Turett, G.S., and Kislak, J.W. 2003. Polymyxin B nephrotoxicity and efficacy against nosocomial infections caused by multiresistant Gram-negative bacteria. Antimicrob. Agents Chemother. **47**(8): 2659-2662.

Parker, W.L., and Rathnum, M.L. 1975. EM49, a new peptide antibiotic IV. The structure of EM49. J. Antibiot. (Tokyo) **28**(5): 379-389.

Patel, G., Perez, F., and Bonomo, R.A. 2010. Carbapenem-resistant Enterobacteriaceae and *Acinetobacter baumannii*: assessing their impact on organ transplantation. Curr. Opin. Organ. Transplant. **15**(6): 676-682.

Paulus, H., and Gray, E. 1964. The biosynthesis of polymyxin B by growing cultures of *Bacillus polymyxa*. J. Biol. Chem. **239**: 865-871.

Pavela-Vrancic, M., Dieckmann, R., Dohren, H.V., and Kleinkauf, H. 1999. Editing of non-cognate aminoacyl adenylates by peptide synthetases. Biochem. J. **342 Pt 3**: 715-719.

Pavela-Vrancic, M., Van Liempt, H., Pfeifer, E., Freist, W., and Von Dohren, H. 1994. Nucleotide binding by multienzyme peptide synthetases. Eur. J. Biochem. **220**(2): 535-542.

Pearson, L.A., Hisbergues, M., Borner, T., Dittmann, E., and Neilan, B.A. 2004. Inactivation of an ABC transporter gene, mcyH, results in loss of microcystin production in the cyanobacterium *Microcystis aeruginosa* PCC 7806. Appl. Environ. Microbiol. **70**(11): 6370-6378.

Perez, J.C., and Groisman, E.A. 2007. Acid pH activation of the PmrA/PmrB two-component regulatory system of *Salmonella enterica*. Mol. Microbiol. **63**(1): 283-293.

Price, N.P., Rooney, A.P., Swezey, J.L., Perry, E., and Cohan, F.M. 2007. Mass spectrometric analysis of lipopeptides from *Bacillus* strains isolated from diverse geographical locations. FEMS Microbiol. Lett. **271**(1): 83-89.

Rachoin, J.S., Schorr, C.A., and Dellinger, R.P. 2010. Targeting endotoxin in the treatment of sepsis. Subcell. Biochem. **53**: 323-338.

Rausch, C., Weber, T., Kohlbacher, O., Wohlleben, W., and Huson, D.H. 2005. Specificity prediction of adenylation domains in nonribosomal peptide

synthetases (NRPS) using transductive support vector machines (TSVMs). Nucleic Acids Res. **33**(18): 5799-5808.

Roach, P.L., Clifton, I.J., Hensgens, C.M., Shibata, N., Schofield, C.J., Hajdu, J., and Baldwin, J.E. 1997. Structure of isopenicillin N synthase complexed with substrate and the mechanism of penicillin formation. Nature **387**(6635): 827-830.

Rodnina, M.V., and Wintermeyer, W. 2001. Ribosome fidelity: tRNA discrimination, proofreading and induced fit. Trends Biochem. Sci. **26**(2): 124-130.

Roongsawang, N., Hase, K., Haruki, M., Imanaka, T., Morikawa, M., and Kanaya, S. 2003. Cloning and characterization of the gene cluster encoding arthrofactin synthetase from *Pseudomonas* sp. MIS38. Chem. Biol. **10**(9): 869-880.

Rounge, T., Rohrlack, T., Kristensen, T., and Jakobsen, K. 2008. Recombination and selectional forces in cyanopeptolin NRPS operons from highly similar, but geographically remote *Planktothrix* strains. BMC Microbiology **8**(1): 141.

Ruttenberg, M.A., and Mach, B. 1966. Studies on amino acid substitution in the biosynthesis of the antibiotic polypeptide tyrocidine. Biochemistry **5**(9): 2864-2869.

Samel, S.A., Wagner, B., Marahiel, M.A., and Essen, L.O. 2006. The thioesterase domain of the fengycin biosynthesis cluster: a structural base for the macrocyclization of a non-ribosomal lipopeptide. J. Mol. Biol. **359**(4): 876-889.

