University of Alberta

Investigating the clavam gene cluster in *Streptomyces antibioticus* Tü1718

by

Sarah Goomeshi Nobary

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

Abstract

Production of clavams (including clavulanic acid, a potent β-lactamase inhibitor) has been studied in *Streptomyces clavuligerus*. All clavams share the early biosynthetic steps. However, many of the late steps of the pathway remain unclear. *Streptomyces antibioticus* Tü1718 was chosen for this study because it produces two 5*S* clavams (different from clavulanic acid in their stereochemistry) and no clavulanic acid. After screening a cosmid library of the gDNA of this bacterium for a *cas* gene (encoding an early biosynthetic enzymes), a cosmid was selected. After finding additional early genes (*bls, pah, ceaS, oat*) and some putative late pathway genes (an aminotransferase and ORFs similar to *c7p, cvm1* and *orfA* of *S. clavuligerus*) in the regions flanking *cas,* the whole cosmid was sequenced. Several novel genes were encountered including *ntr* that showed similarity to a nitroreductase. Knock out mutation of the *cas* and *ntr* gene abolished clavam production in *S. antibioticus* Tü1718.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

A	Adenine
ACV	LLD-α-aminoadipyl-cysteinyl-valine
adh	Alcohol dehydrogenase
Adp	A-factor dependent protein
akr	Aldo-keto reductase
ARE	Autoregulatory element
Arp	A-factor receptor protein
ATP	Adenosine triphosphate
atr	Aminotransferase
bld	Bald
bls	β-lactam synthase
BSA	Bovine serum albomine
cad	Clavulanic acid dehydrogenase
cas	Clavaminate synthase
ccaR	Cephamycin and clavulanic acid regulator
CEA	N^2-(2-carboxyethyl)-arginine
ceaS	Carboxylarginine synthase
cfu	Colony forming unit
claR	Clavulanic acid regulator
ctr	Clavam transcriptional regulator
cvm	Clavam
cyp	Cytochrome P
c6p	Clavam 6 paralogue
c7p	Clavam 7 paralogue
ddl	D-alanyl-D-alanine ligase
dGTP	Deoxyriboguanoside triphosphate
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleoside triphosphate
DGPC	Deoxyguanidinoproclavaminic acid
EDTA	Ethylenediaminetetraacetic acid
fd	Ferrodoxin
FRT	FLP recombination target
gDNA	Genomic deoxyribonucleic acid
HPLC	High performance liquid chromatography
IPNS	Isopenicillin N synthase
IPTG	Isopropyl-β-D-thiogalactopyranoside
ISP2	International *Streptomyces* project

Kbp	Kilo base pair
LB	Lenox broth
LC-MS	Liquid chromatography-Mass spectrometry
Mb	Megabase
MIC	Minimum inhibitory concentration
ntr	Nitroreductase
oat	Ornithine acetyl transferase
OD	Optical density
oppA	Oligopeptide transport protein
ORF	Open reading frame
pah	Proclavaminate amidinohydrolase
PCR	Polymerase chain reaction
ppGppe	Guanosine tetraphosphat
psr	Pathway specific regulator
RNA	Rbinucleic acid
rpm	Rotation per minute
rRNA	Ribosomal rbinucleic acid
SM	Sodium-Manganese
SARP	Streptomyces antibiotic regulatory protein
SAP	Shrimp alkaline phosphatase
SDS	Sodium dodecyl sulfate
SET	Sodium-EDTA-Tris
snk	Sensor kinase
SOB	Super optimal broth
SOC	Super optimal catabolite repression
soh	Serine hydroxymethyltransferase
SSC	Saline-sodium citrate
TAE	Tris-Acetate-EDTA
TE	Tris-EDTA
tRNA	Transfer ribonucleic acid
TSB	Tripticase soy broth
TSB-S	Tripticase soy broth supplemented with starch
UV	Ultra violet
UV/vis	Ultra violet/visible
vol.	Volume
v/v	Volume/volume
w/v	Weight/volume
X-gal	5-bromo-4-chloro-3-indoyl-β-D-galactopyranoside

I. Introduction

I.1 Antibiotics: a brief introduction

Antibiotics (*anti*-= against, *bios*-= life) are chemical substances that are produced by certain microorganisms. These substances inhibit the growth and development of other microbes. Unlike the more generic term of antimicrobial agents, which could be chemically synthesized, antibiotics are natural products of microbial activity (Madigan et al. 2000; Pyatkin 1967).

Antibiotics are divided into different groups depending on their chemical structures and their modes of action. One of the most important groups of antibiotics, both historically and medically, is the β-lactam group. This group includes penicillins, cephalosporins, cephamycins, clavams and carbapenams, all medically important antibiotics (Demain and Elander 1999). These antibiotics all share the presence of a characteristic structural component, the β-lactam ring (Figure I.1). The penicillins and cephalosporins together account for over one half of all of the antibiotics produced and used world-wide (Madigan et al. 2000).

The first β-lactam antibiotic that was discovered was penicillin. This antibiotic was found by Fleming in 1929 in broth filtrates from the fungus *Penicillium*. Fleming showed that this broth filtrate caused the lysis of *Staphylococcus* colonies (Fleming 1929). The first report of a bacterium (from the *Streptomyces* family) able to produce β-lactams was made by Miller et al. in 1962 (Miller et al. 1962). β-Lactam antibiotics inhibit cell wall synthesis by deactivating the transpeptidases (also known as penicillin binding proteins, PBP)

responsible for cross-linking of two glycan linked peptide chains. The β-lactam antibiotics bind to the transpeptidase as an analog of acyl-D-alanyl-D-alanine. Bacteria that are unable to synthesize their cell wall can no longer reproduce. Because Gram positive bacteria lack the lipopolysaccharide layer that the Gram negatives possess, they are more susceptible to these antibiotics (Tipper and Strominger 1965; Wise and Park 1965).

I.2 Antibiotic resistance and importance of a β-lactamase inhibitor

Due to widespread use of penicillins, many species of pathogenic microorganisms have become resistant to these antibiotics. This can be either an acquired ability, or a characteristic of the bacterium. First, it should be noted that in terms of evolution, most resistance genes might have been acquired through a process of genetic exchange from the antibiotic producers. In order to protect themselves from the antibiotics they produce, the antibiotic producing organisms have developed mechanisms to neutralize or destroy their own antibiotics. So, in the right conditions these resistance genes can be transferred to other organisms (Madigan et al. 2000).

One of the mechanisms of resistance to β-lactam antibiotics is by the microorganism acquiring the ability to alter the antibiotic to an inactive form by the enzyme β-lactamase (Baggaley et al. 1997). This enzyme breaks the β-lactam ring and therefore deactivates the antibiotic. A scheme of this reaction can be seen in Figure I.2. Deactivating the β-lactamases is an effective way of fighting

resistant bacteria. One example of a product that deactivates a number of β-lactamases is clavulanic acid. This β-lactam antibiotic is an important β-lactamase inhibitor and is used with other drugs to eradicate infections caused by resistant bacteria.

I.3 The structure of clavulanic acid and other clavams

Clavulanic acid belongs to the clavam group of antibiotics, which is one of the groups in the β-lactam family. The nucleus of clavulanic acid is similar to penicillins. However, the oxazolidine ring (instead of the thiazolidine group in penicillins) contains oxygen instead of sulfur. In addition, this bicyclic clavam nucleus lacks the C-6 acylamino side chain of penicillins (Baggaley et al. 1997, Jensen and Paradkar 1999). Figure I.3 shows the structure of penicillin and the other clavams.

Other clavams differ from clavulanic acid in that they lack the carboxyl group at C-3 and have a 5*S* stereochemistry. These other clavams also have different side chains on their C-2 carbon (Liras and Rodriguez-Garcia 2000). In contrast, clavulanic acid has a 5*R* 3*R* stereochemistry, which is unique among the clavams. Because of their 5*S* stereochemistry, the other clavams are also called 5*S* clavams and will be referred to in that way from here on (Baggaley et al. 1997, Jensen and Paradkar 1999). Figure I.3 shows the structure of clavulanic acid and some of the 5*S* clavams.

Clavamycins are the other members of the 5*S* clavam family. These antibiotics have larger and more complex side chains on their C-2. The structure

of these metabolites is shown in Figure I.4, in which R shows the different side chains that can be present at C-2.

I.4 Activity of clavulanic acid and the 5*S* clavams

As mentioned earlier clavulanic acid is a potent β-lactamase inhibitor. This β-lactam antibiotic also possesses weak bacteriocidal effects, with the minimum inhibitory concentration (MIC) ranging from 25-125μg/ml (Reading and Cole 1977). Therefore, clavulanic acid is used in combination with other penicillins such as amoxicillin and ticarcillin to protect these antibiotics from hydrolysis. The combined drugs are called Augmentin™ and Timentin™, respectively.

Clavulanic acid can inhibit most of the class A β-lactamases (groups 2a, 2b′, 2b, 2c and 2e of the catalytic-profile based classification of Bush) and the cloxacillin-hydrolyzing enzymes of class D (group 2d). Besides activity against these enzymes, clavulanic acid has poor activity against cephalosporinases of class C (group 1) and is inactive against Zn^{2+} metalloenzymes of class B (group 3) or the chromosomal penicillinases of group 4 (Bush 1989a and 1989b).

The 5*S* clavams exert antibacterial or antifungal effects (e.g. valclavam and alanylclavam). The antibacterial activity is due to the non-competitive inhibition of homoserine O-succinyl transferase, an enzyme required for methionine biosynthesis (Röhl et al. 1987). As a result the clavams act as methionine antimetabolites. Clavamycins lack antibacterial activity and were discovered for their activity against *Candida* (King et al. 1986). Only clavulanic acid inhibits β-

lactamase, which is because of its unique 3*R*, 5*R* stereochemistry (Baggaley et al. 1997).

Clavulanic acid forms an irreversible complex with the β-lactamases at their active centers (with the serine hydroxyl group at the active center), producing acylated intermediates. This results in the inactivation of the β-lactamase enzyme. The mechanism of this inactivation is detailed below and shown in Figure I.5.

From the acyl intermediate, three simultaneous reactions occur. One reaction releases free enzyme and hydrolytic products, resembling the case where a normal β-lactam substrate is reacted with the β-lactamase. The second reaction is the formation of the irreversibly inactivated enzyme complex. The irreversibly inactivated enzyme complex results from further covalent interaction between the ring opened clavulanic acid and the enzyme. The third reaction is the formation of transiently inhibited complex. This is formed at a faster rate than the irreversibly inactivated enzyme complex. The transiently stable complex slowly releases free enzyme and hydrolytic products as a result of being in equilibrium with the initial acyl intermediate. Some of the enzyme will be irreversibly inactivated, but the majority of the enzyme is released in active form and reacts with fresh clavulanic acid to repeat the cycle. With each cycle, more enzyme is irreversibly inactivated until all the enzyme activity has been eliminated (Butterworth 1984; Sykes and Bush 1982; Reading and Cole 1977).

I.5 Clavam producing bacteria

Clavulanic acid is the most important member of the clavam family. Only *Streptomyces* species are known to produce clavulanic acid. Some of these are *S. jumonjinensis, S. katsurahamanus, Streptomyces* sp., and *S. clavuligerus* (Box 1977; Kitano et al. 1979; Tahlan et al. 2007). Among these, *S. clavuligerus* is the microorganism used for the production of clavulanic acid in industry. In addition to clavulanic acid, *S. clavuligerus* is able to produce four 5*S* clavams and holomycin and tunicamycin (Kenig and Reading 1979). *Streptomyces antibioticus* Tü1718 is another microorganism able to produce clavam metabolites. This bacterium produces valclavam and hydroxyethylclavam, however it cannot produce clavulanic acid (Röhl et al. 1987).

S. clavuligerus was isolated by Higgens and Kastner in 1971 from a South American soil sample (Higens and Kastner 1971). The name 'clavuligerus' consists of two part: 'clavula', which means 'little clubs', and '-igerus', which means 'bearing' (Stanley et al. 1983). *S. clavuligerus* produces four different types of β-lactam antibiotics as secondary metabolites: clavulanic acid, 5*S* clavams (clavam-2-carboxylate, 2-hydroxymethylclavam, 2-formyloxymethylclavam, and alanylclavam), cephamycin C [7-(5-amino-5-carboxyvaleramido)-3-carboxyloxymethyl-3-cepham-4-carboxylic acid] and penicillin N (Brown et al. 1976; Nagarajan 1972).

I.6 *Streptomyces* and their life cycle

Streptomyces possesses large, high G+C content (about 70%) chromosomes (Kieser et al. 1992 and 2000; Leblond et al. 1990 and 1993; Lezhava et al. 1995). Their chromosomes are linear and both linear and circular plasmids can be found in these microorganisms (Hayakawa et al. 1978; Lin et al. 1993; Schrempf et al. 1975). The whole genome of some of the *Streptomyces* spp. has been sequenced, however most analyses have focused on just two species: *S. coelicolor* and *S. avermitilis.* They both have large chromosomes (about 9.0 Mb), containing about 7500-7800 genes (Ikeda et al. 2003).

The genus *Streptomyces* is a part of the family *Streptomycetaceae.* This family is placed in the *Actinomycetales* order in the *Actinobacteria* class (Stackerbrandt et al. 2003; Waksman and Henrici 1943). *Streptomyces* have a complex life cycle involving the formation of aerial mycelia from the substrate mycelia, and then the development of spores. The start of secondary metabolism is believed to coincide with the onset of aerial mycelia formation (Chater 1989).

As shown in Figure I.6, the spores germinate, forming a germ tube, and then grow until they form a network of multinucleate hyphae which are branching (Hardisson et al. 1978). Since it grows on the surface (and into the growth medium), this mycelium is called substrate mycelium. The substrate mycelium ages and as it uses up the nutrients this resulting nutrient depletion leads to the formation of the aerial mycelium, which grows upwards. It grows from the inside of the branching area (Miguelez et al. 1994). Because the aerial mycelia are far

from the nutrients in the growth medium, they get their nutrients from the substrate mycelia that are being lysed (Mendez et al. 1985). Cross walls are formed between individual nuclei (Wildermuth 1970). The mycelia become more differentiated to form spore chains, which can give rise to more *Streptomyces* colonies (Hardisson and Manzanal 1976).

I.7 Biosynthesis of clavams

I.7.1 Precursors

The biosynthetic pathway of clavulanic acid and the 5*S* clavams is shared in the early stages (Egan et al. 1997). Therefore these antibiotics have the same precursors. Due to the lack of recognizable pathway intermediates analogous to δ-L-α-(aminoadipyl)-L-cysteinyl-D-valine (the peptide precursor of the penicillin/cephamycin biosynthetic pathway), which would support *in vitro* biosynthesis of clavulanic acid, the biosynthetic pathway remained unclear for many years (Jensen and Paradkar 1999). Before 1999, it was believed that the direct precursors of clavulanic acid biosynthesis were ornithine and pyruvate (Pitlik and Townsend 1997; Thirkettle et al. 1997), but then Khaleeli et al. (1999) showed that glyceraldehyde-3-phosphate is used as a substrate for the enzyme responsible for condensation with arginine to give the C3 part (carbons C-5 to C-7) of clavulanic acid. It is now believed that glyceraldehyde-3-phosphate is the direct precursor of the C3 part, rather than pyruvate (Khaleeli et al 1999).

After finding a gene encoding an arginase homologue (proclavaminate aminohydrolase, Pah) in the gene cluster of clavulanic acid, it was initially

suggested that ornithine is the direct precursor of the C5 part of clavulanic acid (carbons C-2, C-3, and C-8 to C-10) (Romero et al. 1986). However, in the research done by Valentine et al. (1993) using mutants blocked in *arg*G and *arg*F (unable to convert ornithine to arginine), the efficient incorporation of arginine was shown, whereas, the incorporation of ornithine was weak. This showed that arginine is the direct precursor of clavulanic acid (Valentine et al. 1993). It is now believed that arginine is the precursor for the C5 part of clavulanic acid (De la Fuente et al. 1996; Romero et al. 1986; Townsend and Ho 1985)

I.7.2 Gene clusters involved in the biosynthesis of clavulanic acid and 5S clavams

In most bacteria genes required for the biosynthesis of a specific secondary metabolite are located in one gene cluster. Surprisingly in *S. clavuligerus,* the genes for producing clavulanic acid and the 5*S* clavams are located on three distant clusters. These clusters are known as the clavam cluster, the paralogue cluster and the cephamycin-clavulanic acid supercluster (Jensen et al. 2000; Jensen et al. 2004a and 2004b; Li et al. 2000; Mellado et al. 2002; Mosher et al. 1999; Tahlan et al. 2004a and 2004b). The first evidence showing that separate gene clusters are involved was found when *cas1* and *cas2* were isolated. These genes encode two separate clavaminate synthase (Cas) enzymes, and Salowe et al. (1990) hypothesized that one of these enzymes is involved in clavulanic acid production, while the other one plays a role in 5*S* clavam production (Marsh et al. 1992).

Genes responsible for the biosynthesis of clavulanic acid and cephamycin C form a supercluster (Aidoo et al. 1994; Hodgson et al. 1995; Ward and Hodgson 1993). Given that clavulanic acid and cephamycin C are structurally related and their biosynthesis is regulated by a shared mechanism, it is perhaps not too surprising that their genes would be clustered together. The regulator molecule, CcaR, is encoded by a gene located in the cephamycin part of the cephamycin-clavulanic acid supercluster and positively regulates the biosynthesis of cephamycin C and clavulanic acid (Alexander and Jensen 1998; Perez-Llarena et al. 1997; Perez-Redondo et al. 1998).

Although 5*S* clavams and clavulanic acid have a shared biosynthetic pathway to a certain point (Egan et al. 1997) the clavam and clavulanic acid clusters do not form one super cluster and instead are on different parts of the chromosome, and the paralogue cluster, which contains paralogues for the early genes and some genes that are specifically required for production of 5*S* clavams, lies on a giant linear plasmid (Medema et al. 2010; Tahlan et al. 2004a). This fact suggests that the 5*S* clavams and clavulanic acid have different regulatory mechanisms. The genes present in these clusters are shown in Figure I.7.

I.7.3 The biosynthetic pathway and the genes involved

The biosynthetic pathway of clavulanic acid and the 5*S* clavams in *S. clavuligerus* can be divided to two stages: Early and Late. The early stages are the steps that are the same for clavulanic acid and the 5*S* clavams. The early stages of clavam production have been well established. The late steps however are yet to

be determined completely. The early steps of clavam biosynthesis are shown in Figure I.8.

I.7.3.1 Early steps:

Early intermediates in clavulanic acid production were first discovered in a study where metabolites containing guanidino groups were purified from the culture filtrate of *S. clavuligerus* mutants blocked in clavulanic acid production. These metabolites were recognized by their ability to participate in the Sakaguchi color reaction (Elson et al. 1993).

The first discovered guanidino intermediate is N^2-(2-carboxyethyl)-arginine (CEA). The enzyme required for the formation of this structure is carboxyethylarginine synthase (CeaS). This enzyme requires thiamine pyrophosphate and Mg^{2+} for its action (Jensen et al. 2000; Perez-Redondo 1998).

The second guanidino intermediate is deoxyguanidinoproclavaminic acid (DGPC). This is a monocyclic β-lactam compound that arises from CEA via closure of the β-lactam ring. The enzyme responsible for the formation of the β-lactam ring, is known as β-lactam synthetase (Bls), and was described by Bachmann et al. (Bachmann et al. 1998). This enzyme forms an intramolecular amide bond in CEA in the presence of Mg^{2+} and ATP (Bachmann et al. 1998).

Until the discovery of Bls, isopenicillin N synthase (IPNS) was the only purified enzyme known to form a β-lactam ring. Isopenicillin N synthase forms isopenicillin N from the LLD-α-aminoadipyl-cysteinyl-valine (ACV) tripeptide in the penicillin and cephalosporin biosynthesis pathway. Although both Bls and IPNS are responsible for the formation of a β-lactam ring, their mechanisms are

quite different. Bls requires ATP and is similar to asparagine synthetase (a primary metabolism amino transferase), while IPNS is a non-heme dioxygenase and is more closely related to clavaminate synthase. No clear ATP-binding motifs are found in IPNS and this enzyme also shows no similarity in size or domain structure to the peptide synthesases (Bachmann et al. 1998; Jensen et al. 2000).

The next reaction in the pathway is the hydroxylation of DGPC to guanidinoproclavaminate. This step is catalyzed by clavaminate synthase 2 (Cas2) (Baldwin et al. 1993). Cas2 is encoded by *cas2,* which is located in the clavulanic acid gene cluster. Cas2 requires α-ketoglutarate, O_2, and Fe^{2+} and disrupting the *cas2* gene results in decreased clavulanic acid production (Salowe et al. 1990). Like Cas2, IPNS is also Fe (II) dependent but it does not require α-ketoglutarate (Roach et al. 1995). An outstanding feature of these kinds of enzymes is their ability to catalyze more than one reaction. The advantage that these enzymes have compared to the heme-based oxygenases is that they provide a more flexible environment for the catalysis and evolution of new reactions, i. e. they are able to catalyze more than one reaction (Kershaw et al. 2005).

The third guanidino metabolite, guanidinoproclavaminic acid, has a hydroxyl group. Its oxygen atom is derived from molecular oxygen. Proclavaminate amidinohydrolase (Pah) removes the guanidino group from guanidinoproclavaminate through hydrolysis to give proclavaminate (Aidoo et al. 1994; Elson et al. 1993; Wu et al. 1995). The reaction catalyzed by this enzyme was suggested by analyzing the sequence of the gene corresponding to the

enzyme (Aidoo et al. 1994; Elson et al. 1993). Pah needs Mn^{2+} for its action and is encoded by *pah* (Aidoo et al. 1994; Elkins et al. 2002).

Proclavaminic acid is converted to dihydroclavaminic acid by the action of Cas. Dihydroclavaminic acid then undergoes another reaction also catalyzed by Cas and is converted into clavaminic acid. Baldwin et al. 1993; Salowe et al. 1991). Interestingly the presence of an enzyme with the same activity as Cas is found in *S. antibioticus* Tü1718. This enzyme, Cas3, is very similar to Cas1. Considering the ability of *S. antibioticus* Tü1718 to produce clavam metabolites, the presence of a Cas enzyme would be expected. Because the gene encoding Cas1 (*cas1*) is suggested to be related to 5*S* clavam production rather than clavulanic acid production, and because *S. antibioticus* Tü1718 produces two 5*S* clavams but not clavulanic acid, it follows that Cas3 might resemble Cas1 more closely than Cas2 (Janc et al. 1995).

Clavaminic acid is believed to be the branch point between clavulanic acid and the 5*S* clavams. The structure of clavaminic acid comprises the fused bicyclic β-lactam/oxazolidine structure, but with a stereochemistry different from clavulanic acid, and the same as the 5*S* clavams. Another similarity of clavaminic acid with the 5*S* clavams is that clavaminic acid is not a β-lactamase inhibitor. Janc et al. (1993), following the observation of incorporation of labeled proclavaminic acid into the clavams, initially suggested that the biosynthetic pathway leading to clavulanic acid and other clavams was the same at least to the stage of proclavaminic acid (Janc et al. 1993). Later, Egan et al. (1997) in their studies showed that clavaminic acid is a shared intermediate between clavulanic

acid and the 5*S* clavams. In this study, the incorporation of labeled clavaminic acid into clavulanic acid and 5*S* clavams was observed. The conclusion that can be made from this observation is that clavaminic acid is incorporated into both types of end products (Egan et al. 1997). *S. clavuligerus* mutants blocked between clavaminic acid and clavulanic acid accumulate clavaminic acid. This accumulated compound will then be N-acylated to produce several other 5*S* clavam structures. The 5*S* clavams produced are then exported into the medium (Elson et al. 1988).

The gene *oat* (in the clavulanic acid gene cluster) is considered one of the early genes because it is expressed in the same pattern as the other early genes but the exact role of this gene in the biosynthetic pathway is not known. It has been shown that an insertional mutation in *oat* results in decreased clavam production (Jensen et al. 2000). The enzyme encoded by this gene (Oat) is an ornithine acetyltransferase, which transfers an acetyl group from N-acetylornithine to glutamate during arginine biosynthesis. Oat undergoes post-translational proteolysis and produces β and α fragments. This reaction happens in other ornithine acetyltransferases as well. The two fragments associate and produce an $\alpha_2\beta_2$ heterotetramer. Kershaw et al. have suggested that Oat increases the levels of arginine in the cell for clavam biosynthesis (Kershaw et al. 2002).

I.7.3.2 The late steps for clavulanic acid production

The steps involved in clavulanic acid production after the formation of clavaminic acid are not certain. However, recent studies on the sequences of the

gene clusters have provided some hints about what the later steps in clavulanic acid and 5*S* clavam production would be.

Clavaminic acid is converted to N-glycylclavaminic acid by the action of ORF17 (Jensen et al. 2000). Mutation in *orf17* causes the elimination of clavulanic acid production (Jensen et al. 2004a). ORF17 shows similarity to a wide range of enzymes, which catalyze carboxylate-amine ligation reactions using ATP and glycine. This is done in the presence of Mg^{2+} and K^{+} (Jensen et al. 2000).

Another gene in the clavulanic acid cluster that is considered a late gene is *oppA1*. The predicted product of *oppA1* is 47% similar to the predicted product of *oppA2* (also in the clavulanic acid gene cluster) (Hodgson et al. 1995; Jensen et al. 2000, Mellado et al. 2002). Based on sequence similarity to oligopeptide permeases, both genes are believed to encode oligopeptide permeases (Mellado et al. 2002). Mutation in these genes causes no clavulanic acid production. In addition, mutation in *oppA2* results in the loss of alanylclavam (Lorenzana et al. 2004).

To form clavulanic acid from clavaminic acid, two changes in the molecule are required: the inversion of stereochemistry of the ring system and the conversion of the end group on the C-2 side chain from an amino to a hydroxyl group. The oxygen on the C-9 of clavulanic acid is derived from molecular oxygen. This was shown by the incorporation of labeled di-oxygen on the C-9 (Townsend and Krol 1988). Therefore, it is suggested that the C-9 amino group of clavaminic acid is removed by an oxidative cleavage reaction to yield

clavaldehyde (Liras and Rodriguez-Garcia 2000). The existence of an aldehyde intermediate, clavaldehyde, with the same stereochemistry as clavulanic acid, and the presence of an NADPH-dependent dehydrogenase known as clavulanic acid dehydrogenase (Cad, which reduces clavaldehyde to clavulanic acid) have been documented (Liras and Rodriguez-Garcia 2000; Nicholson et al. 1994). Cad is the enzyme known to catalyze the last step in clavulanic acid production and has a tetrameric structure (Jensen et al. 2000). This enzyme is produced by *cad,* a gene in the clavulanic acid gene cluster (Jensen et al. 2000; Perez-Redondo 1998).

Clavaldehyde is an unstable compound, possessing β-lactamase inhibitory activity. This supports the theory that β-lactamase inhibitory activity is associated with the 3*R*, 5*R* stereochemistry of the ring system in clavulanic acid (Jensen and Paradkar 1999). The change of stereochemistry found in clavaldehyde is proposed to have a chemical origin (Nicholson et al. 1994) however the presence of an epimerase that would catalyze the stereochemistry inversion cannot be ruled out (Elson et al. 1988).

Two other genes in the clavulanic acid gene cluster are *cyp* and *fd.* Mutants defective in these genes were blocked in clavulanic acid production and had lower levels of 5*S* clavam production (Jensen et al. 2004a). The protein encoded by *cyp* has a high similarity to cytochrome P-450s (Jensen et al. 2000; Li et al. 2000). The product of *fd* appears to be a ferredoxin protein (Li et al. 2000). It has been postulated that the complex of Cyp and Fd is responsible for the enantiomerization and change in stereochemistry in clavulanic acid (from 3*S*, 5*S* to 3*R*, 5*R*), but no experimental verification has yet been achieved.

Another one of the late genes in the clavulanic acid cluster is *orf13*. Mutation in *orf13* resulted in decreased clavulanic acid and 5*S* clavam production. The product of this gene is similar to efflux pumps in other bacteria. It might be possible that this protein transfers the clavams outside the cell (Jensen et al. 2004a; Mellado et al. 2002).

The product of *orf16,* another gene in clavulanic acid cluster, has a conserved amino acid sequence required for the attachment of proteins to the cell wall, but otherwise shows only limited similarity to conserved hypothetical proteins from other species. Mutants in this gene had decreased clavulanic acid and they produced two new clavams: N-acetylglycylclavaminic acid and N-acetylclavaminic acid (in trace amounts) (Jensen et al. 2004a; Mellado et al. 2002).

Recently three other ORFs have been identified further downstream of the clavulanic acid gene cluster. These are *orf21-23.* The product of *orf21* showed high similarity to the sigma factor SigL from *Streptomyces avermitilis* MA-4680 and *S. ceolicolor* A3(2) (Song et al. 2009). More than 60 sigma factors have been identified in *Streptomyces* spp. This reflects their complex mechanisms to respond to changes in the environment (Bentley et al. 2002, Hutchings et al. 2004). Mutation of *orf21* did not have any effect on growth and antibiotic production in *S. clavuligerus,* meaning this ORF is not necessary for the late steps of clavulanic acid production (Song et al. 2009).

orf22 and *orf23* appear to be parts of a two component regulatory system (Song et al. 2009). *orf22* showed similarity to the histidine kinase from

Salinospora arenicola CNS-205. *orf23* is similar to a transcriptional regulator from *Kineococcus radiotolerans,* and it also shows high similarity to a response regulator from *S. coelicolor* A3(2). Regulation of antibiotic production by two component systems has been observed, but mutation of *orf22* did not affect the growth and antibiotic production of *S. clavuligerus*. However mutants lacking *orf23* had deficient growth and antibiotic production. It is possible that Orf23 can be phosphorylated by other sensor kinases and maintain the wild type phenotype even in an *orf22* mutant (Song et al. 2009).

I.7.3.3 Late steps in the biosynthesis of the 5*S* clavams

There are even more unknown steps in the late stages of 5*S* clavam biosynthesis in *S. clavuligerus*. As stated above, clavaminic acid is the last common intermediate in the biosynthesis of clavulanic acid and the 5*S* clavams. Egan et al. (1997) have suggested that formation of the 5*S* clavam end products begins with conversion of clavaminic acid to an aldehyde intermediate by decarboxylation at C-3 and deamination of the C-2 side chain. They have hypothesized that these reactions also involve a pyridoxal phosphate cofactor. They further propose that oxidation of this aldehyde intermediate would yield 2-formyloxymethylclavam, and hydrolysis of the formyl group would then produce 2-hydroxymethylclavam, and a further oxidation could yield clavam-2-carboxylate (Egan et al. 1997). Figure I.9 shows a hypothetical pathway for 5*S* clavam biosynthesis in *S. clavuligerus*.

As shown in Figure I.7, *cas2* is flanked by the rest of the clavulanic acid biosynthetic genes. In the same way, it would not be surprising to find genes

responsible for 5*S* clavam production upstream and downstream of *cas1*. Mosher et al. sequenced this area and found the clavam gene cluster (Mosher et al. 1999). In recent years the clavam gene cluster was extended by the work of Tahlan et al. Mutational analysis of the genes located in this region led to the proposal of a pathway for 5S clavam biosynthesis (Tahlan et al. 2007).

One of the genes in the clavam cluster is *cvm1*. The predicted product of this gene is very similar to aldo-keto reductases. Mutants with the *cvm1* gene disrupted had wild type levels of clavulanic acid but the production of clavam-2-carboxylate, 2-hydroxymethylclavam, and alanylclavam was abolished (Mosher et al. 1999). This suggests that *cvm1* might encode the enzyme responsible for the reduction of an aldehyde or ketone intermediate in the 5*S* clavam biosynthetic pathway (Zelyas 2007).

Another gene in the clavam cluster, which is immediately downstream of *cvm1* is *cvm2.* The predicted enzyme, Cvm2, shows a low degree of similarity to isomerases (Mosher et al. 1999; Tahlan et al. 2007). In *cvm2* mutants, the production of 5*S* clavams was reduced considerably (Tahlan et al. 2007). It is not clear what the exact role of Cvm2 would be in the 5*S* clavam biosynthetic pathway.

The predicted product of *cvm3,* a gene located downstream of *cvm2*, is a flavin reductase. This enzyme might provide reduced flavin for the enzyme produced by *cvm5* (a putative flavin dependent mono-oxygenase). Mutants lacking *cvm3* had normal levels of 5S clavam production. However, *cvm5* mutants were unable to produce any of the *5S* clavams and instead accumulated an

intermediate called 2-carboxymethylideneclavam. The structure of this intermediate is shown in Figure I.10 (Tahlan et al. 2007). The possible reason why *cvm3* mutants were not defective in 5S clavam biosynthesis might be that another enzyme other than Cvm3 is able to reduce flavin and provide it to Cvm5. It is believed that Cvm5 might carry out a Baeyer-Villiger oxidation of 2-carboxymethylideneclavam to yield 2-formyloxymethylclavam (Tahlan et al. 2007). This is shown in Figure I.9. Although *cvm1, cvm2* and *cvm5* are all known to be important or essential for 5*S* clavam production, the exact steps that Cvm1 and Cvm2 would catalyze are still unclear.

More genes have been identified upstream of *cas1*. Mutations in the genes, *cvm7, cvm11, cvm12, cvm13, cvmP* and *cvmG* of the clavam gene cluster, did not affect the production of clavulanic acid or 5*S* clavams and so they are likely not truly part of the clavam cluster. From the genes downstream of *cas1* (*cvm4, cvm5, cvm6, cvm9, cvm10*) only *cvm5* is shown to be essential for the production of 5*S* clavams (Tahlan et al. 2007).

In addition to *cas*, later studies showed that the other early genes (*ceaS, bls, pah, oat*) also exist in two copies in *S. clavuligerus*. (The early genes from the cephamycin-clavulanic acid cluster are hereafter all designated with the number 2). These studies showed that knocking out any of the early genes (*ceaS2, bls2, pah2, cas2, oat2*) did not eliminate clavam production completely. Therefore it was suggested that these genes are present in duplicates. It was known already that *cas* had two copies (Jensen et al. 2000). In 2004 Jensen et al. identified a second copy of the gene *pah*. This paralogue, *pah1,* showed 71% identity to *pah2*

(Jensen et al. 2004b). *pah1* mutants were able to produce clavams but in lower levels than wild type. However a *pah1/pah2* double mutant was unable to produce any clavams (Tahlan et al. 2004b).

More studies were conducted on the genes flanking *pah1*. This cluster is known as the paralogue cluster. The left hand edge end of the paralogue gene cluster (Figure I.7) was initially thought to be bounded by the gene *ceaS1*. CeaS1 was 66% similar to CeaS2 at the amino acid level. Mutation in *ceaS1* decreased the levels of clavulanic acid production. The next gene, which was located downstream of *ceaS1,* was *bls1* with a lower degree of similarity (49% at the amino acid level) to its paralogue (*bls2*). The *bls1* mutants did not show any difference in levels of clavam production compared to wild type. However *bls1/bls2* double mutants were unable to produce any clavams (Tahlan et al. 2004b).

Another gene in the paralogue cluster was *oat1*. This gene showed 47% identity to *oat2* at the amino acid level. Clavam production was not affected in *oat1* mutants and double mutants of *oat1/oat2* showed only decreased levels of clavam production. As mentioned above *ceaS1/ceaS2, bls1/bls2, pah1/pah2* and *cas1/cas2* double mutants could not produce any clavams, while *oat1/oat2* double mutants could produce the clavams but in lower amounts (compared to wild type). This observation further suggests that the role of *oat* is only to provide more arginine for clavam biosynthesis (Tahlan et al. 2004b).

The paralogue gene cluster also contains paralogues of two genes from the clavam gene cluster. These two are *c6p* and *c7p,* and they are paralogues of *cvm6*

and *cvm7,* respectively, from the clavam cluster. The product of *c6p* (and *cvm6*) is similar to aminotransferases. Although knocking out *cvm6* does not have any effect on clavam production, mutants lacking *c6p* were unable to produce 5*S* clavams. Mutation in *c7p* also abolished the biosynthesis of 5*S* clavams, while *cvm7* mutants had no detectable phenotype (Tahlan et al. 2007). It has been proposed by Egan et al. that clavaminic acid undergoes a deamination to produce an aldehyde intermediate (Egan et al. 1997). It is possible that C6p is the enzyme catalyzing this step.

It is still unclear why some genes are present in two copies in *S. clavuligerus.* Two reasons are proposed for this observation:

1. **A gene dosage effect:** to make sure that there is a large enough amount of these enzymes, which are important in the biosynthesis of clavulanic acid (Jensen and Paradkar 1999).
2. **Participation in two biosynthetic pathways:** as discussed earlier, clavulanic acid and the 5S clavams have a shared metabolic pathway up to the production of clavaminic acid. Therefore all these enzymes are required for both pathways (Jensen and Paradkar 1999).

Another gene in the paralogue cluster, *snk,* was found to have an important role in the production of 5*S* clavams. This gene is located downstream of *c7p.* Both *snk* and *c7p* are putative regulators (Tahlan 2005; Tahlan et al. 2007).

Sequencing the region upstream of *ceaS1* has led to the discovery of 12 new genes. Insertional mutagenesis of these genes showed that four of them had possible roles in the production of 5*S* clavams. These four genes are *orfA, orfB,*

orfC and *orfD* and they were shown to be necessary for alanylclavam biosynthesis (Zelyas et al. 2008).

The product of *orfA* shows similarity to serine hydroxymethyltransferases. These enzymes catalyze a reversible reaction involving the interconversion of glycine and 5,10-methylenetetrahydrofolate to serine and tetrahydrofolate respectively. OrfA contains three of the four amino acids required for binding to serine, however it lacks five of the six amino acids that are needed for binding to 5,10-methylenetetrahydrofolate (Scarsdale et al. 2000). It is possible that OrfA catalyzes the fusion of glycine to an aldehyde-containing clavam intermediate and yields 8-OH-alanylclavam (Zelyas et al. 2008).

Another one of the four genes is *orfB*. The product of this gene belongs to the family YjgF/YER057c/UK114 (Kim et al. 2001; Morishita et al. 1999; Oxelmark et al. 2000; Rappu et al. 1999). One of these proteins, YjgF, from *Salmonella enterica* serovar Typhimurium was shown to be necessary for the action of an aminotransferase called IlvB. This aminotransferase is involved in isoleucine biosynthesis. It is proposed that YjgF binds to, and stabilizes α-ketobutyrate and therefore allows for the activity of IlvB (Enos-Berlage et al. 1998; Schmitz and Downs 2004). Based on sequence similarity, OrfB is believed to be a stabilizer/regulator (Zelyas et al. 2008).

orfC is another gene in this region and the nucleotide sequence shows that OrfC belongs to the family I aminotransferases (Jensen and Gu 1996; Mehta et al. 1993). However OrfC only has two of the six amino acids required for substrate

binding in the class I aminotransferases (Kamitori et al. 1990; Okamoto et al. 1998; Rhee et al. 1997).

The product of *orfD* shows similarity to serine and threonine dehydratases. *orfD* is located downstream of *orfC*. The enzyme OrfD contains all of the residues that are needed for binding to its substrate (Gallagher et al. 1998; Simanshu et al. 2006; Yamada et al. 2003). It was proposed that OrfD catalyzes a dehydration/deamination reaction that would convert 8-OH-alanylclavam to a clavam intermediate with a pyruvyl side chain attached at C2. Then OrfC is believed to catalyze a transamination reaction on the keto group of the pyruvyl side chain and yield alanylclavam (Zelyas et al. 2008).

As mentioned above, OrfB resembles proteins that facilitate the activity of aminotransferases. It was suggested that OrfB sequesters the pyruvylclavam intermediate (which is an α-keto acid) allowing OrfC to catalyze the transamination reaction (Zelyas et al. 2008).

Mutants blocked in *orfD* or *orfC* both accumulated a clavam intermediate that is believed to be 8-OH-alanylclavam. Since *orfC* and *orfD* are part of an operon, the polar effect of knocking out *orfC* would cause *orfD* not to be expressed and therefore accumulate 8-OH-alanylclavam (Zelyas et al. 2008).

I.7.5 Regulatory proteins

Other than the environmental factors that affect the levels of clavulanic acid and 5*S* clavams, there are regulatory proteins inside the cells that affect the production of these metabolites. Understanding the mechanisms of these

regulators helps investigate ways of manipulating the production of different metabolites.

A family of proteins that has been identified in some antibiotic regulatory systems is named the *Streptomyces* Antibiotic Regulatory Protein (SARP) family. This family of proteins has an OmpR-like winged helix domain that binds to DNA (Wietzorrek and Bibb 1997). CcaR (cephamycin and clavulanic acid regulator) is one of the regulatory proteins in *S. clavuligerus*. This protein regulates the genes involved in clavulanic acid and cephamycin production. It is encoded by *ccaR,* which is located in the cephamycin gene cluster and it is shown to be essential for the production of both cephamycin and clavulanic acid (Alexander and Jensen 1998; Perez-Llarena et al. 1997). Mutants lacking *ccaR* are incapable of producing many of the enzymes in the cephamycin biosynthetic pathway (Alexander and Jensen 1998). CcaR is also required for the expression of the polycistronic transcript that includes *ceaS2, bls2, pah2,* and *cas2*. CcaR is also known to activate its own transcription (it can bind to its own promoter) (Santamarta et al. 2002).

CcaR also controls *claR,* which is a putative transcriptional regulator located in the clavulanic acid gene cluster. ClaR, the product of the gene *claR,* is a LysR-type transcriptional regulator. ClaR has helix-turn-helix DNA binding domains on one of the ends. Mutants lacking *claR* were able to produce the 5*S* clavams and cephamycin C, but no clavulanic acid (Paradkar et al. 1998; Perez-Redondo et al. 1998). ClaR was shown to be required for the transcription of some of the late genes in clavulanic acid production, including *cyp, cad* and

oppA1 (Paradkar et al. 1998). There is no transcription of *claR* observed in *ccaR* mutants. This clearly indicates the involvement of CcaR for both the early and late steps of clavulanic acid production (Perez-Redondo et al. 1998; Tahlan et al. 2004b).

Interestingly, CcaR can be a regulatory target for *bldA*. *bldA* is a regulator of morphological differentiation and antibiotic production in *S. coelicolor*. Mutants in this gene are capable of vegetative growth, however they do not form aerial hyphae and therefore do not produce any antibiotics (Merrick 1976). Some genes require the leucyl tRNA to translate the rare UUA codons. This tRNA is produced by *bldA* and this gene is the only gene in the whole genome of *S. coelicolor* that is known to produce that specific tRNA (Bentley et al. 2002; Lawlor et al. 1987; Leskiw et al. 1991b). It is known that despite the transcription of *bldA* in the primary growth phase of *S. coelicolor,* the active form of the transcript only appears at the time of secondary metabolism. This gene is a regulator of the genes needed for the secondary metabolism. It has also been shown that genes involved in antibiotic regulation, formation of aerial mycelia, and antibiotic resistance often have TTA codons (Leskiw et al. 1991a and 1993).

Since *ccaR* has a TTA codon, it was expected that it would be regulated by *bldA* (Perez-Llarena et al. 1997). Surprisingly, *S. clavuligerus* mutants lacking *bldA* are capable of producing clavulanic acid and cephamycin C, although these *bldA* mutants were unable to produce aerial mycelia. It was shown that CcaR was still produced in *S. clavuligerus bldA* mutants. It is possible that the TTA codon was being mistranslated by a non-cognate leucyl tRNA (Trepanier et al. 2002).

The expression of *ccaR* is also dependent on another gene called *bldG* (Bignell et al. 2005). The gene *bldG* was first identified in *S. coelicolor*. The product of this gene, BldG, shows similarity to anti-anti-sigma factors in *Bacillus subtilis* and *Staphylococcus aureus* (Bignell et al. 2000). Sigma factors determine the DNA binding specificity of the RNA polymerase. Anti-sigma factors prevent the interaction of the holoenzyme with the sigma factor by binding to the latter. Unphosphorylated anti-anti-sigma factors bind to specific anti-sigma factors and allow the sigma factor to bind to the holoenzyme. However when they are phosphorylated, anti-anti-sigma factors are unable to interact with anti-sigma factors. *S. coelicolor* mutants lacking *bldG* were unable to develop aerial hyphae and could not produce any antibiotics (Bignell et al. 2000). BldG is a protein capable of undergoing phosphorylation. Mutants that were defective in BldG phosphorylation could not produce any antibiotics or aerial mycelia (Bignell et al. 2003).

A homologue of *bldG* was found in *S. clavuligerus* and it was found to be necessary for the development of aerial mycelia and production of all clavams and cephamycin C. It was also shown to be needed for the expression of *ccaR* (Bignell et al. 2005), but CcaR does not regulate the production of 5*S* clavams. Therefore BldG is also likely to be needed for the expression of some other 5*S* clavam production genes (Tahlan et al. 2004a).

Secondary metabolism can also be regulated by microbial hormones in some *Streptomyces* spp. The structure of these hormones is a γ-butyrolactone-type compound and when they are secreted and taken up by the surrounding cells they

affect morphological development and antibiotic production. *Streptomyces griseus* has an A-factor hormone that is necessary for streptomycin production and the formation of aerial hyphae (Hara and Beppu 1982; Khokhlov et al. 1967). The key enzyme for the production of this hormone is produced by a gene called *afsA* (Ando et al. 1997; Horinouchi et al. 1989). The A-factor binds to a protein known as ArpA (A-factor receptor protein). This protein is a repressor and when it is bound to the A-factor the DNA is transcribed (Ohnishi et al. 1999; Onaka et al. 1995; Onaka and Horinouchi 1997). ArpA regulates the expression of *adpA* (A-factor dependent protein) and the product of this gene is a transcriptional regulator for another regulator named *strR*. The product of *strR* is a positive regulator for the biosynthetic genes of streptomycin (Ohnishi et al. 1999; Retzlaff and Distler 1995; Vujaklija et al. 1993). The gene *adpA* also regulates the expression of genes involved in antibiotic resistance, secondary metabolite production and morphological differentiation (Yamazaki et al. 2003 and 2004). AREs (autoregulatory elements) are the DNA-binding sites for ArpA and other γ-butyrolactone receptors (Kinoshita et al. 1999; Yamazaki et al. 2004).

S. clavuligerus produces a γ-butyrolactone autoregulator and an ARE has been identified upstream *ccaR* (Folcher et al. 2001; Hashimoto et al. 1992). ScaR (produced by *scaR*) is a protein that binds to certain types of γ-butyrolactone compounds in *S. clavuligerus*. ScaR also regulates its own expression by binding to an ARE upstream of *scaR* (Hyun et al. 2004; Santamarta et al. 2005). *scaR* mutants produce clavulanic acid and cephamycin C. Therefore it is suggested that

ScaR is a repressor (like ArpA) and inhibits the expression of *ccaR* (Santamarta et al. 2005).

It is known that during amino acid starvation (p)ppGpp (a polyphosphorylated nucleotide) accumulates and causes a significant decrease in the levels of tRNA and rRNA. It also decreases stable mRNA production. This set of events is known as the stringent response. RelA is a ribosome-associated protein that together with SpoT synthesizes (p)ppGpp. L11, a ribosomal protein, is important for the activity of RelA. Mutants defective in these genes are termed relaxed mutants and they produce low levels of (p)ppGpp. Antibiotic production in *Streptomyces* spp. occurs in the stationary phase, which is associated with nutrient starvation. It would therefore not be surprising that (p)ppGpp is produced more during antibiotic production. Relaxed mutants of *Streptomyces* sp. MA406-A-1 and *S. antibioticus* produced no formycin and actinomycin, their natural antibiotic products respectively (Kelly et al. 1991; Ochi 1986).

Production of clavulanic acid in *S. clavuligerus* was shown to occur after the accumulation of (p)ppGpp (Jones et al. 1997). *relA* and *rsh* are homologues of *relA* and *spoT* that have been identified in *S. clavuligerus*. *relA* mutants were unable to produce (p)ppGpp and *rsh* mutants produced lower levels of this nucleotide. These mutants could not produce any cephamycin C and they produced low amounts of clavulanic acid (Jin et al. 2004). L11 mutants of *S. clavuligerus* also produced low levels of clavulanic acid, cephamycin C and (p)ppGpp. This mutant produced lower levels of transcripts of *ccaR* and the genes

positively regulated by the product of this gene. (Gomez-Escribano et al. 2006). The exact role of (p)ppGpp in antibiotic production is still unclear.

There is a novel class of antibiotic regulatory proteins that possesses a C-terminal ATPase domain and an N-terminal SARP-like domain. PimR from *Streptomyces natalensis* is an example of this class and regulates the production of pimaricin (Anton et al. 2004). *c7p,* one of the genes in the paralogue cluster is similar to *pimR.* Mutants lacking this gene produce no 5*S* clavams but are capable of clavulanic acid production and morphological differentiation. *cvm7,* a paralogue of *c7p,* also belongs to this novel class of regulatory proteins, but clavam biosynthesis is not affected in *cvm7* mutants. It is possible that this gene is responsible for regulating the biosynthesis of other antibiotics in *S. clavuligerus* (Tahlan et al. 2007).