Saurin, W., Hofnung, M., and Dassa, E. 1999. Getting in or out: early segregation between importers and exporters in the evolution of ATP-binding cassette (ABC) transporters. J. Mol. Evol. **48**(1): 22-41.

Schindler, M., and Osborn, M.J. 1979. Interaction of divalent cations and polymyxin B with lipopolysaccharide. Biochemistry **18**(20): 4425-4430.

Schneider, A., and Marahiel, M.A. 1998. Genetic evidence for a role of thioesterase domains, integrated in or associated with peptide synthetases, in non-ribosomal peptide biosynthesis in *Bacillus subtilis*. Arch. Microbiol. **169**(5): 404-410.

Schneider, A., Stachelhaus, T., and Marahiel, M.A. 1998. Targeted alteration of the substrate specificity of peptide synthetases by rational module swapping. Mol. Gen. Genet. **257**(3): 308-318.

Schneider, T.L., Shen, B., and Walsh, C.T. 2003. Oxidase domains in epothilone and bleomycin biosynthesis: thiazoline to thiazole oxidation during chain elongation. Biochemistry **42**(32): 9722-9730.

Schwarzer, D., Finking, R., and Marahiel, M.A. 2003. Nonribosomal peptides: from genes to products. Nat. Prod. Rep. **20**(3): 275-287.

Schwarzer, D., Mootz, H.D., Linne, U., and Marahiel, M.A. 2002. Regeneration of misprimed nonribosomal peptide synthetases by type II thioesterases. Proc. Natl. Acad. Sci. U. S. A. **99**(22): 14083-14088.

Shepherd, J., Sarker, P., Rimmer, S., Swanson, L., MacNeil, S., and Douglas, I. 2011. Hyperbranched poly(NIPAM) polymers modified with antibiotics for the reduction of bacterial burden in infected human tissue engineered skin. Biomaterials **32**(1): 258-267.

Stachelhaus, T., and Walsh, C.T. 2000. Mutational analysis of the epimerization domain in the initiation module PheATE of gramicidin S synthetase. Biochemistry **39**(19): 5775-5787.

Stachelhaus, T., Mootz, H.D., and Marahiel, M.A. 1999. The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. Chem. Biol. **6**(8): 493-505.

Stachelhaus, T., Huser, A., and Marahiel, M.A. 1996a. Biochemical characterization of peptidyl carrier protein (PCP), the thiolation domain of multifunctional peptide synthetases. Chem. Biol. **3**(11): 913-921.

Stachelhaus, T., Schneider, A., and Marahiel, M.A. 1996b. Engineered biosynthesis of peptide antibiotics. Biochem. Pharmacol. **52**(2): 177-186.

Stachelhaus, T., Schneider, A., and Marahiel, M.A. 1995. Rational design of peptide antibiotics by targeted replacement of bacterial and fungal domains. Science **269**(5220): 69-72.

Stachelhaus, T., Mootz, H.D., Bergendahl, V., and Marahiel, M.A. 1998. Peptide bond formation in nonribosomal peptide biosynthesis. Catalytic role of the condensation domain. J. Biol. Chem. **273**(35): 22773-22781.

Stansly, P.G., and Brownlee, G. 1949. Nomenclature of polymyxin antibiotics. Nature **163**(4146): 611.

Stansly, P.G., and Schlosser, M.E. 1947. Studies on polymyxin: Isolation and identification of *Bacillus polymyxa* and differentiation of polymyxin from certain known antibiotics. J. Bacteriol. **54**(5): 549-556.

Staunton, J., and Weissman, K.J. 2001. Polyketide biosynthesis: a millennium review. Nat. Prod. Rep. **18**(4): 380-416.

Stein, T., Vater, J., Kruft, V., Wittmann-Liebold, B., Franke, P., Panico, M., Mc Dowell, R., and Morris, H.R. 1994. Detection of 4'-phosphopantetheine at the thioester binding site for L-valine of gramicidinS synthetase 2. FEBS Lett. **340**(1-2): 39-44.

Stein, T., Vater, J., Kruft, V., Otto, A., Wittmann-Liebold, B., Franke, P., Panico, M., McDowell, R., and Morris, H.R. 1996. The multiple carrier model of nonribosomal peptide biosynthesis at modular multienzymatic templates. J. Biol. Chem. **271**(26): 15428-15435.