The gene *snk* produces a sensor kinase and is located immediately downstream of *c7p.* It is usual that genes encoding cognate sensor kinases and response regulators are immediately adjacent and transcribed within a single operon, but C7p is not a response regulator. However, preliminary studies have shown two response regulators for *snk, res1* and *res2,* are located further downstream in the cluster. *snk* mutants produce no 5*S* clavams but wild type levels of clavulanic acid, suggesting a specific role in regulation of 5*S* clavam production (Tahlan 2005).

I.8 Gaps of knowledge, goals of the present study

5*S* clavams are considered unwanted products in clavulanic acid production in industry and they complicate the downstream processing. In addition, as stated before, they neither have a real antibacterial effect, nor β-lactamase inhibitory properties. Therefore the main goal during industrial production is to minimize the production of the 5*S* clavam metabolites without negatively affecting clavulanic acid production. The study of their biosynthetic pathway would give better insight about the steps and genes they share with clavulanic acid and the regulatory mechanisms that control their biosynthesis. Having this knowledge, genetic manipulations can be made in ways to decrease or even eliminate the biosynthesis of 5*S* clavams, while the levels of clavulanic acid are maintained, or preferably increased.

Another feature of the 5S clavams that would be worth studying, relates to the differences in structure between clavulanic acid and 5*S* clavams. As stated before, the 5*S* clavams can have different side chains on their C-2, which determine the type of the 5*S* clavam. It seems rational to believe that the side chains affect the activity of the clavams. It would be interesting to substitute these side chains on the C-2 position of clavulanic acid and study the changes of activity in clavulanic acid. Clavulanic acid has a different stereochemistry than the 5*S* clavams, so it might be possible that the effect that a side chain (when it is on its original clavam) would be different from when it is substituted on clavulanic acid. It is also possible that by substituting these side chains, clavulanic

acid would acquire a better β-lactamase inhibitory effect. This part, however, requires more background in chemistry and biochemistry and is beyond the scope of the present Master's degree project.

S, antibioticus Tü1718 is the only strain of *S. antibioticus* discovered to date that is able to produce 5*S* clavams. Considering the fact that this microorganism produces only two 5*S* clavams and no clavulanic acid, it might be possible that some of the genes encoding these two clavams have been transferred to this strain from *S. clavuligerus,* or even some other *Streptomyces* spp.*,* that is able to produce 5*S* clavams. Or it might be even possible that all of these come from the same source. If the order of the genes in the cluster looks the same in some of the species, then the hypothesis that they have the same origin might be true.

As discussed in detail in the section for the biosynthetic pathways, the biosynthetic pathway of clavulanic acid and the 5*S* clavams is shared up to the production of clavaminic acid (Janc et al. 1995). As seen in Figure I.8 and Figure I.9, there are still many unknown steps in the biosynthesis of the 5*S* clavams. Despite all of the studies done on the biosynthetic genes involved in the early and late steps of clavulanic acid and 5*S* clavam production, the functions of many of these genes are yet to be known. Some of the biosynthetic intermediates might be unstable and be degraded readily, making it impossible to detect them. Therefore finding the enzymes responsible for their formation would be very difficult.

As the literature shows, most of the microbiological studies done on 5*S* clavam production so far, have involved *S. clavuligerus.* This species is unusual in that it produces both clavulanic acid and 5*S* clavams. The genes for clavam

production are located in three distantly located clusters. It should be less complicated to try to understand 5*S* clavam biosynthesis by studying an organism that can only produce the 5*S* clavams, because there is probably only one gene cluster that would contain all of the clavam biosynthesis genes. Using this approach, it would be possible to find the clavam gene cluster from the microorganism and compare the genes to *S. clavuligerus*. *S. antibioticus* Tü1718 is the only other clavam producer that has been studied previously and an enzyme similar to Cas in *S. clavuligerus* was found in this microorganism (Janc et al. 1995). There are no genetic studies done on *S. antibioticus* Tü1718. This microorganism would therefore be a good candidate. Studies on 5*S* clavam biosynthesis in *S. antibioticus* Tü1718 are described.

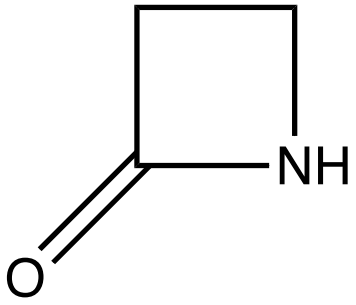


Figure I.1: The structure of the β-lactam ring. This structure is shared in all the β-lactam antibiotics.

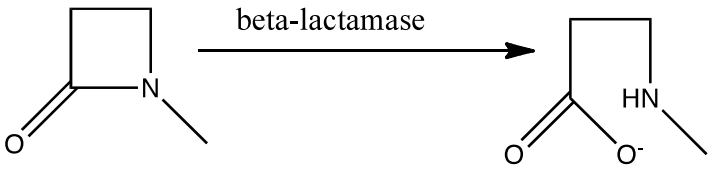
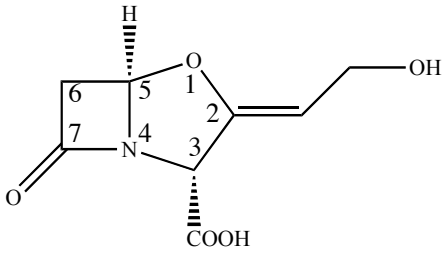
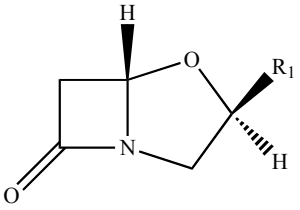


Figure I.2: A scheme of β-lactamase activity.

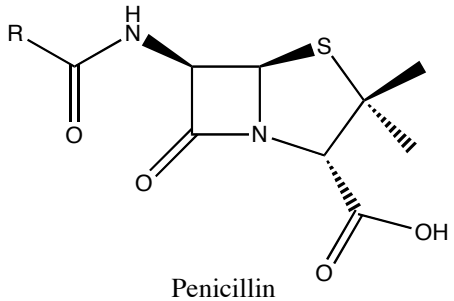


Clavulanic acid



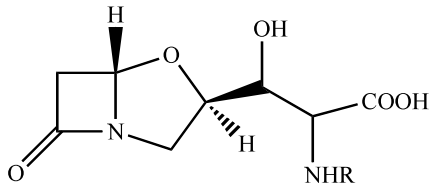
5*S* clavams

R_1	Clavam metabolite
COOH	clavam-2-carboxylate
CH_2OH	2-hydroxymethylclavam
CH_2OCHO	2-formylmethylclavam
CH_2CH_2OH	2-hydroxyethylclavam
CH_2CHNH_2COOH	alanylclavam
$CH_2CH(OH)CHCOOH$ | $NHCOCH(NH_2)CHMe_2$	valclavam

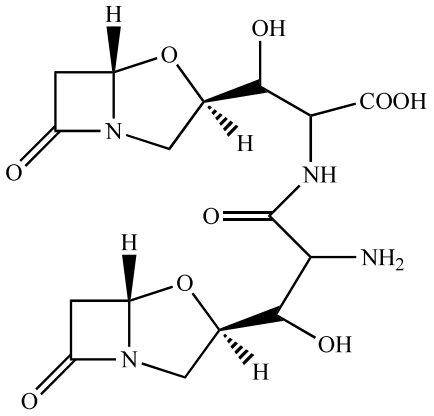


Penicillin

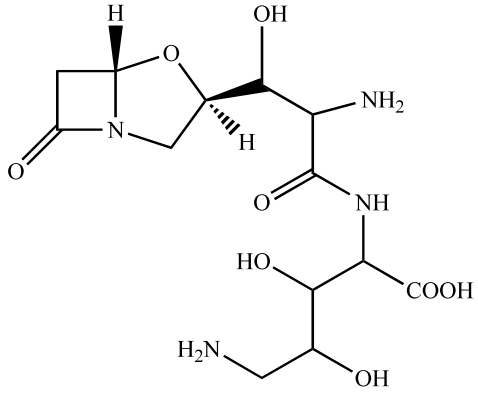
Figure I.3: The structure of clavulanic acid compared to some of the 5*S* clavams. These 5*S* clavams are not all produced by *S. clavuligerus.* The structure of penicillin is also shown at the bottom for reference and comparing to the clavams.



clavamycins



clavamycin A



clavamycin B

R	clavamycin metabolite
$COCH(NH_2)CH(OH)CH(OH)CH_2NH_2$	clavamycin C
L-valyl	clavamycin D
L-alanyl	clavamycin E
L-N^5-acetylornithine	clavamycin F

Figure I.4: The structure of some of the clavamycins.

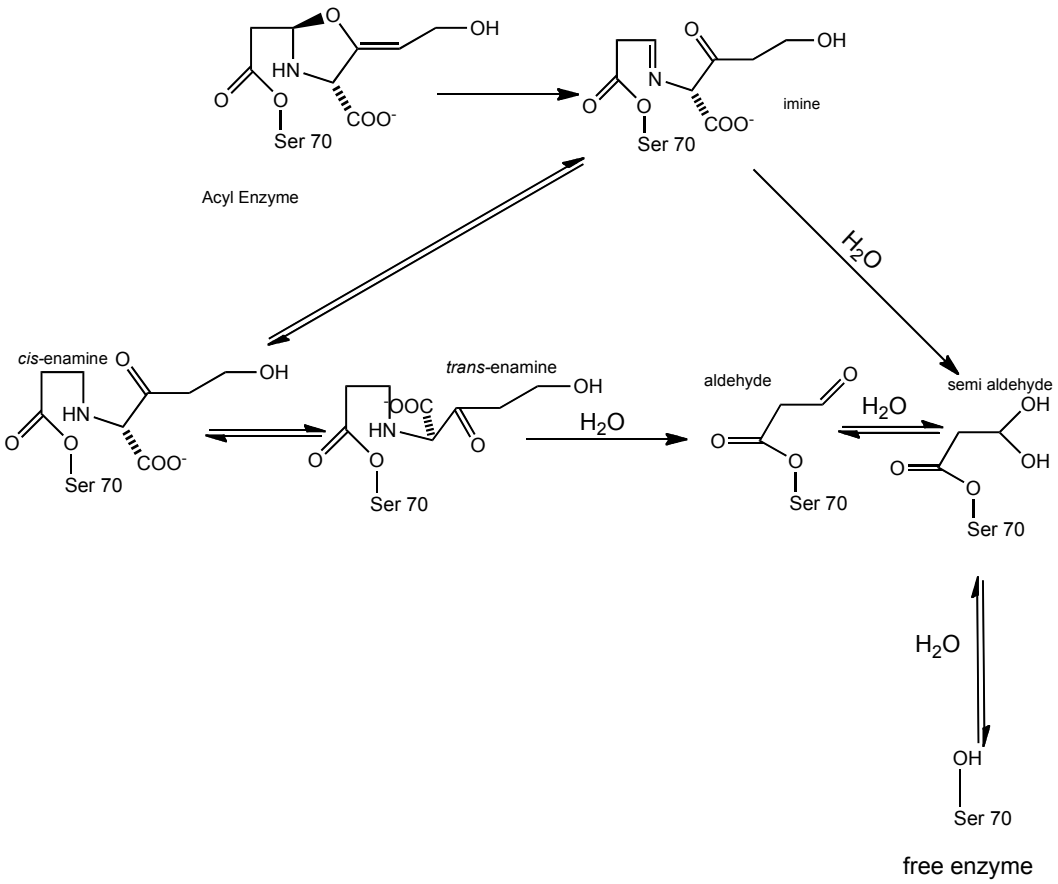


Figure I.5: Deactivation of β-lactamase enzyme by clavulanic acid.

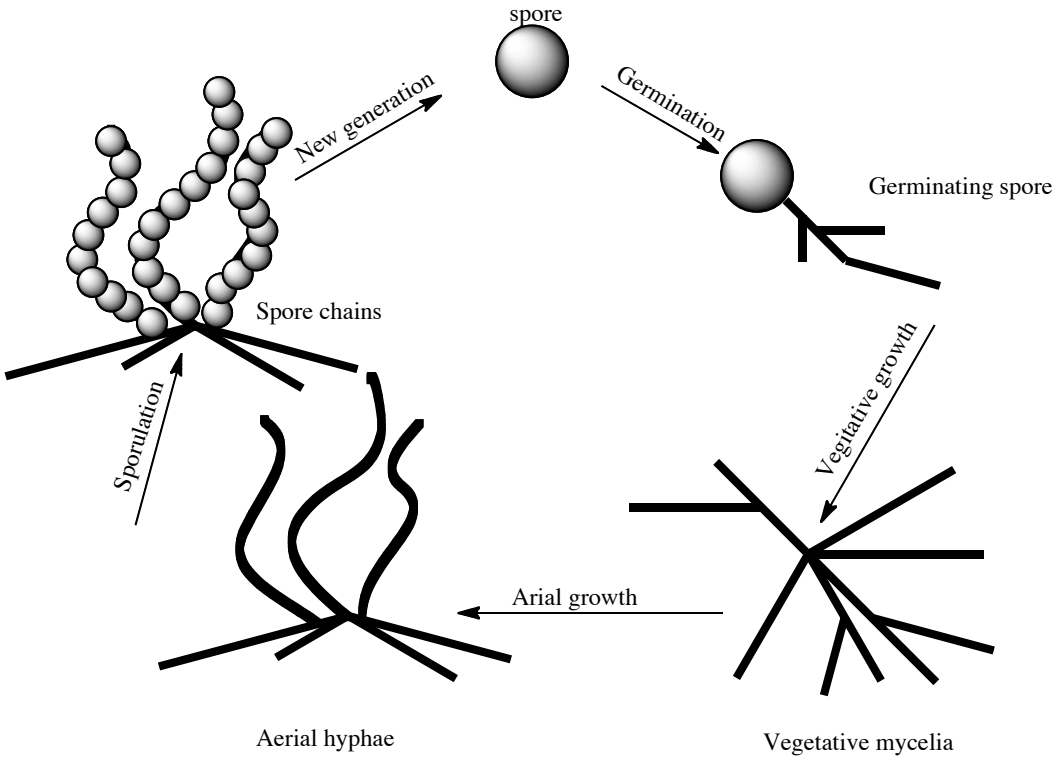


Figure I.6: *Streptomyces* life cycle.

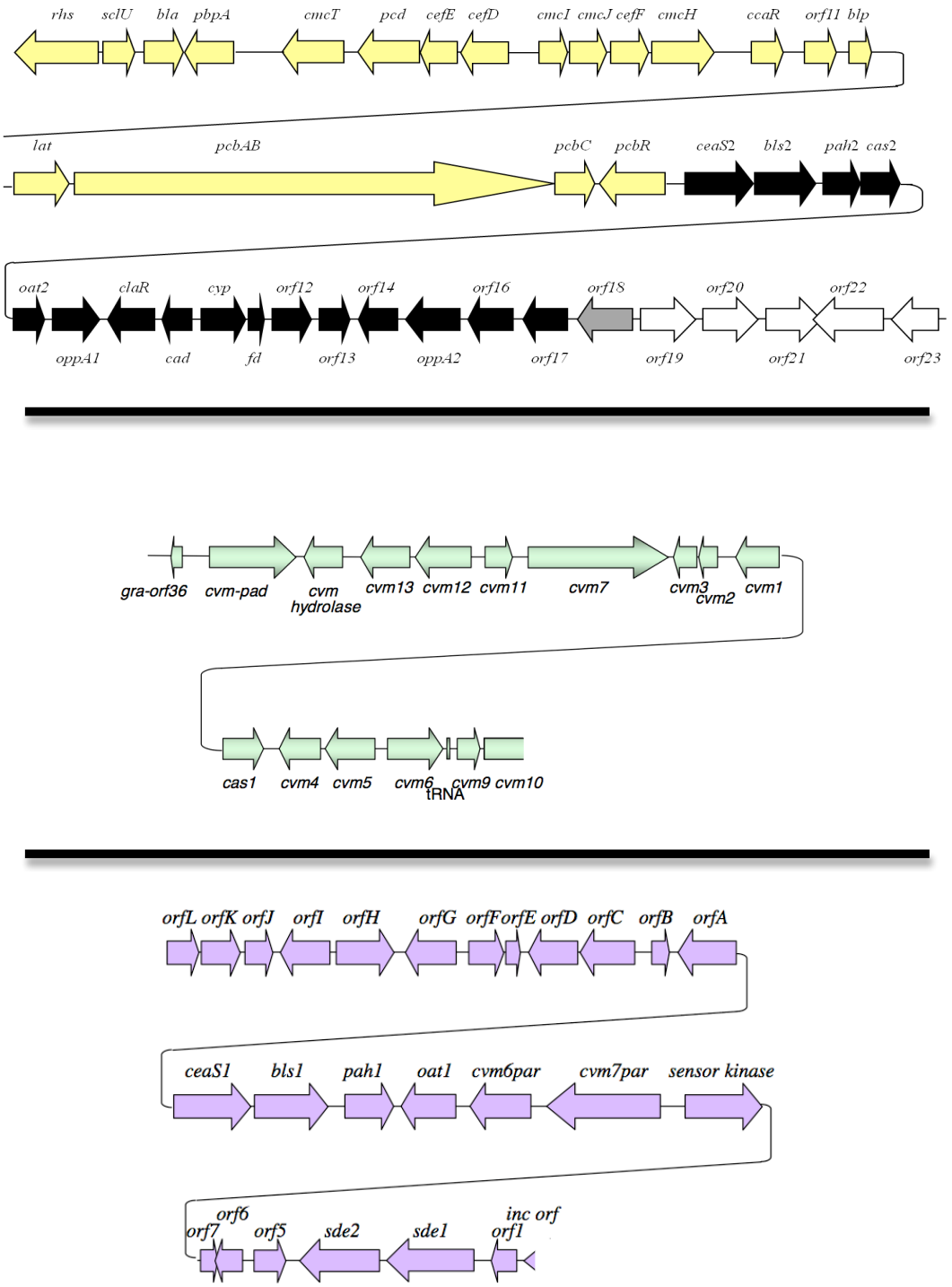


Figure I.7: The three gene clusters in *S. clavuligerus* that are responsible for clavam biosynthesis. From top to the bottom: cephamycin-clavulanic acid super cluster (blue and yellow, the black genes are clavulanic acid production genes and the yellow genes are cephamycin C producing genes), the clavam gene cluster (green) and the paralogue gene cluster (purple).

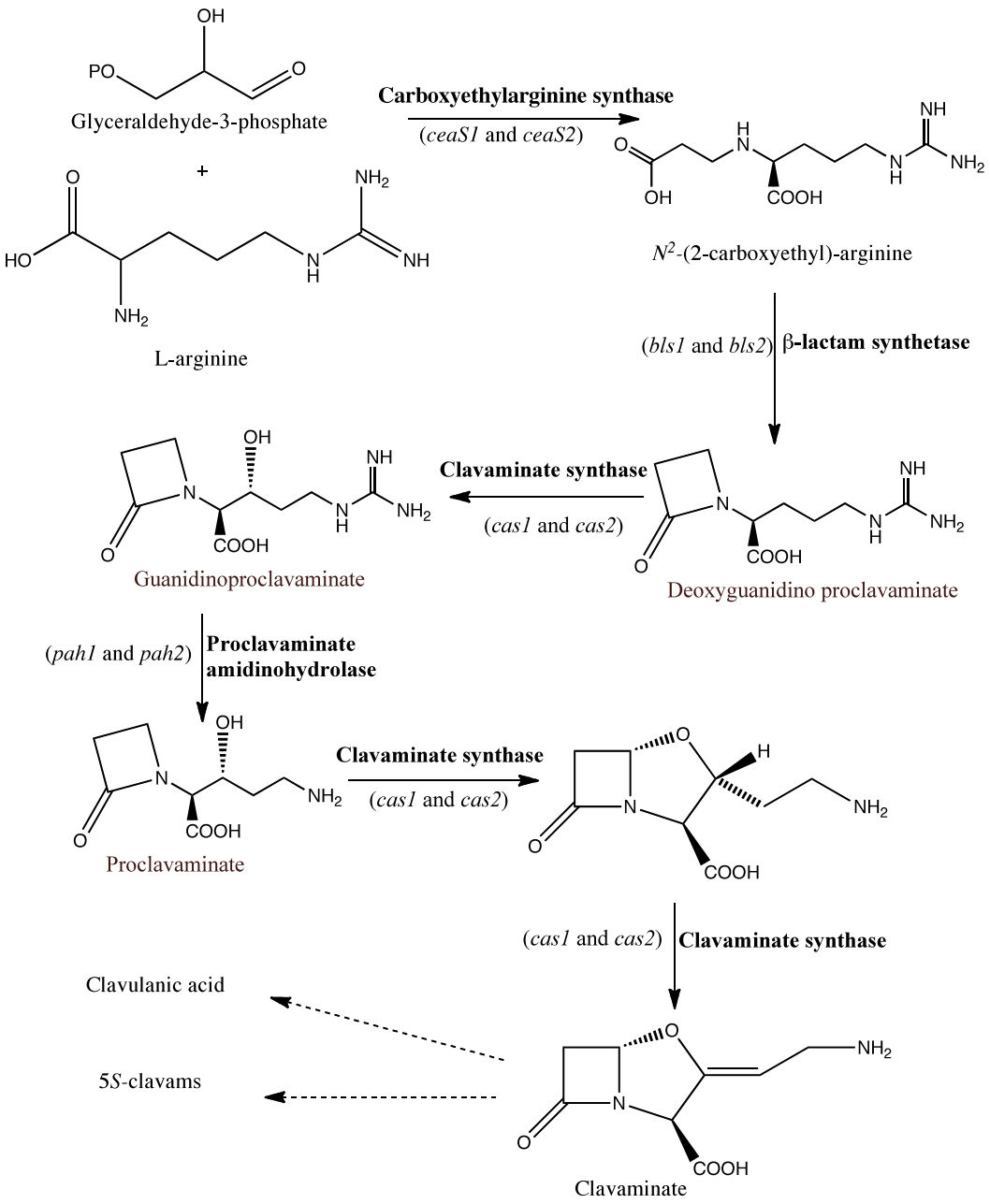


Figure I.8: The early steps of the biosynthetic pathway to clavulanic acid and 5*S* clavams in *S. clavuligerus.*

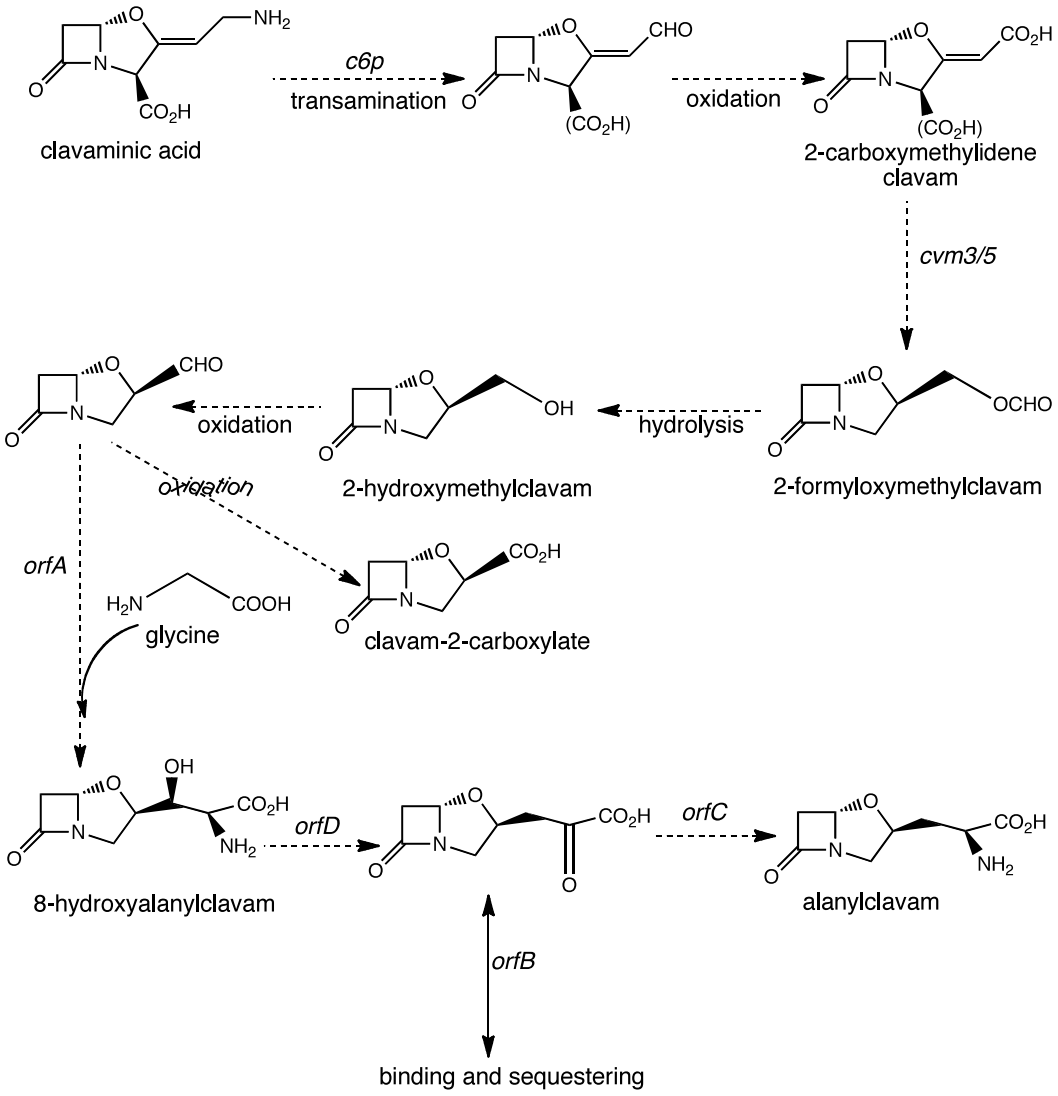


Figure I.9: Hypothesized biosynthetic pathway for the 5*S* clavams in *S. clavuligerus.* The dashed arrows indicate that the steps are hypothesized.

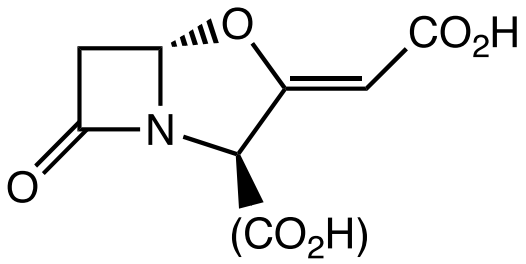


Figure I.10: the structure of the intermediate 2-carboxymethylideneclavam. This intermediate accumulates in *cvm5* mutants in *S. clavuligerus.* The bracket indicates that it is not clear whether or not the intermediate contains a carboxyl group at C-3.

II. Materials and Methods

II.1 Bacterial strains, culture media and culture conditions

Strains of *Streptomyces* and *Escherichia coli* used in this study are described in Tables II.1 and II.2 respectively.

Liquid medium grown cultures were used for measuring the production of *5S* clavams. These cultures were prepared from *Streptomyces* spores. Spores of *S. antibioticus* Tü1718 were collected from ISP2 agar plates (DIFCO Laboratories). Inoculated plates were incubated at 28°C for seven days. Seed medium was prepared as described by Janc et al. (1995) (0.4 g L-methionine, 4.0 g D-glucose, 4.0 g yeast extract, 10 g malt extract, 1000 ml H_2O, pH 7.3) and inoculated with 10^8 spores and incubated for 20 hours on a Model G-25 rotary shaker (New Brunswick Scientific Co., Edison, NJ) at 28°C and 250 rpm. Fermentation medium was also prepared according to Janc et al. (1995) as follows: 30 g mannitol, 30 g soy flour, 1000 ml H_2O, pH7.2. Fermentation flasks were incubated at 28°C and 250 rpm for 72 hours on the same shaker used for the seed cultures.

Trypticase Soy Broth (DIFCO Laboratories) supplemented with 1% (w/v) soluble starch (TSB-S) was used for growing *Streptomyces* in order to extract the genomic DNA (gDNA). The cultures were incubated at 37ºC for 20 hours on the same shaker used for the seed cultures.

Lennox broth (LB; 10 g tryptone, 5 g yeast extract, 5 g NaCl, 1000 ml H_2O) was used for growing *E. coli* strains. The media were incubated at 37°C on a Cell

Production Roller Drum (Bellco Biotechnology, Vineland, NJ) or a G-24 Environmental incubator shaker (New Brunswick Scientific Co., Edison, NJ). LB agar plates were prepared by adding 15 g/L agar to LB. In order to grow antibiotic-resistant strains, the media were supplemented with either one, or a combination of the following antibiotics (depending on the resistance gene that is inserted): ampicillin (100 μg/ml), chloramphenicol (25 μg/ml), apramycin (50 μg/ml), kanamycin (50 μg/ml). For selecting *Streptomyces* mutants containing antibiotic markers, ISP2 agar plates were supplemented with nalidixic acid (25 μg/ml) and apramycin (25 μg/ml).

SOB medium was used for making electrocompetent cells: 20 g tryptone, 5 g yeast extract, 0.5 g NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 20 mM $MgSO_4$ and 1000 ml H_2O. When used for growth of *E. coli* BW25113/pIJ790, the SOB was supplemented with 10mM L-arabinose to induce the λ-Red genes on pIJ790 (required for recombination).

II.2 Storage of bacterial strains as glycerol stocks

In order to keep a stock of the cells containing specific constructs or cells that were required for later steps in the study, they were kept in a 20% glycerol solution and stored at -80°C. This methods provides a reliable way of storing the cells with minimum chance of the cells dying or having mutations. One and half milliliters of 16 hour grown cultures of *E. coli* strains was centrifuged and the pellet was re-suspended in 500 μl of 20% (v/v) glycerol and stored at -80°C.

Streptomyces strains were stored as spore suspensions. After scraping the *Streptomyces* spores from ISP2, they were re-suspended in 5 ml of dH_2O. The suspensions were placed in a sonication bath for a few minutes to breakdown the clumps of spores. Then the suspension was passed through non-absorbent cotton to remove the mycelia. The mycelia-free suspensions were then centrifuged in an Eppendorf 5415 C microcentrifuge in 1.5 ml centrifuge tubes at 13000 rpm for 5 minutes. The spores were re-suspended in 20% (v/v) glycerol and stored at -80°C.

II.3 Cosmids and plasmids

Table II.3 provides a list of the cosmids and plasmids that were used in this study.

II.4 DNA methods

II.4.1 DNA quantification

The amount of DNA was quantified via spectrophotometry using a UNICAM UV/Vis Spectrophotometer UV3 (ATI Unicam, Cambridge, UK) or with a NanoDrop® ND-1000 Spectrophotometer (NanoDrop Technologies, Inc., Wilmington, DE). For UV/Vis spectrophotometery, absorbance was measured at 260nm and the machine was calibrated with Milli-Q™ water, which was used for dissolving DNA. The A_{260} was then multiplied by 50 μg/ml DNA and then by the dilution factor. When NanoDrop® was used the machine was blanked using Milli-Q™ water and the absorbance was scanned from 220 to 350nm. The concentration of DNA in the sample and the A_{260}:A_{280} ratio, used to assess quality of the DNA preparation, were measured using the NanoDrop spectrophotometer.

II.4.3 Isolation of gDNA from *S. antibioticus* Tü1718

Because the gDNA was going to be used for constructing a library it had to be extracted by a method that would give the largest possible fragments of DNA. Therefore the salting out method was used (Kieser et al. 2000). In this method 30 ml TSB-S medium was inoculated with the spore suspension (10^8 spores) and after harvesting the mycelium using a Brinkmann centrifuge (Brinkmann Instruments Inc. Mississauga, ON) (500 x g, 10 minutes), the cell material was re-suspended in 5ml SET buffer (75 mM NaCl, 25 mM EDTA pH8, 20mM Tris-HCl pH 7.5) with 100µl lysozyme (final concentration 1 mg/ml). This mixture was incubated at 37ºC for 1 hour. Then 140 µl proteinase K (final concentration 0.5 mg/ml) and 600 µl 10% SDS were added and incubation was continued at 55ºC for 2 hours. During incubation the mixture was inverted occasionally. After this step 2 ml 5 M NaCl (final concentration 1.25 M) was added and the mixture was cooled to 37ºC. Five milliliters of chloroform was then added and mixed by inversion for 30 minutes at 20ºC. To separate the supernatant, the mixture was centrifuged for 15 minutes at 4500 x g at 20ºC. The supernatant was then transferred to a fresh tube and 0.6 vol isopropanol was added and mixed by inversion. DNA was spooled on a sealed Pasteur pipette after about 3 minutes. Finally the DNA was rinsed with c. 5ml 70% ethanol and air-dried then dissolved in 1ml TE (10 mM Tris pH 8.0, 1mM EDTA) at 55ºC.

II.4.4 Isolation of plasmid and cosmid DNA from *E.coli*

The alkaline lysis method was used for isolating plasmids (Birnboim and Doly 1979) and cosmids (Sambrook et al. 1989).

II.4.5 Introduction of DNA into electrocompetent *E. coli*

In order to make electrocompetent *E. coli* cells, 10 ml LB cultures of the cells to be made electrocompetent were grown 16 hours at 37ºC (or 30ºC in the case of the BW25113/pIJ790 strain of *E. coli*). One hundred microliters of this culture was used to inoculate 10 ml SOB and then it was incubated at 37ºC (or 30ºC) until the OD_{600} reached 0.4-0.6. The cells were then harvested (at 4ºC) and washed with an equal volume of 4ºC 10% (v/v) glycerol. Cells were centrifuged and washed two more times, once with half-volume and the next time with 8 ml of ice-cold 10% glycerol. Finally the cells were re-suspended in 1 ml 10% (v/v) glycerol. Forty microliters was transferred into each 1.5 ml centrifuge tube and these were flash-frozen in dry-ice containing 95% ethanol. The aliquots were stored at -80ºC.

When electrocompetent cells were prepared for PCR-targeting mutagenesis (REDIRECT©), they were cultured as described above but they were grown in SOB supplemented with 10 mM arabinose. After reaching the OD_{600} of 0.4-0.6 the cells were washed twice with one volume of 4ºC 10% (v/v) glycerol and then re-suspended in the residual glycerol solution and 50 μl was used immediately for electroporation.

To carry on the electroporation, the electrocompetent cells were thawed on ice and the DNA was added to them. The mixture was then transferred to an ice-cold 2 mm gap BTX® Disposable Cuvette Plus™ (Genetromics, Inc., San Diego, CA). This cuvette was pulsed using a Bio-Rad GenePulser II. The device was set to 200 Ω, 25 μF and 2.5 kV. The cells were transferred to 1 ml ice-cold LB and

incubated at 37ºC (or 30ºC for BW25113/pIJ790 strain) for an hour. Then the cells were spread on an LB agar plate containing the required antibiotic. The plates were incubated at 37ºC (or 30ºC). If blue-white screening was performed (to detect the cells containing the DNA inserted in the vector), 100 mM isopropyl-β-D-thiogalactopyranoside (IPTG) and 40 μg/ml 5-bromo-4-chloro-3-β-D-galactopyranoside (X-gal) were also spread on the plates.

II.4.6 Agarose gel electrophoresis

DNA samples to be analyzed by gel electrophoresis were mixed with 5x loading buffer [60% (w/v) sucrose, 100 mM EDTA, 0.25% (w/v) bromophenol blue, 0.25% (w/v) ficoll orange and 0.25% (w/v) xylene cyanol). Depending on the size of fragment of interest, 0.5%-1.5% agarose gels were used for analyzing the DNA molecules. The gels were electrophoresed in 1xTAE buffer (40 mM Tris-Acetate and 1 mM EDTA, pH 8.0) at 70-90 V. Gels were stained in 2 μg/ml ethidium bromide for 5-20 minutes. The stained gels were observed under UV light. Digesting λ phage DNA with either *PstI* or *BstEII* made molecular weight markers.

II.4.7 DNA purification from agarose gels

After separating the DNA fragments on an agarose gel, staining and visualizing them under UV light, the desired fragment was excised from the gel. Purification was then carried out according to the instructions provided by QIAquick Gel Extraction Kit (Qiagen Inc., Valencia, CA).

II.4.8 DNA digestion and cloning (with and without the TOPO TA Cloning® kit)

DNA was digested using restriction enzymes from New England Biolabs (New England Biolabs, Ipswich, MA) or Roche (Roche, Basel, Switzerland). Depending on the size of the DNA molecule the digest was carried out for an hour or 16 hours. The desired band was purified from an agarose gel as described above.

After purifying the fragment, it was inserted into vectors using 1 unit of T4 DNA ligase and 10X ligation buffer (both purchased from Roche, Basel, Switzerland). The ratio of insert to vector was three to one and the total volume was 10-30 μl. Ligation reactions were carried out at 16°C for 16 hours.

When the TOPO TA Cloning® kit was used, 4 μl of the PCR product (see section II.4.9) was added to 1 μl of the vector pCR®2.1-TOPO® (Invitrogen). After incubating the mixture for 5 minutes at 20ºC, 2 μl of it was added to a vial of One Shot® TOP10 *E. coli* chemically competent cells (Table II.1). The latter mixture was kept at 4ºC for 25 minutes and then heat shocked for 30 seconds at 42°C. Two hundred and fifty microliters of SOC [20 g/l tryptone, 5 g/l yeast extract, 0.5 g/l NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, and 20 mM glucose (pH 7.0)] was added to the cells and the tube was placed on a tube roller at 37°C for an hour. After this step, 10 and 50 μl of the transformation mixture was spread on LB agar plates containing 100 μg/ml ampicillin and 0.008% (w/v) X-gal. The plates were incubated for 16 hours at 37°C and white colonies were chosen and transferred to LB and further analyzed.

II.4.9 Polymerase chain reaction (PCR)

The amount of template DNA used for PCR was 5-10 ng (plasmid or cosmid) or 500 ng (gDNA). The reaction also contained 50 pmol of each oligonucleotide primer (Table II.4), 50 mM Tris-HCl (pH 9.0), 1.5 mM $MgCl_2$, 0.4 mM β-mercaptoethanol, 0.1 mg/ml purified bovine serum albumin (BSA) (New England Biolabs, Ipswich, MA), 10 mM $(NH_4)_2SO_4$, 0.2 mM of each dNTP, and 5% (v/v) DMSO in a total volume of 50 μl. Where absolute sequence fidelity was not required 0.5 μl of Taq polymerase was added to the reaction (Taq polymerase was supplied by the Fermentation Service Unit, Department of Biological Sciences, University of Alberta). When high fidelity sequence was needed for the PCR product, the EXPAND High Fidelity PCR system (Roche, Basel, Switzerland) was used according to the manufacturers' instructions. The cycling conditions were as follows: 2 minutes at 96°C, 30 cycles of 96°C (30-60 seconds) and 55-60°C (30 seconds), 72ºC for 5 minutes. PCR was carried out in thin-walled 0.2 ml tubes in a TGradient thermocycler (Biometra, Goettingen, Germany). If the template DNA was gDNA or a difficult template to amplify, the initial 2 minutes at 96°C was extended to 5 minutes, and the dNTPs and Taq were only added to the reaction tubes at the last 30 seconds. The amount of DMSO could also be increased to 10%.

For amplifying regions containing secondary structures, instead of regular dGTP, an analog, 7-deaza-dGTP (NEB), was used. This nucleotide hydrogen bonds less strongly with cytosine, making the secondary structure of the PCR product less stable and easier to sequence.

To amplify the REDIRECT© extended resistance cassette, pIJ773 was first digested with *EcoRI* and *HindIII* and the apramycin resistance cassette was gel purified. Fifty nanograms of this product was used as template for PCR. The PCR conditions were as follows: 2 minutes at 94°C, 10 cycles of 94°C (45 seconds), 50°C (45 seconds) and 72°C (90 seconds), 15 cycles of 94°C (45 seconds), 55°C (45 seconds) and 72°C (90 seconds), and a final 5 minutes at 72°C.

II.4.10 DNA sequencing

The BigDye terminator v3.1 Cycle Sequencing kit (Applied Biosystems Foster City, CA) was used for sequencing DNA. The amount of template DNA that was used depended on the nature of the molecule: cosmids were used at 800 ng, plasmids were used at 400 ng and PCR products were used at 50-100 ng. The total volume of each reaction was 20 μl and it consisted of 10 pmol of the primer, 5% DMSO, 2 μl BigDye and 6 μl 2.5x BigDye buffer. If the template was a cosmid or if it contained secondary structures, the amount of DMSO was increased to 10% and 4 μl of BigDye (with 4 μl BigDye buffer) was used to get longer and better runs of sequence. Reactions were carried out in thin-walled 0.2 ml PCR tubes in a TGradient thermocycler (Biometra, Goettingen, Germany).

The cycling conditions were dependent on the T_m of the oligonucleotide primer. If the T_m was above 60°C, the conditions were: 30 cycles of 96°C for 30 seconds and 60°C for 1 minute. If the T_m was below 60°C it was: 30 cycles of 95°C for 30 seconds, 50-55°C for 30 seconds and 60°C for 1 minute. After the reaction, the DNA was precipitated by adding 80 μl of 96% (v/v) ethanol and 2 μl of sodium acetate/EDTA (final concentration of sodium acetate was 6mM and

EDTA was 100mM). Then the tubes were incubated on ice for 15-20 minutes and centrifuged for 15 minutes. Pellets were washed twice with 1 ml 70% (v/v) ethanol and air-dried. Sequence information was determined by an AB3730 genetic analyzer (Applied Biosystems) at the Molecular Biology Service Unit (Department of Biological Sciences, University of Alberta).

The sequence of the cosmid 6J11 (the cosmid that contained the clavam gene cluster) was determined by SeqWright (Houston, TX).

II.4.11 Software used for sequence analysis

DNA sequence alignments and assemblies were made using GeneTools version 2.0 (BioTools Inc., Edmonton, AB). This software was also used for identifying restriction sites and analyzing sequence chromatograms. Another software program that was used to analyze the sequence chromatograms was 4Peaks version 1.7.2 (Griekspoor and Groothuis 2006). DNA Strider version 1.2 (Marck 1988) was another program used for identifying the restriction sites and analyzing the DNA sequence. To identify the Open Reading Frames (ORFs) [by finding the possible start and stop codons and analyzing the usage of G+C in the third position of codons (Ishikawa and Hotta 1999), FramePlot 3.0beta (http://watson.nih.go.jp/~jun/cgi-bin/frameplot-3.0b.pl) was used. RNAfold (http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi) was used to predict the secondary structures and their minimum free energy. Using BLAST programs [Altschul et al. 1997, made available by the National Institute for Biotechnology Information (http://www.ncbi.nlm.nih.gov/BLAST/)] it was possible to compare the obtained sequences to the deduced DNA and amino acid sequences deposited in GenBank.

Multiple DNA sequences were aligned using ClustalW2 (http://www.ebi.ac.uk/Tools/clustalw2/index.html), which is a web based general purpose multiple sequence alignment program for DNA or proteins.

II.4.12 DNA probe generation and labeling

Different DNA probes were used to find the genes of interest in restriction digests of cosmid 6J11. Genomic DNA of *S. antibioticus* Tü1718 was used as a template for making probes for different genes. Probes were made for *ceaS* (primers ceaS2-RT-For and ceaS2-RT-Rev), *bls* (primers SGN12 and SGN18), *pah* (primers SGN04_FOR and SGN04_REV), and *cas* (primers SGN01 and SGN2). Another probe was made using primers for *cvm5* (primers cvm5 RT-forward and cvm5 RT-reverse) and *S. clavuligerus* gDNA as template for amplifying the probe. These probes were generated by PCR using primers that were designed from highly conserved regions of the corresponding genes in *S. clavuligerus.* The products were then eluted from agarose gels. After the sequence of the product was confirmed, the probes were non-radioactively labeled using DIG High Prime DNA Labeling and Detection Starter Kit I according to the manufacturer's directions (Roche).

II.4.13 Southern analysis

The transfer of DNA fragments from agarose gels to nylon membranes was based on a protocol by Sambrook et al. (1989). The restriction digest to be analyzed was separated on a 0.8% agarose gel for 20 hours at 19 volts. The gel was processed for transfer of DNA to Hybond-N nylon membranes according to the manufacturer's instructions (Amersham, Buckinghamshire, UK). The gel-

membrane sandwich was then assembled in a transfer apparatus containing 20 X SSC (3 M NaCl, 0.3 M $Na_3C_6H_5O_7$.$2H_2O$, pH 7.0) as transfer buffer. DNA was transferred to a nylon membrane through capillary action. After 15-17 hours the membrane was rinsed briefly in 2 x SSC and air-dried. In order to fix the nucleic acid on the membrane, it was baked in an 80°C vacuum oven for 2 hours. The membrane was then hybridized and visualized following the instructions provided with the DIG High Prime DNA Labeling and Detection Starter Kit I (Roche).

II. 5 Constructing a gDNA library

II.5.1 Preparing the gDNA

The gDNA was subjected to a partial digest with the enzyme Sau3A. The amount of enzyme and time of digestion was chosen to yield fragments of gDNA approximately as big as λ DNA (48 kb). After identifying the appropriate time and concentration of enzyme, 100μg of gDNA was subjected to digestion with 0.08 units of Sau3A for 5 minutes and then the digested DNA was ethanol precipitated.

After re-suspending in TE, the digest was centrifuged on a 10% – 35% sucrose gradient (18 ml of sucrose solution containing 22.5% sucrose, 1 M NaCl, 20 mM Tris pH 8.0, 2 mM EDTA) for 20 hours at 26,000 rpm. The sucrose gradient was divided to 18 fractions by repeatedly pipetting 1 ml aliquots from the top of the tube. Each fraction was ethanol precipitated and re-suspended in TE and run on a gel. Fractions containing the appropriate sized fragments were pooled together and ethanol precipitated. Then the DNA was treated with shrimp alkaline phosphatase (SAP).

II.5.2 Preparing the vector: pOJ436

The cosmid pOJ436 (Bierman et al. 1992) was chosen as the vector for library construction because of its ability to integrate into *Streptomyces* chromosomes but replicate as an independent plasmid in *Escherichia coli*. After extracting the cosmid from the host *E. coli,* it was treated with HpaI to linearize it. The linearized vector was then treated with SAP and then digested with BamHI to yield two fragments of about two and nine kbp, the cosmid vector arms.

II.5.3 Ligation, packaging and picking the library clones

Twenty microliters of the insert DNA (approx. 3.2 mg) was added to 5 µl (approx. 10 mg) of the vector arms. The mixture was heated to 45˚C for 15 minutes and then cooled to 4ºC. Two units of T4-DNA ligase and 3 µl 10X ligation buffer were then added to the mixture and it was incubated for 16 hours at 16^{o}C.

After ligation, 4 µl of the ligation mixture was added to a packaging reaction mixture tube (Gigapack® gold III kit, Stratagene, La Jolla, CA) and the tube was incubated at 20ºC for 2 hours. Five hundred microliters of SM buffer (5.8 g/L NaCL, 2 g/L $MgSO_4.7H_2O$, 0.05 M Tris.Cl pH 7.5, 0.1 g/L gelatin) and 20 µl chloroform were added to the mixture. The tubes were centrifuged for 5 minutes at maximum speed the supernatant was then transferred to another tube and stored at 4ºC.

In order to prepare the *E. coli* host strains the glycerol stocks were streaked on LB agar plates supplemented with 0.2% (w/v) maltose and 10 mM $MgSO_4$. The plate was incubated for 16 hours at 37ºC. After incubation, LB broth

supplemented with 0.2% (w/v) maltose and 10 mM $MgSO_4$ was inoculated with a single colony and was incubated at 37ºC for 4-6 h, until the OD_{600} reached 1.0. The bacteria were then centrifuged at 500 x g for 10 minutes and the pellet was re-suspended in half the original volume with 10mM $MgSO_4$. The OD_{600} was then brought to 0.5 with 10mM $MgSO_4$.

In order to evaluate the efficiency of the packaging, a control DNA was also used to transfect the *E. coli* cells. The control DNA was thawed on ice and 0.2 μg of it was added to a packaging reaction mixture tube and it was incubated for 2 hours at 20ºC. One milliliter of SM buffer and 40 μl chloroform were added to the tube and the tube was centrifuged. Two consecutive 10^{-2} dilutions were made in SM buffer. Ten microliters of the 10^{-4} dilution was added to 200 μl host cells (*E. coli* VCS257) and the tube was incubated at 37ºC for 15 minutes. Three milliliters of LB top agar (LB broth+0.7% agarose) was added to this and the mixture was divided and spread on three LB agar plates. Plates were incubated at 37ºC for 16 hours. After incubation more than 500 plaques were observed on one plate.

To titer the packaging reaction the same steps were taken as for the control DNA, except that the host strain was XL-1 Blue MR and the packaging reaction was diluted 10 and 50 times. From each dilution 25 μl was added to 25 μl of the host strain and the mixture was incubated for 30 minutes at 25ºC. Two hundred microliters of LB broth was added to each tube and the tubes were incubated at 37ºC for 1 hour. The tubes were mixed gently every 15 minutes. After incubation the contents were spread on LB agar plates supplemented with 25 μg/ml apramycin. The plates were incubated for 16 hours and the colonies were counted.

The number of colony forming units (cfu) was calculated to be 5.4 x 10^5 cfu/ml (or 1.35 x 10^2 cfu/μg of DNA). Twelve colonies were taken and the cosmids were extracted and digested with NcoI. This was done to ensure the colonies contain cosmids with inserts.