Steller, S., Sokoll, A., Wilde, C., Bernhard, F., Franke, P., and Vater, J. 2004. Initiation of surfactin biosynthesis and the role of the SrfD-thioesterase protein. Biochemistry **43**(35): 11331-11343.

Steller, S., Vollenbroich, D., Leenders, F., Stein, T., Conrad, B., Hofemeister, J., Jacques, P., Thonart, P., and Vater, J. 1999. Structural and functional organization of the fengycin synthetase multienzyme system from *Bacillus subtilis* b213 and A1/3. Chem. Biol. **6**(1): 31-41.

Storm, D.R., Rosenthal, K.S., and Swanson, P.E. 1977. Polymyxin and related peptide antibiotics. Annu. Rev. Biochem. **46**: 723-763.

Tillett, D., Dittmann, E., Erhard, M., von Dohren, H., Borner, T., and Neilan, B.A. 2000. Structural organization of microcystin biosynthesis in *Microcystis aeruginosa* PCC7806: an integrated peptide-polyketide synthetase system. Chem. Biol. **7**(10): 753-764.

Trivedi, O.A., Arora, P., Sridharan, V., Tickoo, R., Mohanty, D., and Gokhale, R.S. 2004. Enzymic activation and transfer of fatty acids as acyl-adenylates in *Mycobacteria*. Nature **428**(6981): 441-445.

Tseng, C.C., Bruner, S.D., Kohli, R.M., Marahiel, M.A., Walsh, C.T., and Sieber, S.A. 2002. Characterization of the surfactin synthetase C-terminal thioesterase domain as a cyclic depsipeptide synthase. Biochemistry **41**(45): 13350-13359.

Tsubery, H., Ofek, I., Cohen, S., Eisenstein, M., and Fridkin, M. 2002. Modulation of the hydrophobic domain of polymyxin B nonapeptide: effect on outer-membrane permeabilization and lipopolysaccharide neutralization. Mol. Pharmacol. **62**(5): 1036-1042.

Vaara, M. 2010. Polymyxins and their novel derivatives. Curr. Opin. Microbiol. **13**(5): 574-581.

Vaara, M., and Vaara, T. 2010. Structure–activity studies on novel polymyxin derivatives that carry only three positive charges. Peptides **31**(12): 2318-2321.

Velkov, T., Thompson, P.E., Nation, R.L., and Li, J. 2010. Structure--activity relationships of polymyxin antibiotics. J. Med. Chem. **53**(5): 1898-1916.

Vogler, K., and Studer, R.O. 1966. The chemistry of the polymyxin antibiotics. Experientia **22**(6): 345-354.

von Dohren, H., Keller, U., Vater, J., and Zocher, R. 1997. Multifunctional peptide synthetases. Chem. Rev. **97**(7): 2675-2706.

Walker, J.E., Saraste, M., Runswick, M.J., and Gay, N.J. 1982. Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. EMBO J. **1**(8): 945-951.

Walsh, C.T., Chen, H., Keating, T.A., Hubbard, B.K., Losey, H.C., Luo, L., Marshall, C.G., Miller, D.A., and Patel, H.M. 2001. Tailoring enzymes that modify nonribosomal peptides during and after chain elongation on NRPS assembly lines. Curr. Opin. Chem. Biol. **5**(5): 525-534.

Wilkinson, S., and Lowe, L.A. 1964. Structure of Polymyxin B1. Nature **202**(4938): 1211-1211.

Yeh, E., Kohli, R.M., Bruner, S.D., and Walsh, C.T. 2004. Type II thioesterase restores activity of a NRPS module stalled with an aminoacyl-S-enzyme that cannot be elongated. Chembiochem **5**(9): 1290-1293.

Yin, X., and Zabriskie, T.M. 2006. The enduracidin biosynthetic gene cluster from *Streptomyces fungicidicus*. Microbiology **152**(Pt 10): 2969-2983.

Yonus, H., Neumann, P., Zimmermann, S., May, J.J., Marahiel, M.A., and Stubbs, M.T. 2008. Crystal structure of DltA. Implications for the reaction mechanism of non-ribosomal peptide synthetase adenylation domains. J. Biol. Chem. **283**(47): 32484-32491.

Zavascki, A.P., Goldani, L.Z., Li, J., and Nation, R.L. 2007. Polymyxin B for the treatment of multidrug-resistant pathogens: a critical review. J. Antimicrob. Chemother. **60**(6): 1206-1215.