After titering the cosmid packaging reaction, appropriate volumes (500 μl of the 1:50 dilution of the packaging reaction was added to 500 μl of the cells) were spread on LB + 50 μg/ml apramycin plates (22 cm x 22 cm, 200 mL of medium/plate) to give about 2,000 colonies per plate. Colonies (2304 in total) were then picked [using a QPix robot (Genetix, Boston, MA)] and transferred to 384 wells microtiter dishes and were grown overnight. Each well contained 80μl LB medium+50μg/ml apramycin. After incubation for 16 hours, glycerol was added to give a final concentration of 20%, and the library was stored at -80˚C.

II.5.4 Gridding the library and hybridizing with the *cas3* probe

In order to screen the library for the positive cosmids, the library had to be gridded on a nylon membrane. The gridding of the library was done using the same robot that picked the colonies. The library was gridded onto two Performa 22 x 22 cm nylon membranes that were placed on plates containing LB agar + 50 μg/ml apramycin. After growing for 16 hours at 37˚C, the membranes were treated to lyse the colonies and fix the DNA (according to manufacturer's instructions), and then were marked. One of the membranes was stored as an extra copy.

The plasmid pSGNcas3 contained the *cas3* probe (Table II.3). The plasmid was digested with EcoRI to release the insert. After gel purifying the insert

(500bp PCR amplified *cas3*), it was used to probe the library. The probe was labeled using a non-radioactive labeling kit and was hybridized with the membrane. Clones that hybridized with the probe were grown on LB + 50 μg/ml apramycin and the cosmids were extracted. In order to find the enzymes that would digest the cosmids into suitable fragments, and to compare the pattern of digestion in different cosmids, six restriction enzymes were selected. Among these, NcoI, KpnI, XcmI and BamHI gave useable fragments.

In order to confirm the presence of *cas3* in the cosmids that hybridized to the probe, they were used as templates for PCR using the set of primers previously used to amplify *cas3* from the gDNA of *S. antibioticus* Tü1718.

II.6 Construction of *Streptomyces* mutants

Mutants of *S. antibioticus* Tü1718 were constructed following the REDIRECT© PCR-targeting protocol (Gust et al. 2003). PCR was used to amplify a gene disruption cassette, which is flanked by 39-nucleotide regions homologous to the DNA sequences that flank the gene to be disrupted. The cassette carries *aac3(IV)* which is an apramycin resistance gene. It also has *oriT* (the RP4 origin of transfer). Two FRT (FLP recognition target) sites flank the *aac3(IV)* gene-*oriT*. The PCR product was then electroporated into *E. coli* BW25113/pIJ790/14E10. This strain already carries the plasmid containing the gene targeted for disruption. When the PCR product is introduced in the *E. coli* BW25113/pIJ790, the λ Red proteins that are produced by the genes on pIJ790 catalyze the recombination between the region flanking the gene of interest on the plasmid and the 39-

nucleotide sequences flanking the cassette. The electroporated *E. coli* cells were plated on LB agar supplemented with 50 μg/ml apramycin and were incubated at 37˚C for 16 hours, and the plasmids were extracted and analyzed by restriction digestion. After confirming the plasmids were carrying target gene disrupted by the cassette, they were transferred to *E. coli* ET12567/pUZ8002, a non-methylating strain. From this strain the plasmid containing the disrupted gene was transferred to *S. antibioticus* Tü1718 via intergenic conjugation. Colonies resistant to apramycin were patched and then streaked for single colonies in order to make sure the colonies had double-cross overs. The gDNA of the colonies was then extracted and subjected to PCR analysis to ensure the cassette was inserted in the place of the gene to be disrupted. After verifying the mutants had the cassette at the right place, they were grown together with wild type *S. antibioticus* Tü1718 and antibiotic production was measured in the mutants and compared to the wild type.

cas3 and *ntr* were the two genes that were subjected to REDIRECT mutagenesis. SGN08_FOR and SGN08_REV were used to amplify the cassette for *cas3* and SGN09 and SGN14_nir K/O_FOR were used for *ntr* REDIRECT© mutagenesis.

After obtaining the mutant *Streptomyces* strains, the *cas3* mutants were analyzed using the combination of primers Red-SEQ-UP with SGN06_FOR (for the upstream region of the cassette) and Red-SEQ-DWN with SGN14_FOR (for the downstream region of the cassette). SGN06_FOR and SGN14_FOR were also used together to amplify the cassette in the gDNA of the mutants. The size of the product was compared to the size of the product obtained from the gDNA of wild

type *S. antibioticus* Tü1718. The *ntr* mutants were analyzed using primers Red-SEQ-UP with SGN16 (for the upstream region of the cassette) and Red-SEQ-DWN with SGN14_REV (for the downstream region of the cassette). Again, SGN14_REV and SGN16 were used together to amplify the cassette and the size was compared to the product from wild type *S. antibioticus* Tü1718. As a negative control the gDNA of wild type *S. antibioticus* Tü1718 was also used as template for PCR with the combination of primers.

II.7 Analysis of antibiotic production via high performance liquid chromatography (HPLC)

The amount of clavams in culture supernatants of *S. antibioticus* Tü1718 was determined using HPLC (Foulstone and Reading 1982; Paradkar and Jensen 1995). Supernatants were obtained from 48 hour cultures. Mycelia grown in the fermentation media and other insoluble materials were removed from culture samples by centrifugation of the broth for 5 minutes at 13000 rpm (using a microcentrifuge). Samples were diluted using Milli-Q™ water. From each culture supernatant sample, two sets of samples were prepared: one set of the samples was imidazole-derivatized [25 µl of 25% (w/v) imidazole was added], while the other set was underivatized (25 µl of Milli-Q™ water was added). After incubation for 15 minutes at 25˚C, the samples were centrifuged to remove any insoluble material. One hundred microliters of each sample was transferred to polypropylene spring inserts (Fisher Scientific, Pittsburgh, PA). The tubes were then placed into HPLC screw neck vials.

The chromatographic profile was determined by injecting 50 µl of each sample using a Waters 2690 Separations Module and analyzed with a Waters 996 Photodiode Array Detector (Waters, Milford, MA). A Phenomenex® Bondclone 10µ C18 (100 x 800 mm 10 micron) column (Phenomenex, Torrance, CA) was used. The mobile phase consisted of 100 mM NaH_2PO_4 + 6% (v/v) methanol in an isocratic elution run at 2 ml/minute. Each run was 15 minutes in length and derivetized clavams were detected by viewing the absorbance of samples at 311 nm. Waters Millenium[32] Software (version 3.20; Waters, Milford, MA) was used to analyze and process the data obtained from the samples.

TableII.1: *Streptomyces* strains used in this study

Strain	Description	Reference or source
S. clavuligerus NRRL 3585	Wild type	Northern Regional Research Laboratory, Peoria, Ill. USA
S. antibioticus Tü1718	Wild type	Townsend CA, Department of Chemistry, Johns Hopkins University
Δ*cas3::apra*	Mutant strain of *S. antibioticus* with an apramycin resistance cassette replacing *cas3*	This study
Δ*ntr::apra*	Mutant strain of *S. antibioticus* with an apramycin resistance cassette replacing *ntr*	This study

Table II.2: *E. coli* strains used in this study

Strain	Description	Reference or source
BW25113/pIJ790	Carries the λ Red genes encoded on plasmid pIJ790, under the control of an arabinose-inducible promoter	Gust et al.
DH5α	General cloning host	Gibco BRL
ET12567/pUZ8002	Methylation-deficient host carrying the plasmid mobilization functions on pUZ8002, used for inter-generic conjugation from *E. coli* to *Streptomyces*	Kieser et al. 2000
One Shot® TOP10	Host for TOPO™ vectors, used for TA cloning of PCR products	Invitrogen
XL1-Blue MR	Restriction minus derivative of Stratagene's XL1-Blue strain, useful for cosmid-based cloning, endonuclease (*endA*), and recombination (*recA*) deficient. The *hsdR* mutation prevents the cleavage of cloned DNA by the *Eco*K endonuclease system, and the *recA* mutation helps ensure insert stability. The *endA* mutation greatly improves the quality of miniprep DNA. Does not contain an F´ episome,	Stratagene

Table II.3: Plasmids and cosmids used in this study

Replicon	Description	Reference or source
plasmids pCR®2.1-TOPO®	*E. coli* phagemid for TA-cloning of PCR products	Invitrogen
pIJ773	*E. coli* plasmid carrying the 1398 bp REDIRECT© cassette [*aac(3)IV* and *oriT* flanked by FRT sites] as an *Eco*RI/*Hind*III insert	Gust et al. 2003
pUC119	*E. coli* phagemid cloning vector, Amp^R	Vieira and Messing (1987)
pUC120	*E. coli* phagemid cloning vector containing an *NcoI* site in the MCS, Amp^R	Vieira and Messing (1987)
pUZ8002	Carries the RK2/RP4 transfer functions for *in trans* conjugation of *oriT*-bearing replicons	Kieser et al. 2000
TOPO-*ntr*	*E. coli* pCR®2.1-TOPO® derivative carrying an *EcoRI/NotI* fragment with *ntr* in the middle	This study
pET26b	*E. coli* expression vector, kan^r	Novagen
pET26b-*cas3*	*E. coli* pET26b derivative carrying an *NcoI* fragment with *cas3* in the middle	This study
pSET-AT	*Streptomyces- E. coli* shuttle vector, $apra^r$, $thio^r$	This study
pSET-AT-*cas3*	pSET-AT derivative carrying the *cas3*-NcoI fragment	
pUWL-KS	*Streptomyces- E. coli* shuttle vector, with pIJ101 *ori* and blue white screening	Wehmeier 1995
pUWL-*cas3*	pUWL-KS derivative carrying the *NcoI* fragment containing *cas3*	This study
Cosmids pOJ436	λ*oriT* RK2 plasmid containing *aac(3)IV* (apramycin resistance)	Bierman et al. 1992

Table II.4: List of the oligonucleotide primers used in this study

Primer	Sequence (5'-3')	Use
CAN 68	CGCCAGGGTTTTCCCAGTCACGAC	Universal forward primer for sequencing the end of inserts cloned into pUC119 and pUC120
CAN 69	GAGCGGATAACAATTTCACACAGGA	Universal reverse primer for sequencing the end of inserts cloned into pUC119 and pUC120
CAN 82	GTAATACGACTCACTATAGGGC	Sequencing the T7 strand from pOJ436
CAN 168	TGATGCCGGCCACGATGCG	Sequencing the T3 strand from pOJ436
ceaS2-RT-For	AGGCCGCGTCGATTCTCTTC	Forward primer for ceaS2 RT-PCR
ceaS2-RT-Rev	CGGCGGGTTGGGGACGGT	Reverse primer for ceaS2 RT-PCR
cvm5 RT-forward	TTTCGAGGTCTCCAACAACCCGATCA	Forward primer for cvm5 RT-PCR
cvm5 RT-reverse	TTGATGATCTCCCAGTGCTCCTCGAA	Reverse primer for cvm5 RT-PCR
Red-SEQ-UP	CTGCAGGTCGACGGATCC	Reverse primer for checking the presence of the REDIRECT© cassette at the right position-for all the cassettes
Red-SEQ-DWN	CGAAGCAGCTCCAGCCTAC	Forward primer for checking the presence of the REDIRECT© cassette at the right position-for all the cassettes
SGN01_cas1 FOR	GCGCTCGGGCACGGTCTACCAC	Forward primer for amplifying *cas3*
SGN2_cas1 REV	GCGGTGCAGCCAGCGGTCCTTC	Reverse primer for amplifying *cas3*
SGN03-cas3-out-FOR	CGTGCTGTCGAAGGCGCTC	Sequencing downstream of *cas3*
SGN03-cas3-out-REV	GCAGGGTACTCGGAGGTCTTCG	Sequencing upstream of *cas3*
SGN04_pah3_FOR	CGACGGGGGCACCAGCTACC	Forward primer for amplifying *pah3*
SGN04_pah3_REV	CCGGGCGTAGTCGAGCGAG	Reverse primer for amplifying *pah3*
SGN05_cas1out_FOR	CACCGACCGCAATGGACAGCTCTC	Sequencing downstream *cas3*
SGN05_cas1out_REV	CGCGCAGCTCGGGGCCGTAC	Sequencing upstream of *cas3*
SGN06_pah out_FOR	CCACAACCCGCAGCCCGACT	Sequencing downstream of *pah3*
SGN06_pah out_REV	CGTGGATGAGCCCGGATTCG	Sequencing upstream of *pah3*
SGN07_pah further out_FOR	TGGGCCTGCACACCGGTTAC	Sequencing downstream of *pah3*
SGN08_cas3 K/O_FOR	TTCCTCGACGCCGCGCACGAAGCCGCCGGC GACCTGCCCATTCCGGGGATCCGTCGACC	Forward primer for *cas3* REDIRECT© mutagenesis
SGN08_cas3 K/O_REV	CCACAGGGTTCTGCGCGGGTGCTAGCGGCG	Reverse primer for *cas3* REDIRECT©

	TGGCGAGAATGTAGGCTGGAGCTGCTTC	mutagenesis
SGN09_nir K/O_REV	CGTCATGAGCGGCCGTGTCACCAGCTGTCG AGGTGGAGGTGTAGGCTGGAGCTGCTTC	Reverse primer for *ntr* REDIRECT© mutagenesis
SGN10_nir out_FOR	GCTGATTCCGACGGCGTACTC	Sequencing downstream *ntr*
SGN11_nir_upstream	GTCGGGGTTCAGGTTCAGCAGA	Sequencing upstream *ntr*
SGN12_bls3_upstream	ACCCCTGACCCCTCGTGGACG	Sequencing upstream *bls3*
SGN14_oat out_FOR	ACACCCGTGGAGGCGATGAG	Sequencing upstream *pah3*
SGN14_oat out_REV	ACACGGACATCGACCAGGACAA	Sequencing downstream *pah3*
SGN14_nir K/O_FOR	CCATGGCTCTGCTGAACCTGAACCCCGACG AACTGCTCAATTCCGGGGATCCGTCGACC	Forward primer for *ntr* REDIRECT© mutagenesis
SGN15_bls3 furthermore out	CGCGGGCACCTGATGGACAC	Sequencing upstream of *bls3*
SGN16_nir K/O check_upstream	GGACGCCTGTTGCTGGTGTGGA	Forward primer for checking the presence of the REDIRECT© cassette at the right position for *ntr*
SGN17_bls3 evenfurther out_REV	CGAGCTGTACAACCGGGACGA	Sequencing upstream of *bls3*

Table II.4: continued		
SGN18_bls3 probe making_FOR	CGTCGAGACAGTCCGGCAGGT	Forward primer for amplifying a probe for *bls3*
SGN19_ceaS3_upstream	TGGTGGGTGTCGTTGGGGAAG	Sequencing upstream of *ceaS3*
SGN20_stemloop breaking_FOR	GCACCCGCGCAGAACCCT	Forward primer for sequencing the secondary structure between *cas3* and *oat3* from within the loop
SGN20_stemloop breaking_REV	GGTCGTGCGGAGCGTGCC	Reverse primer for sequencing the secondary structure between *cas3* and *oat3* from within the loop
SGN21_aldo/keto_out	GCTCCTTCTACGGGCTCTCGG	Sequencing upstream of the aldo/keto reductase
SGN22_upstream ceas3	GGGGAAGTGGAAGGGCTGGTG	Sequencing upstream of *ceaS3*
SGN22_downstream ceaS	CGCCACCTCCATACTCGACCG	Sequencing downstream of *ceaS3*
SGN23_into atr	TGCGCTGTCGACGAGGTGCCC	Sequencing the aminotransferase
SGN24_atr probe making	GCAGGGCATGACCGGCGAGGA	Primer for amplifying a probe for the aminotransferase
SGN25_soh upstream	CCTCGCTCGTTTTCTCATGGGA	Sequencing upstream of the homologue OF *orfA*
SGN26_downstream akr	CGGGAGACGGAGGCGGCGGT	Sequencing downstream of the aldo/keto reductase
SGN27_DnStr the new gene	TACCTGATACGGGACTATCAGCTC	Sequencing downstream of the homologue of *c7p*
SGN28_UpStr the new gene	GCTTCATGCAGTTCGTGGATCAGC	Sequencing upstream of the homologue of *c7p*
SGN29_5’ of atr	TTCAACGCCCTGCACACCA	Sequencing the 5’ end of the aminotransferase

III. Results

III. 1 Amplifying *cas3* from *S. antibioticus* Tü1718, to serve as a *cas3* probe

To better understand the late steps in clavam biosynthesis, the genes for clavam biosynthesis were examined in a 5*S* clavam producing species other than *S. clavuligerus.* For this purpose *S. antibioticus* Tü1718 was chosen for analysis because some preliminary information was available regarding biosynthesis of 5*S* clavam metabolites in this organism. The presence of the enzyme, Cas3, in *S. antibioticus* Tü1718 was shown by Janc et al. (1995). The amino acid sequence of the N-terminal end of Cas3 was shown to be similar to Cas1 and Cas2 (more similar to Cas1) (Janc et al. 1995). Therefore the search for the 5*S* clavam biosynthetic gene cluster in *S. antibioticus* Tü1718 started with investigating the gene *cas3*. Because Cas3 was very similar to Cas1 and Cas2 it was expected that the DNA sequence would be very similar as well. Therefore, two primers SGN01 and SGN02 (Table II.4) were designed based on the most conserved regions of *cas1* and *cas2*.

Using the primers SGN01 and SGN02 and the gDNA of *S. antibioticus* Tü1718, the *cas3* gene was amplified (Figure III.1). As a positive control the primers were also used for amplifying the *cas* gene from *S. clavuligerus* gDNA. The expected size of the product was about 500 bp and as shown in Figure III.1, the product was obtained and it was the same size as the product from *S. clavuligerus*. In order to prove that the PCR product was really *cas3,* the product was gel purified and sequenced. Sequence analysis revealed very high similarity

(more than 70% at nucleotide level) of this product to *cas1* and *cas2*. Figure III.2 shows the alignment of the three *cas* genes (*cas1* and *cas2* from *S. clavuligerus* and *cas3* from *S. antibioticus* Tü1718). This product was cloned and labeled using the non-radioactive DNA labeling kit, and was used as a probe to identify cosmids in a gDNA library that would contain the clavam cluster.

As described in the Materials and Methods, a library was made from the gDNA of *S. antibioticus* Tü1718, and in order to make it easier to locate the DNA fragment that contained the 5*S* clavam cluster, the library was gridded on a nylon membrane. The cosmid vector used for making the library was pOJ436, which is a *Streptomyces-E. coli* shuttle vector and has the ability to integrate into the *Streptomyces* chromosome. It is also possible to use this vector to transfer pieces of the DNA of *S. antibioticus* Tü1718 into other *Streptomyces* such as *S. clavuligerus,* which is one of the reasons it was chosen.

The gDNA library carried on a nylon membrane was hybridized with the *cas3* probe. After hybridizing, eight cosmid clones were shown to contain *cas3* (Figure III.3). The cosmid clones were identified and the *E. coli* cells containing them were grown on LB for 16 h at 37˚C. The cosmids were then extracted following the alkaline lysis method (explained in the Materials and Methods section). In order to find enzymes that would cut the cosmids into suitable fragments, and to compare the pattern of digestion in different cosmids, the cosmids were digested with six different restriction enzymes. Among these enzymes, NcoI, KpnI, XcmI and BamHI were judged to have the best digestion pattern, i. e. these enzymes digested the cosmids into fragments that were suitable

for cloning. Therefore these four enzymes were selected for further analyzing the two chosen cosmids, 2L10 and 3D23. These cosmids were chosen because they showed the best digestion patterns and their digests shared common bands suggesting that they carried overlapping DNA fragments. Figure III.4 shows a digestion pattern of these two cosmids after digestion with BamHI, KpnI, NcoI and XcmI.

In order to determine the sequence of all of the genes involved in 5S clavam biosynthesis in *S. antibioticus* Tü1718, total sequence analysis of one or more cosmid inserts would be required. In order to decide which cosmids might be the best candidates for complete sequence analysis, a number of preliminary studies were required. In the first instance it was necessary to determine if the *cas3* gene was clustered with other clavam biosynthetic genes, if the late as well as the early genes of clavam biosynthesis are clustered together in *S. antibioticus* Tü1718 or if they are in separate clusters as they are in *S. clavuligerus*. Finally, it was important to determine if the *cas3* gene was located near the center of any of the cosmid inserts, because those cosmids would be more likely to contain most or all of the gene cluster.

III. 2 Analyzing the positive cosmid clones from the library

The next step in locating the 5*S* clavam gene cluster was to identify the fragments in the restriction digests that would contain *cas3* and flanking regions. The two chosen cosmids were digested with the enzymes mentioned in the previous section. The digests were then subjected to Southern analysis. Figure

III.4 shows the restriction digest pattern of 2L10 and 3D23 on an agarose gel stained with ethidium bromide, with the Southern blot corresponding to it.

The NcoI digest of the cosmid 3D23 had a 4.5 Kbp fragment that hybridized with the *cas3* probe, whereas the NcoI fragment that hybridized with the *cas3* probe from 2L10 was more than 10 Kbp. The NcoI sites are shown in Figure III.5.

This suggested that the hybridizing NcoI fragment from 2L10 might represent one end of the cosmid insert. Therefore the hybridizing NcoI fragment from 3D23 was selected for further analysis. The KpnI digest of both cosmids contained a 2.5 Kbp fragment that hybridized with *cas3* probe. Cosmid 2L10 was digested with KpnI. After digesting the cosmids DNA fragments were separated by agarose gel electrophoresis and the desired fragments were extracted from the gel and purified. Purified fragments were sub-cloned by ligating to pUC119 (for the KpnI fragment) and pUC120 (for the NcoI fragment) and inserting the construct into electrocompetent *E. coli* DH5α cells.

After ligation, five NcoI clones and six KpnI clones were obtained. However the sizes of the DNA inserts were slightly different, both for some of the NcoI sub-clones and for some of the KpnI sub-clones (Figure III.6). The clones were therefore subjected to Southern analysis using the *cas3* probe in order to find those that would contain the *cas3* gene. As Figure III.7 shows, only some of the clones hybridized with the probe, showing that only some contained *cas3.*

III.2.1 Sequence analysis of the NcoI clone

One of the NcoI clones that hybridized was selected for further sequencing. The two ends of the clone were sequenced using CAN68 and CAN69 universal

primers and cycle sequencing using the BigDye terminator kit (Table II. 4). Primers SGN01, SGN03_FOR, SGN06_FOR and SGN04_FOR (Table II. 4) were also used to obtain sequence information from within the clone. Figure III.8 shows a diagrammatic representation of the genes identified as a result of these sequencing reactions and assembling the DNA sequences obtained from different primers. One of the ends of the clone contained an ORF, which showed 72% identity to a putative nitroreductase-type oxidoreductase from *Thermomonospora curvata* DSM 43183. The sequence from the other end of the clone showed high similarity to *bls1* and *bls2* (from *S. clavuligerus*) and to *pah1* and *pah2* (from *S. clavuligerus*). SGN01 (the forward primer for amplifying *cas3*) and SGN04 (the forward primer for amplifying *pah3*) were used to obtain additional sequence for *cas3* and *pah3* respectively. The sequence obtained from SGN06 showed high similarity to *cas1* and *cas2* (from *S. clavuligerus*). Interestingly, the sequence obtained using SGN03_FOR (downstream of *cas3*) stopped at a certain point every time. Different amounts of BigDye or DMSO, and heating or digesting the sequencing template did not help to obtain sequence beyond that point. This suggested the presence of a secondary structure downstream of *cas3.* The positions of the primers used are shown in Figure III.5.

Another primer, SGN10, was designed to sequence the region downstream of the nitroreductase gene (*ntr*) (Figure III.5). This primer did not yield a long sequence and, similar to *cas3,* the sequence stopped abruptly at a certain point. The inability to sequence across this region suggested the presence of another secondary structure, this time downstream of *ntr.* The total assembled sequence

that was obtained from all primers was about 1.5 Kbp less than the estimated size of the NcoI fragment. Therefore, there must be two regions of secondary structure surrounding an internal un-sequenced region, rather than just one stem-loop separating *cas3* from *ntr*.

III.2.2 Attempts to sequence the secondary structures and the region between them

After finding that there is a region in the insert that is flanked by two secondary structures (most probably two stem-loops), I searched for restriction enzymes that would cut close to the secondary structure and release that 1.5 Kbp fragment. If successful, it would then be possible to clone this smaller fragment and sequence it. Digestion at restriction sites in or very near the stem-loops might relax the secondary structure making it possible to sequence through them.

According to the partial DNA sequence of the NcoI fragment, there were sites for PvuII very close to the stem-loops. Therefore this made PvuII an excellent candidate for taking out the stem-loop region. If there were no PvuII sites in the un-sequenced region between the stem-loops, a PvuII digest should give fragments of about 200, 700, 1500 and 2500 bp. The fragment that contained the stem-loops would be the 1.5 Kbp fragment. The PvuII sites are shown in Figure III.5.

The NcoI clone was initially digested with NcoI to release the 4.5 Kbp insert. After that it was digested with PvuII and the expected fragments were observed (Figure III.9). The 1.5 Kbp PvuII fragment was cloned into the SmaI site of pUC120 (PvuII and SmaI both create blunt ends when they digest DNA). The

ends of this clone were sequenced using the universal primers, but once again the sequences stopped at the same points as they did for sequences from the full-length NcoI clone obtained using SGN03 and SGN10 primers.

The next strategy was to further digest the PvuII fragment and sub-clone smaller pieces into pUC119. If sequence information could be obtained from between the two stem-loops it could be used to design primers that would approach each stem-loop from within the un-sequenced region and in that way complete the sequence.

After digesting the clone with a number of enzymes commonly found in cloning vectors, SmaI was selected. This enzyme released a 750 bp fragment from the PvuII sub-clone. None of the other enzymes tested were able to release a fragment from the clone. The 750 bp fragment was cloned into the SmaI site of pUC119 and the insert was fully sequenced using the universal primers. The sequence obtained showed high similarity to *oat1* and *oat2* from *S. clavuligerus*. Based on this sequence, primers were designed (SGN14_FOR and SGN14_REV, Table II. 4) and used for further sequencing of the PvuII clone. Figure III.5 shows the sites for these primers. The sequences obtained from these primers were assembled with the 750 bp from the SmaI fragment. After assembling all the sequences the full *oat3* gene sequence was obtained but there was still about 200 bp missing from the PvuII sub-clone, showing that the combined length of the stem-loops flanking *oat3* was about 200 bp.

Using SGN14_FOR with SGN03_FOR it should be possible to amplify the stem-loop between *cas3* and *oat3,* and using SGN14_REV with SGN10 the other

stem-loop could be amplified. PCR products were obtained and subjected to sequencing. In this way, it was possible to read through the sequence of the stem-loop between *oat3* and *ntr,* to give a secondary structure that was 77 bp long, leaving about 125 bp for the other stem-loop. The sequence obtained from the stem-loop between *cas3* and *oat3* however, stopped at the same point as seen previously. So it was not possible to sequence the PCR product from this stem-loop structure.

The final approach for sequencing the larger stem-loop was to use 7-deaza-dGTP in the PCR amplification step to produce a template with weaker secondary structure (explained in II.4.9). In order to compare the effect of the concentration of 7-deaza-dGTP used in the PCR on the quality of the sequence obtained from the stem-loop, four different concentrations of this nucleotide were used to make PCR products. As a result, there were eight sequence chromatograms obtained from this stem-loop (each concentration of 7-deaza-dGTP was sequenced by forward and the reverse primers). Interestingly, there was a significant amount of difference between the sequence files. The sequences were aligned and the most probable sequence of the structure was decided. Figure III.10 shows an alignment of the different sequences together with a sequence of the stem-loop that was determined later in the study by pyrosequencing of the entire cosmid.

The predicted structures of the stem-loops are shown in Figures III.11 and III.12. The free energy of the stem-loop between *cas3* and *oat3* was -87.2 Kcal/mol and for the other stem-loop it was -55.5 Kcal/mol, as predicted using RNAfold.

III.2.3 Further analysis of the insert within cosmid 6J11

The ends of the eight cosmid clones that hybridized with the *cas3* probe were sequenced using primers CAN82 and CAN168 (Table II.4). Among these eight cosmids, cosmid 6J11 and 1F10 gave the best quality sequence and digest patterns. In addition, according to Figure III.4, the 5*S* clavam cluster seemed to be at the end of the cosmid insert for 2L10. These reasons led to cosmids 6J11 and 1F10 being chosen for further analysis.

Using primers SGN04_FOR and SGN04_REV (which were designed based on homologous regions of *pah1* and *pah2*) a PCR product was prepared using *S. antibioticus* Tü1718 gDNA as template. The PCR product was sequenced, and after confirming that the sequence showed homology with *pah1* and *pah2*, it was used as a probe for *pah3* (Figure III.5). A Southern analysis was carried out on digests of 6J11 and 1F10 using the *pah3* probe. These two cosmids were digested with KpnI and NcoI. The NcoI fragments that hybridized were 4.5 Kbp for both cosmids. Figure III.13 shows the ethidium bromide stained gel and the hybridized nylon membrane for this Southern analysis.

Sequence analysis of the 4.5 Kbp NcoI clone showed that it had part of *bls3* on one end. Therefore the goal was to find the adjacent NcoI fragment that would contain the rest of *bls3* and the upstream genes. Based on the sequences obtained from the NcoI clone, primers were designed to determine the DNA sequence of *bls3*. One of these primers, SGN15, was used to walk upstream of the NcoI fragment on the cosmid, and from the sequence information obtained, another primer, SGN18, was designed and used for PCR with SGN15 to make a *bls3* probe. The primers and the probe are shown in Table II.4 and Figure III.5.

In order to get all of the NcoI fragments from 6J11 cloned into pUC120, the cosmid was digested with NcoI and the fragments were shotgun cloned into the vector. Forty-eight colonies were picked and were subjected to Southern analysis using the *bls3* probe. One of the positive clones that hybridized with the *bls3* sub-clone apparently contained a large insert as judged by electrophoresis of the uncut plasmid compared to pUC120. Given the size of the plasmid it was expected that one end of the insert would give sequence of *bls3* and the other end would be new sequence, perhaps showing similarity to *ceaS* (in *S. clavuligerus ceaS* is immediately upstream of *bls* (Figure. I.7). The sub-clone was end sequenced using primers, CAN68 and CAN 69, and the sequence of one end showed similarity to *bls* from *S. clavuligerus,* but the other end was similar to a peptidoglycan binding protein. This was at first interpreted to mean that *ceaS3* was not immediately upstream *bls3* and that the order of the 5*S* clavam cluster genes might be different in *S. antibioticus* Tü1718 than in *S. clavuligerus*, but the results from the restriction digest of the sub-clone were also unexpected. The NcoI restriction digest only showed a 1.5 Kbp fragment and pUC120, whereas the mobility of the uncut plasmid suggested a much larger size.

In order to answer these questions, the 48 NcoI clones were subjected to Southern analysis with a *ceaS3* probe (constructed using ceaS2-RT-For and ceaS2-RT-Rev, see Table II.4 and Figure III.19, with *S. clavuligerus* gDNA as template). The same nylon membrane that had been probed previously with a *bls3* specific probe was de-colorized by placing the membrane in hot dimethylformamide. The membrane was then re-hybridized with the *ceaS3* probe,

but the same clones hybridized as did for *bls3* probe (data not shown). To clarify the matter further, Southern analysis of NcoI and KpnI digests of cosmids 1F10 and 6J11 was conducted with the *ceaS3* probe after decolorizing the blot previously hybridized with the *pah3* probe (Figure III.13). This new Southern blot and its corresponding gel are shown in Figure III.14. From the KpnI digests a 6.5 Kbp band hybridized with the *ceaS3* probe and from the NcoI digests a 3 Kbp fragment hybridized with the probe. Based on these analyses, it was deduced that the NcoI sub-clone actually contained two NcoI fragments, a 1.5 kb fragment, and a 3 kb fragment that happened to co-migrate with pUC120 upon electrophoresis. The KpnI fragment from 6J11 was gel purified and cloned for further analysis and sequencing.

III.2.3.1 Sequencing the KpnI fragment and the genes upstream of bls3

After finding that a 6.5 Kbp KpnI fragment is likely to contain *ceaS3* (and therefore the upstream genes from the cluster), the band was gel purified and cloned into pUC119. The ends were then sequenced using the universal primers. One end gave sequence from *pah3* and the other end gave sequence similar in part to a 4-aminobutyrate aminotransferase (*atr*, similar to *Nakamurella multipartita* DSM 44233), and in part to the 5' end of an aldo-keto reductase (*akr*). The gene *atr* did not show any similarity to any of the clavam genes in *S. clavuligerus*. However, *akr* showed 70% identity to *cvm1* from the *S. clavuligerus* clavam gene cluster.

Based on the obtained sequences, new primers were designed to sequence further into the KpnI fragment. Primer SGN_21 was designed to sequence further

into the *akr* and primers SGN22_UPSTREAM and SGN22_DOWNSTREAM were used to sequence the regions flanking *ceaS3*. Using these primers and the primers used to amplify the *ceaS3* probe (ceaS2-RT-For and ceaS2-RT-Rev), most of the sequence of the KpnI fragment was obtained. In addition to *atr* and *akr*, another ORF was found. BLAST analysis showed that this ORF encoded a putative serine OH-methyl transferase and was 74% identical to *orfA* (at nucleotide level) from the *S. clavuligerus* paralogue gene cluster.

Primers were designed to amplify a probe for *atr* (SGN23 and SGN24, Table II.4, Figure III.5) in order to analyze the sequence upstream of the KpnI fragment. Using this probe a Southern analysis was done on the NcoI digest of cosmid 6J11. A 3 Kbp NcoI fragment hybridized with the *atr* probe. This fragment was sub-cloned into pUC120 and the ends were sequenced (using universal primers), one end of the sub-cloned fragment gave sequence from *atr* as expected. The sequence of the other end of the clone showed similarity to a peptidoglycan binding protein. Furthermore, the gene upstream from *atr* was a hypothetical protein. These observations suggested that *atr* was one end of the 5*S* clavam gene cluster and it would be unlikely to find more 5*S* clavam genes upstream.

Since none of these genes have counterparts in the clavam biosynthetic gene clusters in *S. clavuligerus,* on this basis it was concluded that cosmid 6J11 was likely to contain most or the entire 5*S* clavam gene cluster from *S. antibioticus* Tü1718, and no rearrangements were evident. Therefore cosmid 6J11 was sent to a DNA sequencing laboratory for complete analysis of the insert.

III.2.3.2 Sequencing the whole cosmid 6J11

After confirming that cosmid 6J11 contained the early genes and at least some of the late genes for 5*S* clavam production in *S. antibioticus* Tü1718, the whole cosmid was sent to SeqWright for sequencing. The obtained sequence was analyzed by FramePlot (to locate the ORFs) and BLAST (to analyse the ORFs that were found using FramePlot). FramePlot is a web-based software, which is able to locate the ORFs in a given DNA sequence by defining the start and stop codons and by determining the %G+C content in each position in the codons of ORFs. The map of the ORFs is shown in Figure III.15B. As the map shows, the order of the early genes and their orientation was the same in *S. antibioticus* Tü1718 as in the clavulanic acid gene cluster in *S. clavuligerus* (Figure III.15A). The sequence obtained from the two stem-loops (using the strategies explained previously in III.2.2) was compared to the sequence from SeqWright. The two sequences were identical except for two bases in each stem-loop. Table III.1 shows a list of the ORFs present in the 5*S* clavam gene cluster of *S. antibioticus* Tü1718 and their putative roles.

A surprising feature in the 5*S* clavam gene cluster of *S. antibioticus* Tü1718 was the presence of two *pah* genes (shown as *pah3* and *pah4* in the cluster). *pah* is the only gene present in two copies in this cluster. *pah4* was in the reverse orientation compared to *pah3* and was 72% identical (at nucleotide level) to *pah3*. It was also slightly shorter than *pah3*.

In addition to orthologues to the early genes of *S. clavuligerus,* the 5*S* clavam cluster in *S. antibioticus* Tü1718, also contained orthologues to some late

genes. *ctr* (clavam transcriptional regulator) showed 66% identity at the nucleotide level to *c7p* in the paralogue gene cluster of *S. clavuligerus,* although over a very short stretch of nucleotides. The gene *soh* (serine OH-methyl transferase) was 74% identical to *orfA* at nucleotide level. *akr* (aldo-keto reductase) was 70% identical to *cvm1* from the clavam gene cluster of *S. clavuligerus*. The gene *atr* encodes a putative aminotransferase and as mentioned earlier does not show similarity to any of the genes in *S. clavuligerus*. However, it is proposed to be a functional counterpart of *c6p* (which is also an aminotransferase) from the paralogue gene cluster of *S. clavuligerus.*

In addition to the genes mentioned above, there were two genes that apparently reside in the 5*S* clavam cluster of *S. antibioticus* Tü1718 and yet do not have any counterparts in *S. clavuligerus.* These two were *ntr* (nitroreductase) and *lig* (ligase). Although the extent of the 5*S* clavam cluster in *S. antibioticus* Tü1718 is not known, the fact that these genes were flanked by genes showing high similarity to genes for clavam biosynthesis in *S. clavuligerus* strongly suggests they are within the cluster.

III. 3 Mutagenesis of *cas3* and *ntr*

While sequence analysis strongly suggested that the cloned DNA contained in cosmid 6J11 carried the 5*S* clavam gene cluster from *S. antibioticus* Tü1718, more definitive proof was needed and so *cas3* and *ntr* were disrupted by PCR targeting mutagenesis (explained in the previous chapter).

III. 3. 1. The *cas3* mutant

REDIRECT™ is a trade name for a set of PCR targeting tools used widely for gene disruption in *Streptomyces* spp. In a typical application, a genomic DNA fragment carried in the *E. coli* cosmid vector, Supercos, is used as the platform for undertaking the mutagenesis, and an apramycin resistance cassette is used to replace the gene to be mutated. However, in my initial attempts at mutation of *cas3*, cosmid 6J11, a pOJ436-based construct was used as the platform, and this cosmid already contained an apramycin resistance cassette. Therefore a streptomycin resistance cassette was used as an alternative disruption marker, but the attempts were not successful because the cassette could not give useable streptomycin resistance to *S. antibioticus* Tü1718. In order to use the apramycin resistance cassette as a disruption marker, the 4.5 Kbp NcoI fragment (that contained part of *bls3* and all of *pah3, oat3, cas3* and *ntr*) cloned in pUC120 was used instead of the whole cosmid as the platform for PCR targeting mutagenesis with primers SGN08_cas3 K/O_FOR and SGN08_cas3 K/O_REV used to generate the mutagenic PCR product. The resulting Δ*cas3*-pUC120-Nco::apra plasmid was then transformed into *E. coli* ET12567-pUZ8002, and from there, introduced into wild type *S. antibioticus* Tü1718 by interspecies conjugation. Since pUC120 is not stable in *Streptomyces* spp., apramycin resistant exconjugants should only arise through single or double cross-over of the mutant plasmid with the *S. antibioticus* Tü1718 chromosome via homologous recombination. These exconjugants were then allowed to sporulate to encourage loss of integrated plasmid through a double cross-over, and finally the resulting spores were checked for apramycin resistance as indicative of potential *cas3*

double cross-over mutants. However, single crossover mutants would also be resistant to apramycin, and therefore PCR analysis was carried out on the gDNA of the putative mutants to verify the presence of the mutation and the absence of the wild type gene. In order to verify that the target gene was replaced by the antibiotic resistance cassette, the primers RED_SEQ_UP and RED_SEQ_DOWN (Table II.4), which are specific for the apramycin resistance cassette, were used with other gene specific primers that would amplify the regions upstream or downstream of the cassette.

Initially, the PCR analysis gave band patterns suggesting that the putative mutants were a mixture of both single and double cross-over mutants (data not shown). Therefore another round of sporulation was undertaken to obtain the desired double crossover mutants. Genomic DNA from four potential Δ*cas3* mutants was analyzed by PCR and gave band patterns that confirmed that they were true double crossover mutants. Figure III.16 shows the PCR products from the gDNA of the *cas3* mutant and wild type *S. antibioticus* Tü1718. Use of the primer pair, SGN6_FOR and RED_SEQ_UP, should amplify a fragment extending from within the apramycin resistance cassette into the *cas3* upstream region and give a band of 546 bp for *cas3* mutants, and no band for the wild type. The primer pair, SGN6_FOR and SGN14_FOR, encompasses the *cas3* gene, and should give a band of 2.4 Kbp for *cas3* mutants, and a band of 1.8 Kbp for the wild type. The length of the product from the mutants was expected to be bigger than the wild type because the size of the cassette is bigger than *cas3.* According to the gel, the four putative mutants all did contain the cassette in the place of

cas3. The set of primers (RED_SEQ_DOWN and SGN14_FOR), should amplify a 514-bp fragment extending from within the apramycin resistance cassette into the *cas3* downstream region, but it did not give a very prominent product (data not shown). Most probably this is because the stem-loop that was in the downstream region between *cas3* and *oat3* blocked the amplification.

The *cas3* mutants and the wild type strain were then grown under 5S clavam production conditions and the culture supernatants were derivatized with imidazole and assayed by HPLC. Figure III.17 shows the HPLC chromatogram from the wild type *S. antibioticus* Tü1718 and a representative *cas3* mutant. No authentic standards are available for valclavam or 2-hydroxyethylclavam, but previous HPLC-mass spectrometric analysis of culture supernatants of *S. antibioticus* Tü1718 indicated that the wild type strain produces large amounts of valclavam as well as smaller amounts of another clavam-type metabolite that does not correspond to 2-hydroxyethylclavam. No 2-hydroxyethylclavam was detected. Based on the retention time for valclavam determined in these previous studies, a large peak due to valclavam was seen in the wild type, but no clavam peaks were seen in the *cas3* mutant. The other unknown clavam product seen previously was again detected in the wild type supernatant, and 8-hydroxyethylclavam was again not observed under these growth and HPLC conditions. Deletion of *cas3* resulted in the complete loss of clavam production in *S. antibioticus* Tü1718.

III. 3. 2. The *ntr* mutant

In order to mutate *ntr* the cosmid 6J11 was digested with NotI and EcoRI. The DNA sequence of the cosmid insert indicated that a 4.2 Kbp DNA fragment

containing the *ntr* gene would result from this double-digestion. Therefore after digesting the cosmid with these two enzymes the 4.2 Kbp fragment together with a 3.9 Kbp fragment that was not well resolved were purified from the agarose gel. After sub-cloning these fragments into pCR®2.1-TOPO®, clones were analyzed by restriction digestion to choose the right clone (the one with the 4.2 Kbp insert). The construct, TOPO-*ntr,* was then used as the platform for PCR targeting mutagenesis with primers SGN14_nir K/O_FOR and SGN09_nir K/O_REV used to generate the mutagenic PCR product. The resulting *Δntr*-TOPO-Eco-Nco::apra plasmid was then transformed into *E. coli* ET12567-pUZ8002, and from there, introduced into wild type *S. antibioticus* Tü1718 by interspecies conjugation. Mutants were selected as described above for the *cas3* mutant, and then gDNA from putative *ntr* mutants was screened by PCR.

The PCR analysis of the gDNA of the putative *ntr* mutants and wild type *S. antibioticus* Tü1718 is shown in Figure III.18. The primer pair, SGN16 and SGN14_REV, encompasses the *ntr* gene, and should give a band of about 2.5 Kbp for *ntr* mutants (Lanes 5 and 7), and a band of about 1.1 Kbp for the wild type (Lanes 3, 4, 6 and 8). The primer pair, SGN16 and RED_SEQ_DOWN, should amplify a fragment extending from within the apramycin resistance cassette into the *ntr* downstream region and give a band of about 400 bp for *ntr* mutants, and no band for the wild type. The primer pair, SGN14_REV and RED_SEQ_UP, should amplify a fragment extending from within the apramycin resistance cassette into the *ntr* upstream region and give a band of about 400 bp for *ntr* mutants, and no band for the wild type. However, the products were not very

obvious for this primer pair, again presumably due to the presence of the stem-loop in the upstream region. The analysis proved that only two of the five mutants were true double cross-over mutants.

The true *ntr* mutants and the wild type strain were then grown under 5S clavam production conditions and the culture supernatants were derivatized with imidazole and assayed by HPLC. Figure III.19 shows the HPLC chromatogram from the wild type *S. antibioticus* Tü1718 and a representative *ntr* mutant. In the case of these *ntr* mutants a complete loss of antibiotic production was again observed.

Table III.1: List of the ORFs found in the 5*S* clavam biosynthetic gene cluster in *Streptomyces antibioticus* Tü1718.

ORF	Length (aa)	Most similar protein (by Blast analysis)	% Similarity	Predicted function	Counterpart in *S. clavuligerus* clavam biosynthesis
psr	843	SanG from *Streptomyces ansochromogenes*	44% over 835 aa	Pathway specific transcriptional regulator	
atr	467	4-Aminobutyrate aminotransferase from *Segniliparus rotundus*	78% over 443 aa	Aminotransferase	*c6p*
akr	340	Cvm1 from *Streptomyces clavuligerus*	73% over 325 aa	Aldo-keto reductase	*cvm1*
soh	416	OrfA from *Streptomyces clavuligerus*	78% over 405 aa	Hydroxymethyltransferase	*orfA*
ceaS3	571	Carboxyethylarginine synthase 2 from *Streptomyces clavuligerus* Thiamine pyrophosphate protein domain protein TPP-binding from *Streptomyces flavogriseus*	83% over 560 aa 92% over 571 aa	Carboxyethylarginine synthase Carboxyethylarginine synthase	*ceaS1 and ceaS2*
bls3	507	β-lactam synthetase 2 from *Streptomyces clavuligerus*	74% over 423 aa	Forming the β-lactam ring	*bls1 and bls2*
pah3	320	Proclavaminate amidinohydrolase 2 from *Streptomyces clavuligerus* Agmatinase from *Streptomyces flavogriseus*	75% over 300 aa 90% over 315aa	Amidinohydrolase	*pah1 and pah2*
cas3	324	Clavaminate synthase 1 from *Streptomyces clavuligerus*	88% over 323 aa	Oxygenase	*cas1 and cas2*
oat3	394	Ornithine acetyltransferase 2 from *Streptomyces clavuligerus*	87% over 394 aa	Acetyltranfersase	*oat1 and oat2*
ntr	214	Oxidoreductase from *Rhodococcus erythropolis*	80% over 213aa	Nitroreductase/oxidoreductase	
*orf*Tü1	405	Major facilitator superfamily MFS_1 *Streptomyces violaceusniger*	58% over 408aa	Clavam transporter	

lig	478	LigA protein from *Streptomyces hygroscopicus*	62% over 427aa	Ligase	
pah4	305	Proclavaminate amidinohydrolase 2 from *Streptomyces clavuligerus*	71% over 296 aa	Amidinohydrolase	*pah1 and pah2*
	309	Agmatinase from *Streptomyces flavogriseus*	72% over 304 aa	Amidinohydrolase	
ctr	456	Hypothetical protein from *Streptomyces sp.*	69% over 430 aa	Transcriptional regulator	
*orf*Tü2	174	NUDIX hydrolase-like protein from *Streptomyces platensis*	81% over 168 aa	NUDIX hydrolase	
adh	219	3' end of alcohol dehydrogenase from *Streptomyces clavuligerus*	80% over 218 aa	Dehydrogenase	

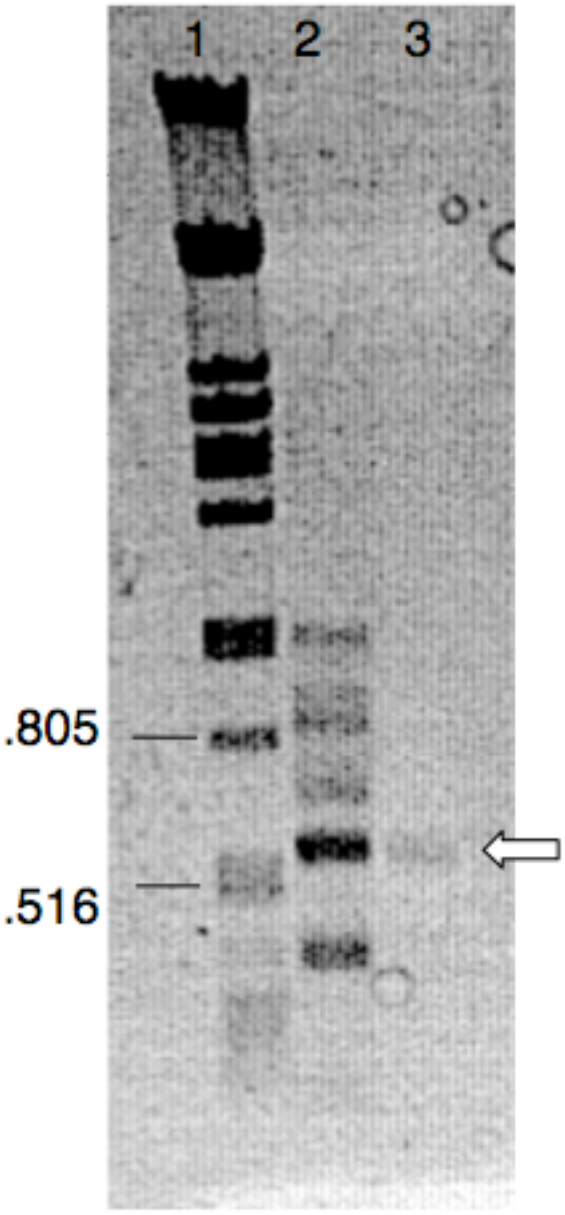


Figure III.1: Amplification of the *cas* gene from *S. antibioticus* by PCR. Left to right: lane 1: λPstI marker, lane 2: *S. antibioticus* Tü1718 gDNA, lane 3: *S. clavuligerus* gDNA. The desired fragments are shown with the white arrow.

```
cas1  ATGACCTC------AGTGGACTGCACCGCGTACGGCCCCGAGCTGCGCGCGCTCGCCGCCCGGCTGCCCCGGACCCCCCGGGCCGACCTGTACGCCTTCCTGGACGCCGCGC
cas2  ATGGCCTCTCCGATAGTTGACTGCACCCCGTACCGCGACGAGCTGCTCGCGCTCGCCTCCGAGCTTCCCGAGGTGCCGCGCGCGGACCTCCATGGCTTCCTCGACGAGGCGA
cas3  GTGACCGT------AGTCGACTGCTCGGAGTACTCCGCCGACCTGCTCGCGTTGGCCTCGCGGCTGCCGCGGATCCCCCGCCAGGACCTGTACGGCTTCCTCGACGCCGCGC
       ** **        *** ****** *   ****  *  *** **** **** * *** *   *** **   *   ** **    *****  * * ****** ****  ***

cas1  ACACAGCCGCCGCCTCGCTCCCCGGCGCCCTCGCCACCGCGCTGGACACCTTCAACGCCGAGGGCAGCGAGGACGGCCATCTGCTGCTGCGCGGCCTCCCGGTGGAGGCCGAC
cas2  AGACGCTGGCCGCCCGTCTCCCGGAGGGGCTGGCCGCCGCTCTCGACACCTTCAACGCCGTGGGCAGCGAGGACGGTTATCTGCTGCTGCGCGGGCTGCCCGTCGA---CGAC
cas3  ACGAAGCCGCCGGCGACCTGCCCGAGGGCCTGGGCACGGCGCTCGACCGGTTCAACGCCGACGGCAGCCACGACGGCTACCTCATGCTGCGGGGCCTGCCGGTGGAGGACGAC
      *       **** *   ** ** *  *  ** * * * ** ** ***   **********  ****** * *****  * **  ******* ** ** ** ** **   ****

cas1  GCCGACCTCCCCACCACCCCGAGCAGCACCCCGGCGCCCGAGGACCGCTCCCTGCTGACCATGGAGGCCATGCTCGGACTGGTGGGCCGCCGGCTCGGTCTGCACACGGGGTA
cas2  AGCGAGCTGCCCGAGACGCCGACCTCCACCCCGGCCCCGCTGGACCGCAAGCGGCTGGTGATGGAGGCCATGCTCGCGCTGGCCGGCCGCCGGCTCGGTCTGCACACGGGGTA
cas3  GACGACCTGCCGGCCACCCCCACCAGCACCCCGGCTCCGGTGGACCGCCCCCTGCAGAACATGGAGGCGATGCTGGCGGTCATCGGCCGCCGCCTGGGCCTGCACACCGGGTA
        *** ** **    ** ** * *  ********* **   *******   * ** *   ******** ***** *   *    ******** ** ** **************

cas1  CCGGGAGCTGCGCTCGGGCACGGTCTACCACGACGTGTACCCGTCGCCCGGCGCGCACCACCTGTCCTCGGAGACCTCCGAGACGCTGCTGGAGTTCCACACGGAGATGGCCT
cas2  CCAGGAGCTGCGCTCGGGCACGGTCTACCACGACGTGTACCCGTCGCCCGGCGCGCACTACCTGTCCTCGGAGACCTCCGAGACGCTGCTGGAGTTCCACACGGAGATGGCGT
cas3  CCGGGAGCTGCGGTCCGGCACCGTCTACCACGACGTGTACCCGTCGCCGGGCGCCCACCACCTGTCCTCCGAGACCTCCGAGACCCTGCTCGAATTCCACACGGAGATGGCCT
      ** ********* ** ***** ************************* ***** *** ********** ************** ***** ** ***************** *

cas1  ACCACCGGCTCCAGCCGAACTACGTCATGCTGGCCTGCTCCCGGGCCGACCACGAGCGCACGGCGGCCACACTCGTCGCCTCGGTCCGCAAGGCGCTGCCCCTGCTGGACGAG
cas2  ACCACATCCTCCAGCCGAACTACGTCATGCTGGCCTGCTCCCGCGCGGACCACGAGAACCGGGCGGAGACGCTGGTCGGCTCGGTCCGCAAGGCGCTGCCCCTGCTGGACGAG
cas3  ACCACGTCCTCCAGCCGAACTACGTGATGCTCGCATGCTCACGCGCGGACCACGAGCGCAAGGCGGCCACCCTGGTCGGCTCGATCCGCAAGGCGCTGCCCCTGATACCCGAG
      *****   ***************** ***** ** ***** ** ** *********  *  *****  ** ** **** **** ******************** *   ****

cas1  AGGACCCGGGCCCGGCTCCTCGACCGGAGGATGCCCTGCTGCGTGGATGTGGCCTTCCGCGGCGGGGTGGACGACCCGGGCGCCATCGCCCAGGTCAAACCGCTCTACGGGGA
cas2  AAGACCCGGGCCCGTCTCTTCGACCGCAAGGTGCCCTGCTGCGTGGACGTGGCCTTCCGCGGCGGGGTCGACGACCCGGGCGCGATCGCCAACGTCAAGCCGCTCTACGGGGA
cas3  GAAGTCCGCGCCCGCCTGTTCGACCGGCCCATGCCGTGCTGCGTGGACGTCGCGTTCCGCGGCGGCGTCGAGAACCCGGGCGCCATCGCCAACGTCAAGCCGCTGTACGGCGA
           *** ***** **  *******     **** *********** ** ** *********** ** **  ********** ****** * ***** ***** ***** **

cas1  CGCGGACGATCCCTTCCTCGGGTACGACCGCGAGCTGCTGGCGCCGGAGGACCCCGCGGACAAGGAGGCCGTCGCCGCCCTGTCCAAGGCGCTCGACGAGGTCACGGAGGCGG
cas2  CGCGAACGACCCGTTCCTCGGGTACGACCGCGAGCTGCTGGCGCCGGAGGACCCCGCGGACAAGGAGGCCGTCGCCCATCTGTCCCAGGCGCTCGACGATGTGACCGTCGGGG
cas3  TCCCCGGGATCCCTTCCTAGGCTACGACCGTGAGCTGCTCGCGCCGCGCGAACCCGACGACGTGGAGGCCGTCGCCGTGCTGTCGAAGGCGCTCGACGAGGTGAGCGAGGCGG
        *    ** ** ***** ** ******** ******** ******   ** ****  ***  *************   *****  ************* ** *  *  * **

cas1  TGTATCTGGAGCCCGGCGATCTGCTGATCGTCGACAACTTCCGCACCACGCACGCGCGGACGCCGTTCTCGCCCCGCTGGGACGGGAAGGACCGCTGGCTGCACCGCGTCTAC
cas2  TGAAGCTCGTCCCCGGTGACGTCCTCATCATCGACAACTTCCGCACCACGCACGCGCGGACGCCGTTCTCGCCCCGCTGGGACGGGAAGGACCGCTGGCTGCACCGCGTCTAC
cas3  TGCGGCTGACCCCCGGGGACCTTCTGGTCGTCGACAACTTCCGCACCACGCACGCCCGTACGCCGTTCTCCCCGCGCTGGGACGGCAAGGACCGCTGGCTGCACCGCGTCTAC
      **   **    ***** **  * **  ** ************************* ** *********** ** *********** **************************
cas1  ATCCGCACCGACCGCAATGGACAGCTCTCCGGCGGCGAGCGCGCGGGCGACGTCGTCGCCTTCACACCGCGCGGCTGA
cas2  ATCCGCACCGACCGCAATGGACAGCTCTCCGGCGGCGAGCGCGCGGGCGACACCATCTCGTTCTCGCCGCGCCGCTGA
cas3  ATCCGCACGGACCGCAACGACCAGCTGTCCGGCGGTGAACGCGCCGGTGACGTCGTGGACTTCTCGCCACGCCGCTAG
      ******** ******** *  ***** ******** ** ***** ** ***  * *    *** * ** *** ***
```

Figure III.2: Alignment of the three genes *cas1, cas2,* and *cas3.* The stars represent identical nucleotides in all three genes.

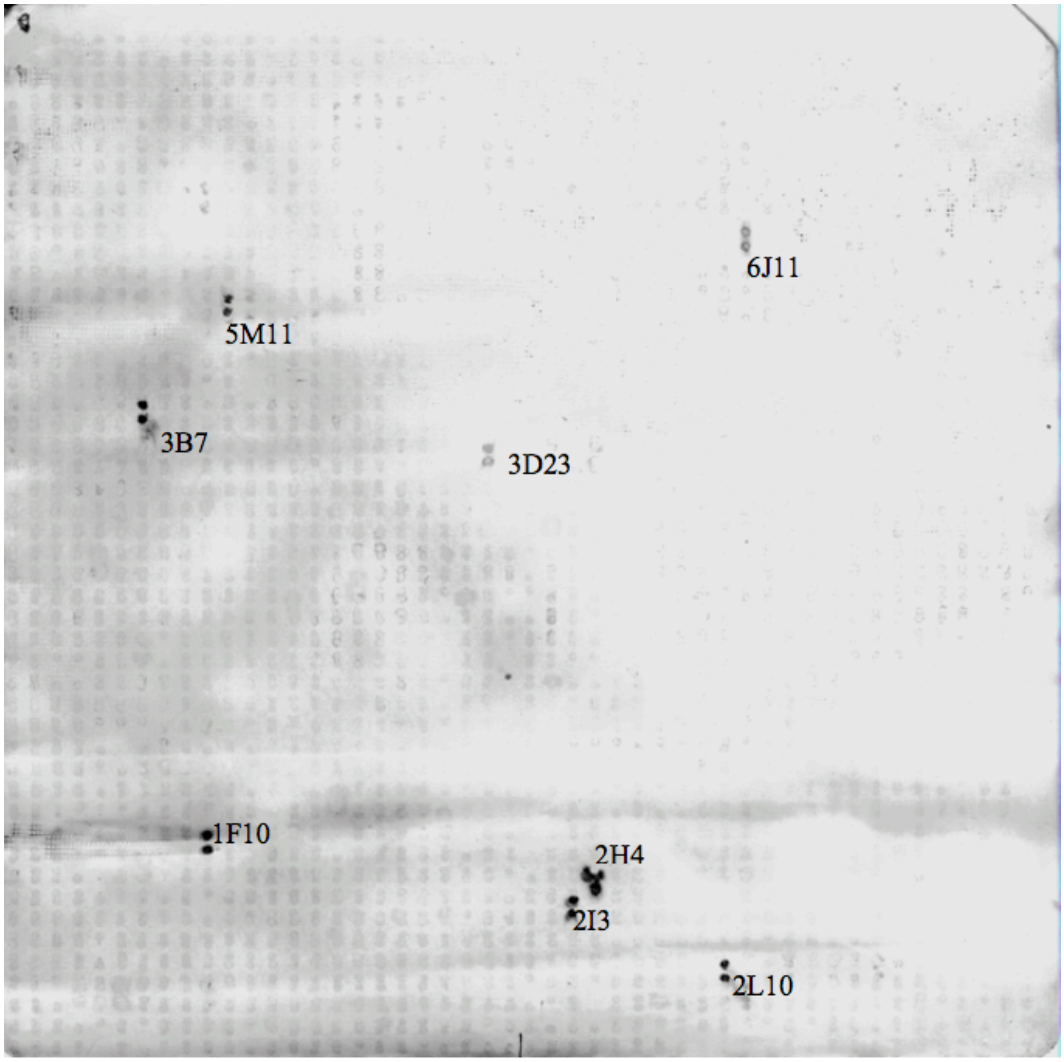


Figure III.3: Colony hybridization analysis of the cosmid library of genomic DNA from *S. antibioticus* Tü1718 using a *cas3* specific probe. The probe was labeled using a non-radioactive DNA labeling kit. Each colony was gridded in duplicates (vertically) on the membrane by a robot.

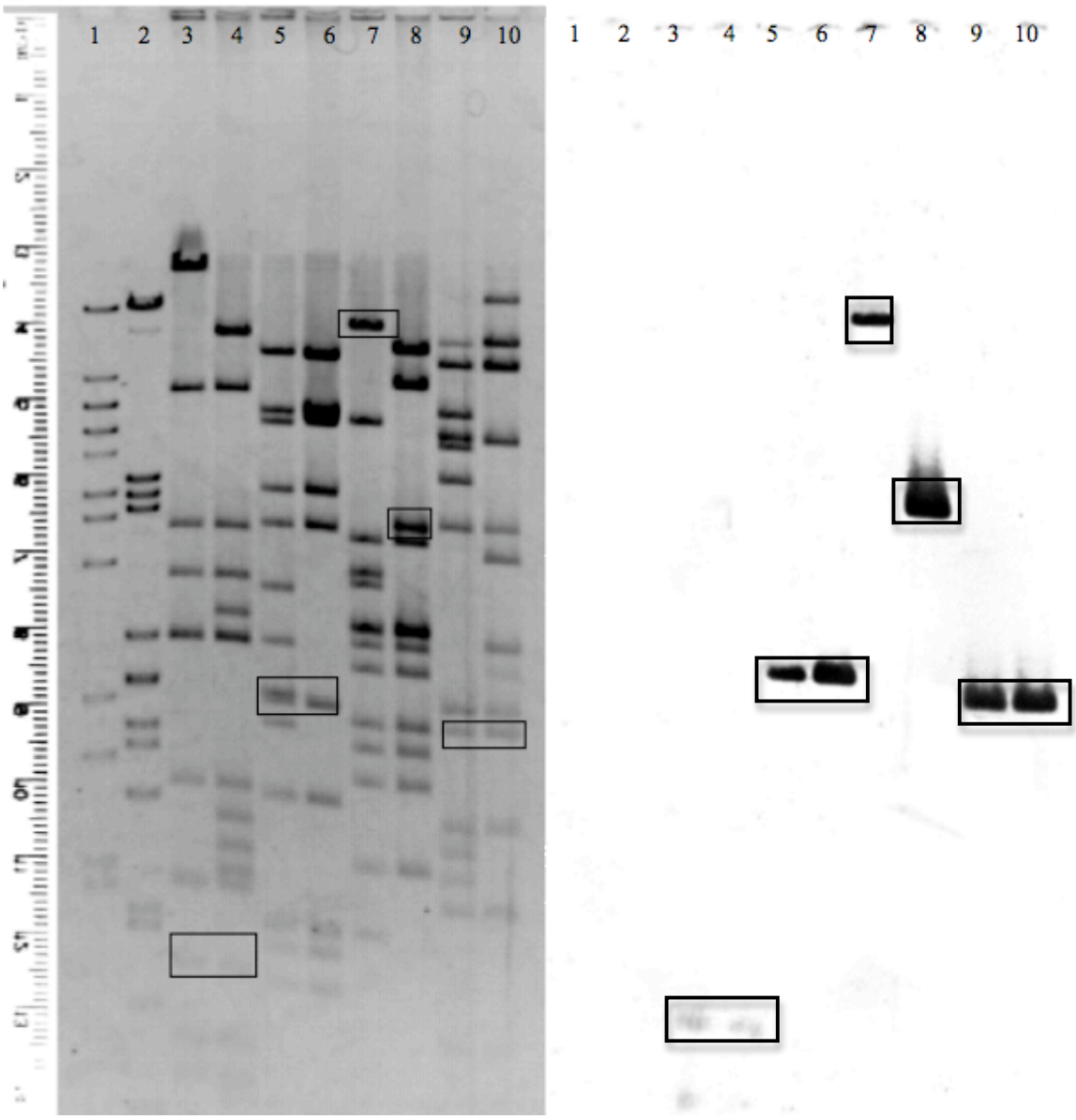


Figure III.4: Southern analysis of *cas3* containing cosmid digests. Left: Ethidium bromide stained agarose gel with cosmid 2L10 and 3D23 digests. Right: Hybridized nylon membrane with DNA transferred from the agarose gel. Lanes 1: λBstEII, 2: λPstI, 3&4: 2L10 and 3D23 digested with BamHI, 5&6: 2L10 and 3D23 digested with KpnI, 7&8: 2L10 and 3D23 digested with NcoI, 9&10: 2L10 and 3D23 digested with XcmI. The bands which showed hybridization in the Southern blot are shown in the boxes on the gel.

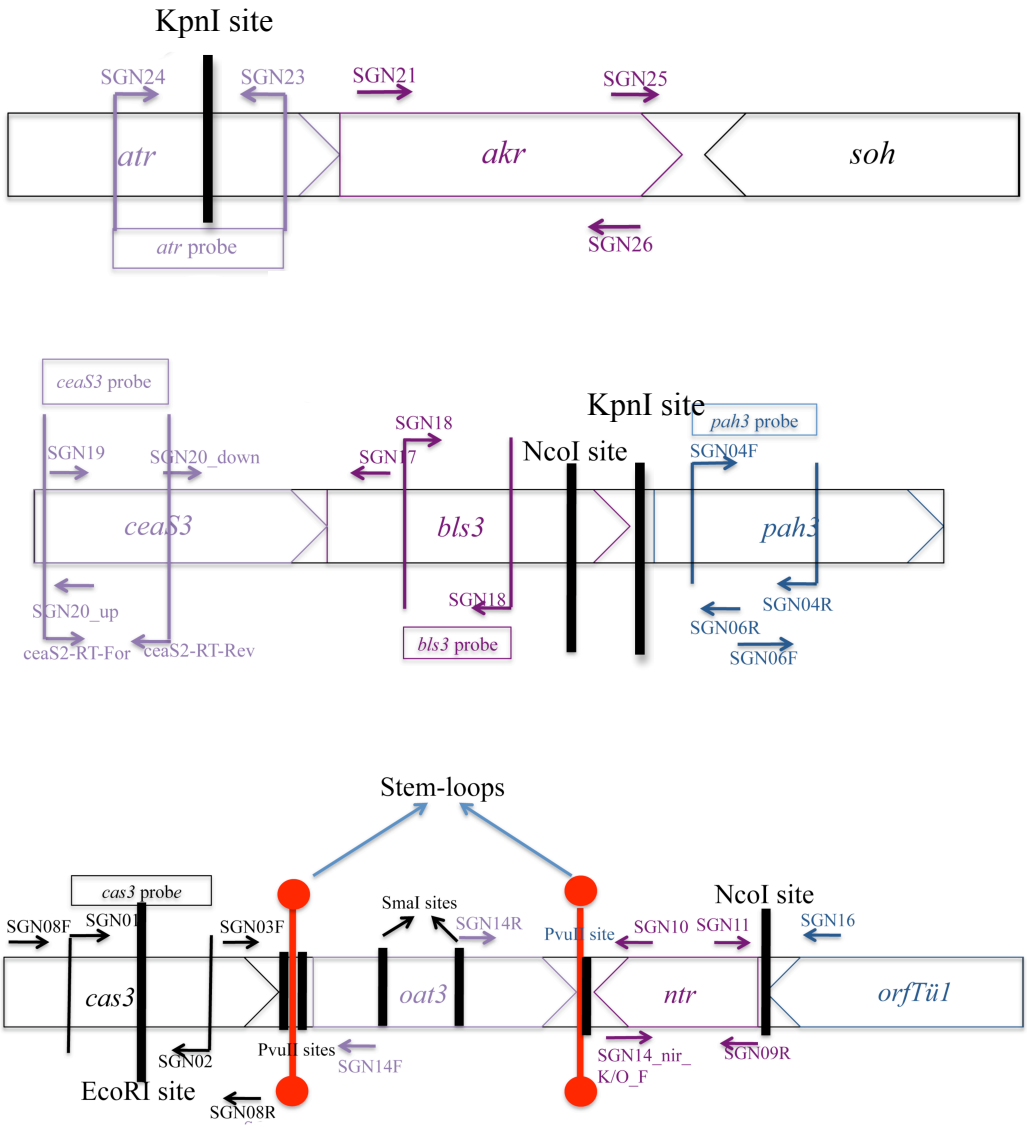


Figure III.5: The clavam gene cluster from *S. antibioticus* Tü1718. Important restriction sites are labeled. Secondary structures (red lollipop-like structures), probes used to identify different genes (shown in boxes) and primer binding sites (small arrows) are also indicated.

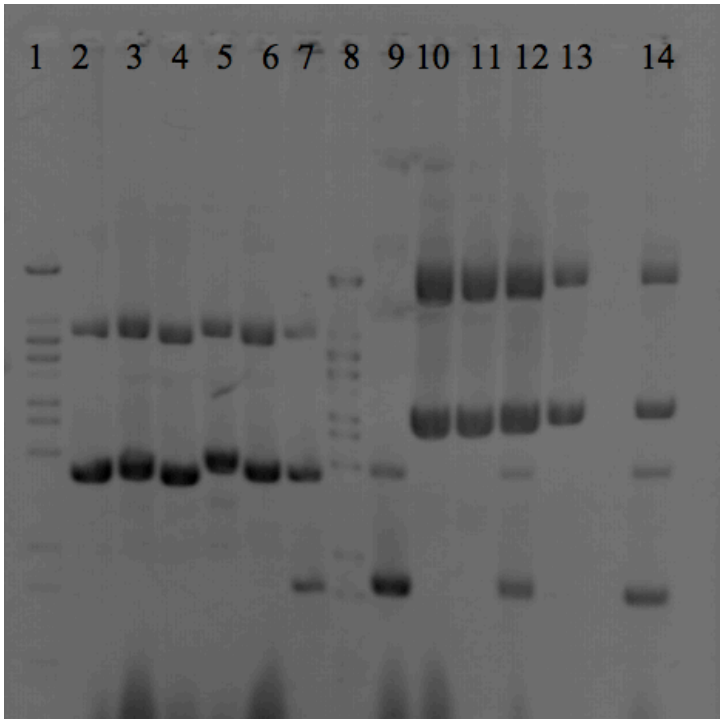


Figure III.6: Analysis of NcoI and KpnI sub-clones by agarose gel electrophoresis. The ethidium bromide stained agarose gel carries the KpnI (2-7) and the NcoI (10-14) sub-clones. The plasmids are uncut. Marker is λBstEII (lanes 1 and 8). Uncut pUC120 was also run (9). Some clones (lanes 7, 12 and 14) seemed to be mixed and contained plain pUC120 as well as a larger plasmid.

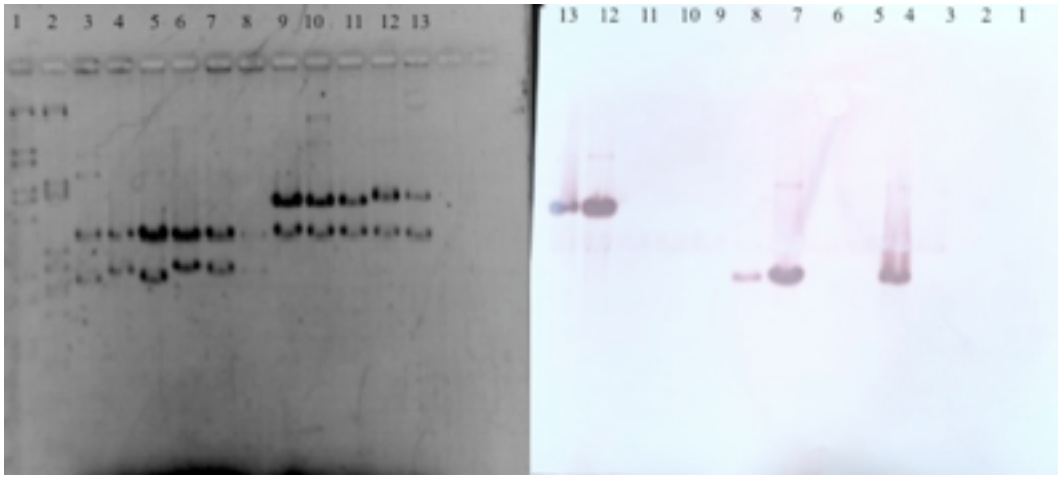


Figure III.7: Southern analysis of potential *cas3* sub-clones. Ethidium bromide stained agarose gel (left) and the Southern blot of an agarose gel containing the NcoI and KpnI clones (right). The membrane was hybridized with the *cas3* specific probe in order to identify which clone contained the desired insert. Lanes 12 and 13 represent the NcoI clones and lanes 4, 7 and 8 represent the KpnI clones containing the *cas3* gene.

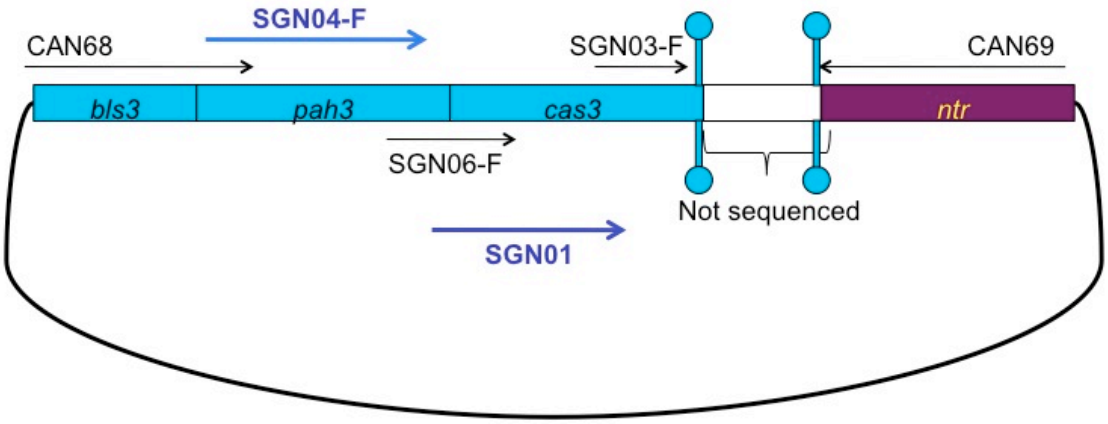


Figure III.8: Diagrammatic representation of the genes of the 4.5 Kbp NcoI sub-clone. Arrows indicate the location and approximate amount of sequence information obtained from each primer. The "Lollipop" like structures indicate the presence of putative stem-loop structures that blocked sequence analysis.

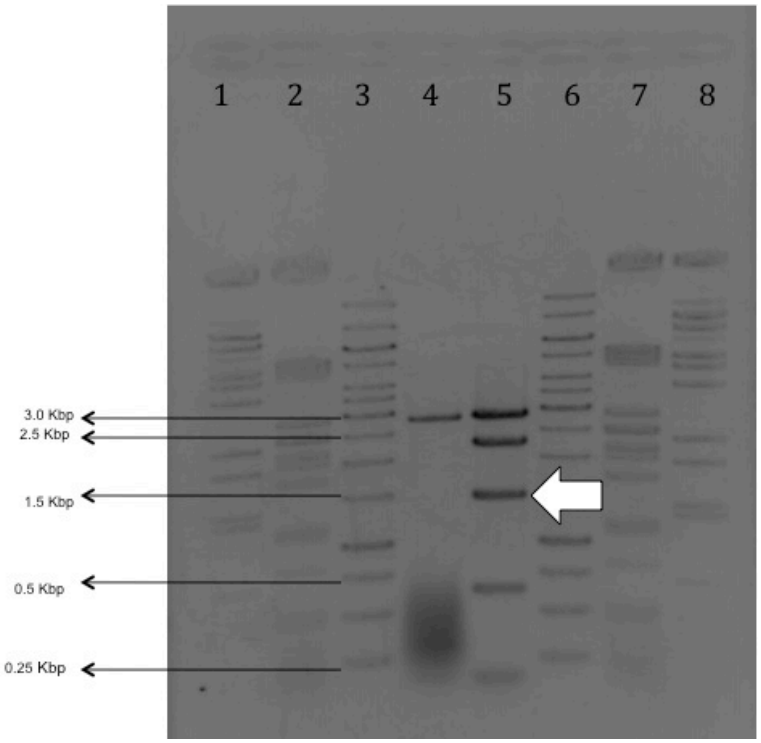


Figure III.9: Digestion of the NcoI sub-clone with NcoI and PvuII. PvuII-NcoI digested NcoI sub-clone and PvuII digested pUC120 were separated on an agarose gel and stained with ethidium bromide. Lanes 1 and 8: λBstEII, lanes 2 and 7: λPstI, lanes 3 and 6: the 100bp marker, lane 4: pUC120 cut with PvuII, lane 5: the NcoI sub-clone cut with NcoI and PvuII. The desired fragment is shown with a white arrow.

Figure III.10: Alignment of the DNA sequences of the stem-loop region obtained from PCR products amplified using different concentrations of deaza-dGTP. Lines 1-4 show the sequences of the PCR products obtained using 0.2, 0.4, 0.6, and 0.8 mM deaza-dGTP, respectively, with SGN03_FOR as sequencing primer. Lines 5-8 show the sequences of the same PCR products obtained using primer SGN14_FOR. The line "c" which is bold represents the sequence of the stem-loop as determined subsequently by pyrosequencing of the entire cosmid 6J11.

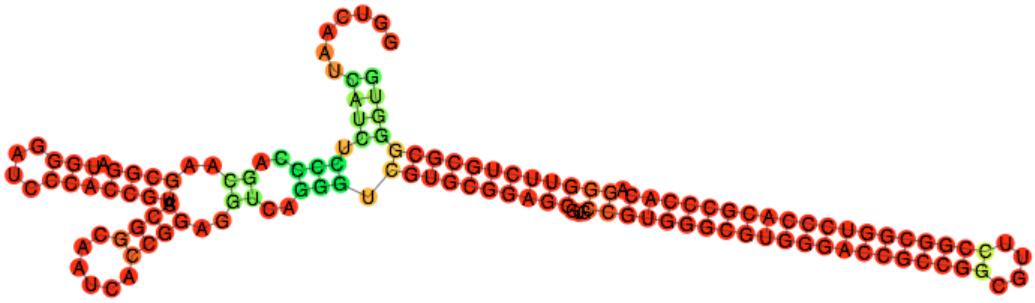


Figure III.11: The predicted structure of the stem-loop between *cas3* and *oat3*. Determined by http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi.

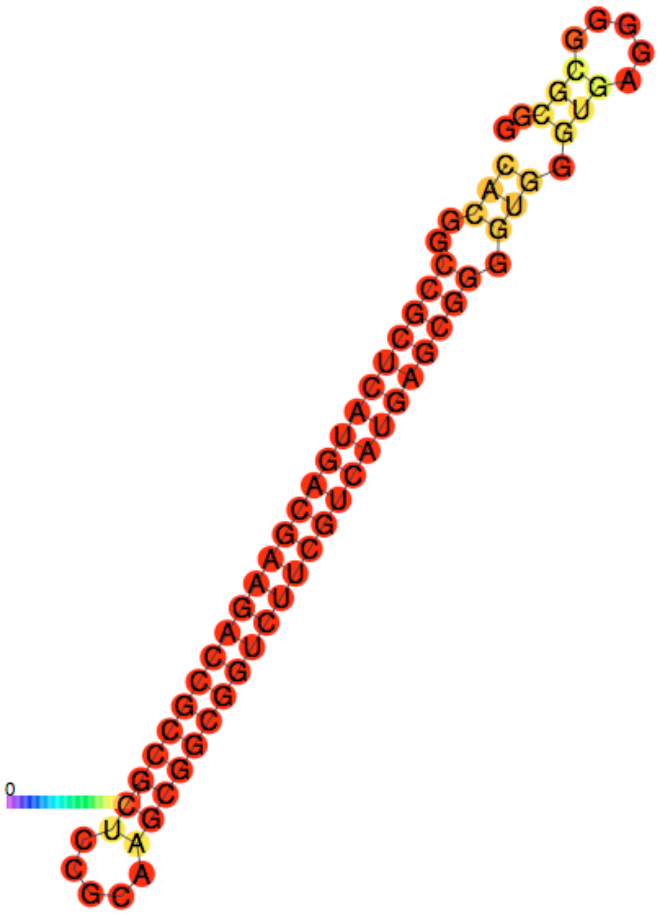


Figure III.12: The predicted structure of the stem-loop between *oat3* and *ntr*. Determined by http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi.

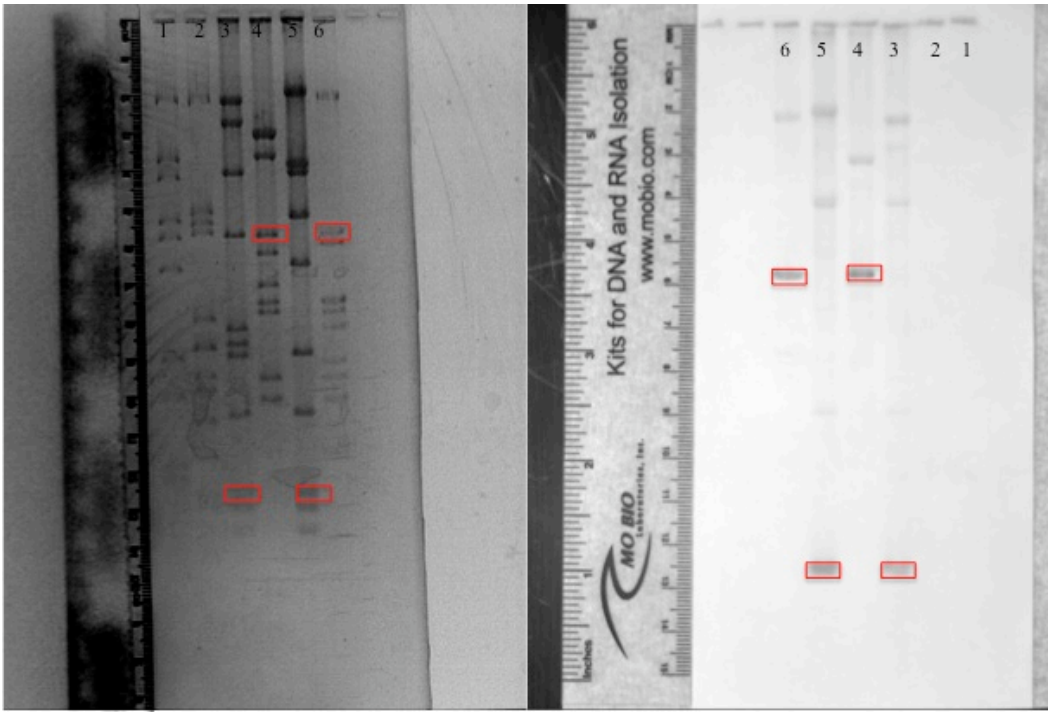


Figure III.13: Southern analysis to locate the *pah3* gene. The ethidium bromide stained gel with the KpnI and NcoI digests of 1F10 and 6J11 is shown on the left, and the Southern analysis of the transferred DNA of that gel, hybridized with the *pah3* probe is on the right. Lane 1: λBstEII, lane 2: λPstI, lane 3: Kpn1 digest of 1F10, lane 4: NcoI digest of 1F10, lane 5: KpnI digest of 6J11, lane 6: NcoI digest of 6J11.

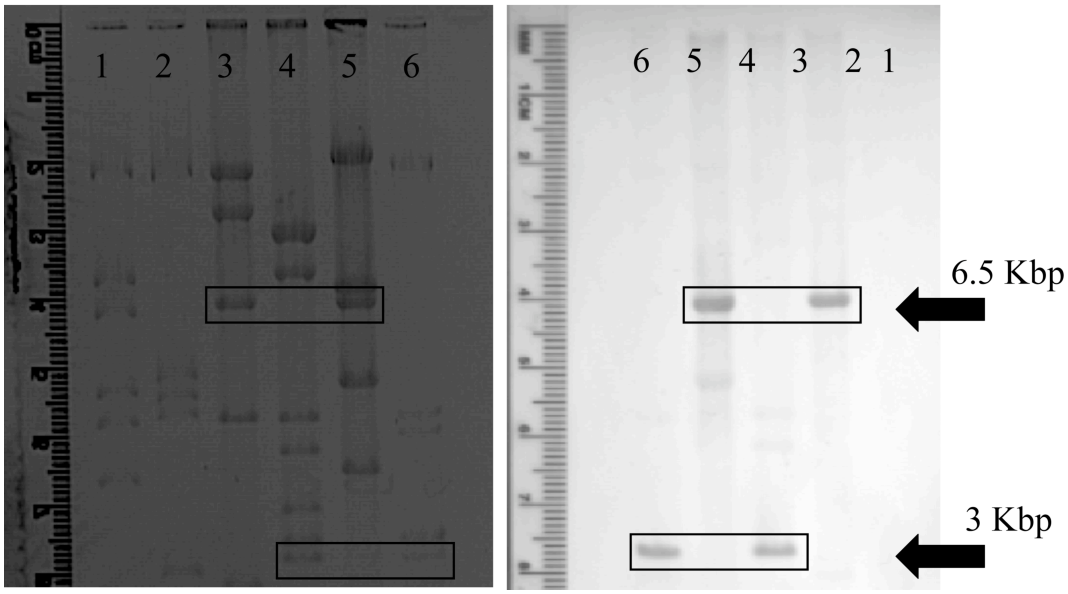


Figure III.14: Southern analysis to locate the *ceaS3* gene. The ethidium bromide stained gel with the KpnI and NcoI digests of cosmids 1F10 and 6J11 is shown on the left. The Southern analysis of the cosmids digested with NcoI and KpnI and hybridized with the *ceaS3* probe is shown on the right. Lane 1: λBstEII lane 2: λPstI, lane 3: Kpn1 digest of 1F10, lane 4: NcoI digest of 1F10, lane 5: KpnI digest of 6J11, lane 6: NcoI digest of 6J11.

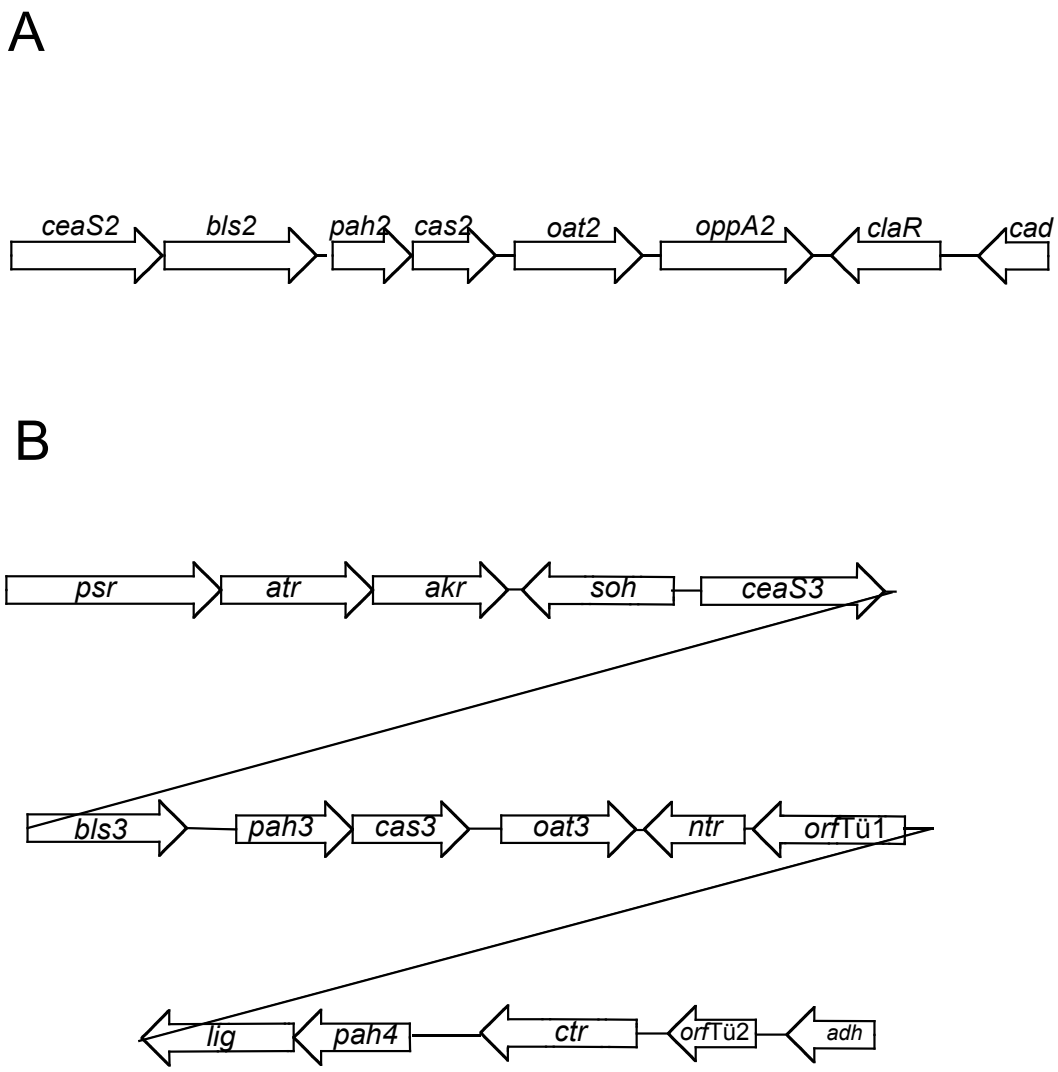


Figure III.15: A: Comparison of gene clusters for clavam biosynthesis in *S. clavuligerus* and *S. antibioticus* Tü1718. **A:** Part of the clavulanic acid gene cluster containing the early genes for clavam production in *S. clavuligerus.* B: The clavam gene cluster in *S. antibioticus* Tü1718. The early genes are shown in blue in both clusters.

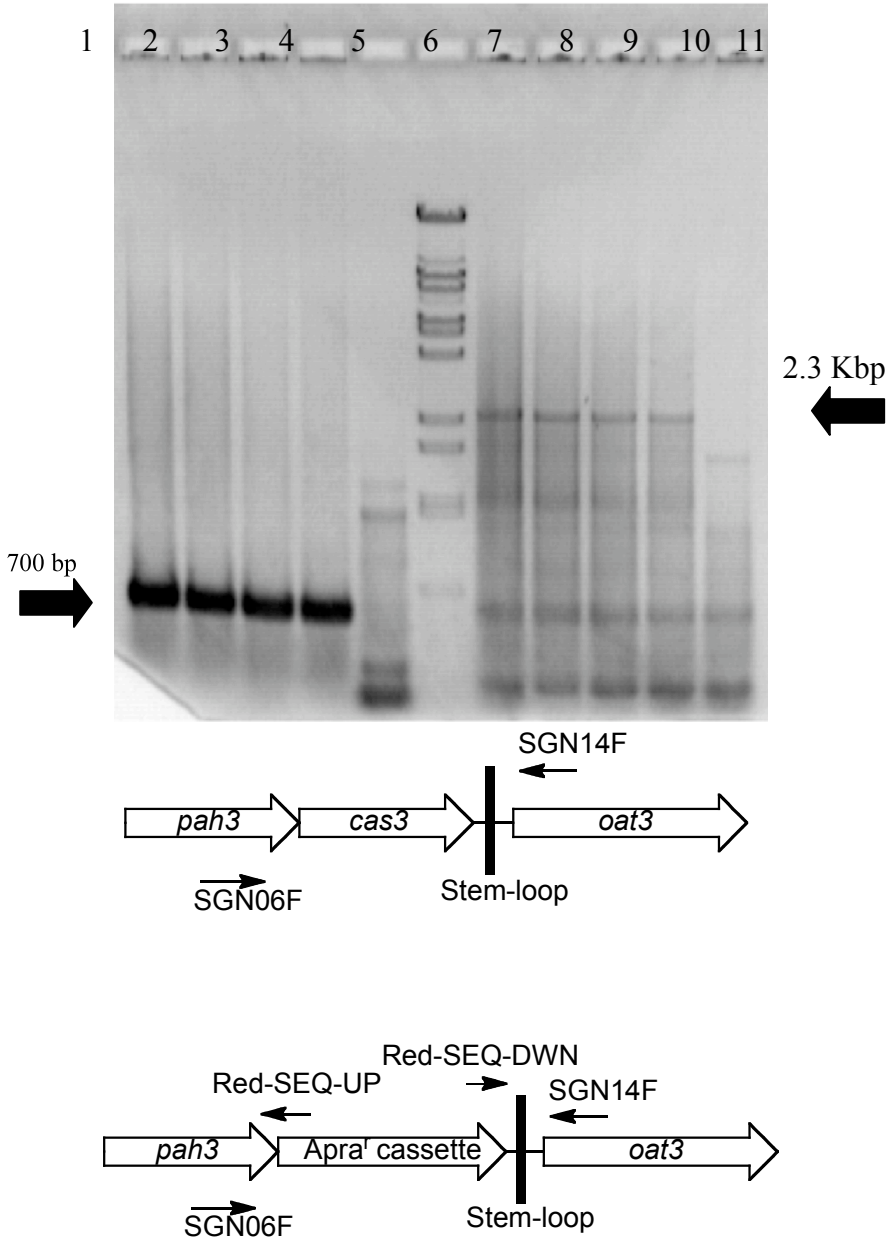


Figure III.16: PCR analysis of the *cas3* mutants. Lanes 1-4 show the PCR products obtained using gDNA from four putative mutants as template with primers Red-SEQ-UP (located within the apramycin cassette) and SGN06F (located just upstream of *cas3*). Lane 5 is the PCR product obtained using the same primer pair with gDNA from wild type *S. antibioticus* Tü1718. Lane 6 is the λBstEII marker. Lanes 7-10 show the PCR products obtained using gDNA from four putative mutants as template with primers SGN06F (located just upstream of *cas3*) and SGN14F (located just downstream of *cas3*). Lane 11 is the the PCR product obtained using the same primer pair with gDNA from wild type *S. antibioticus* Tü1718. This experiment was done by Dr. S.E. Jensen.

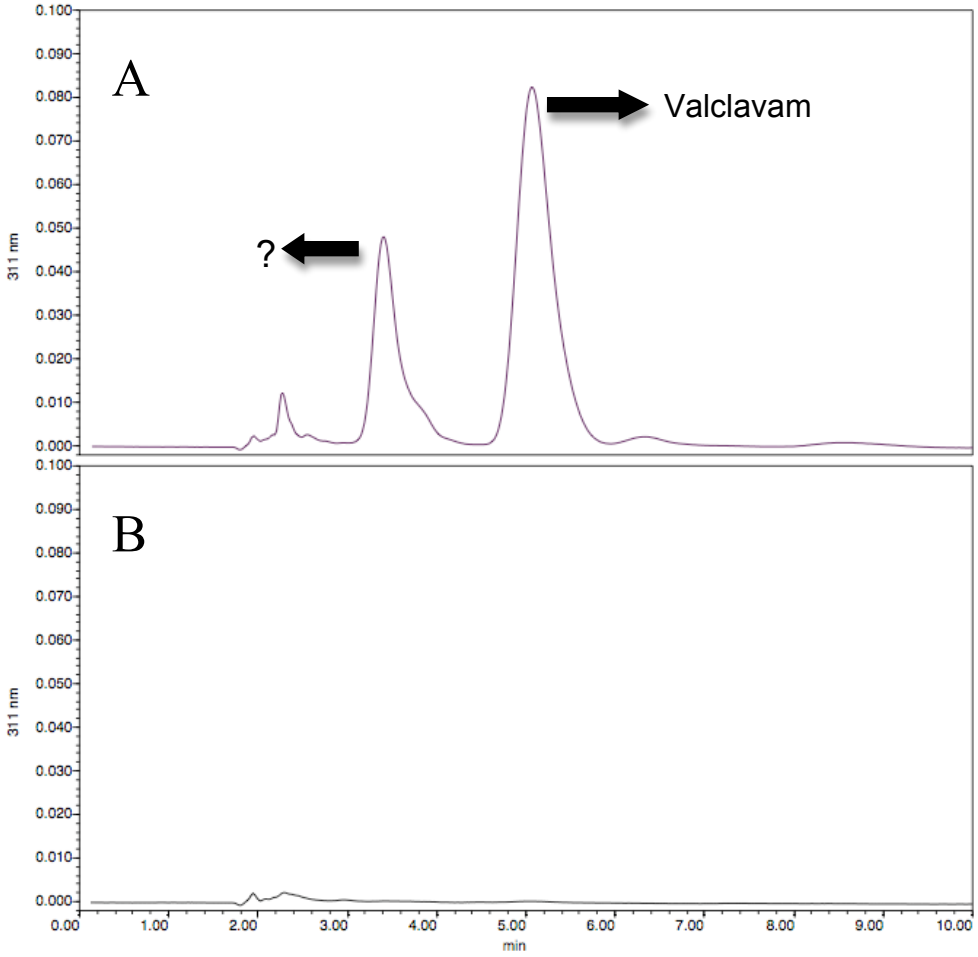


Figure III.17: HPLC analysis of the fermentation broths of wild type *S. antibioticus* Tü1718 (A) and the Δ*cas3* mutant (B). Cultures were grown in mannitol soy production medium for 48h at 28˚C. Culture supernatants were derivatized with imidazole to detect compounds with the β-lactam structure.

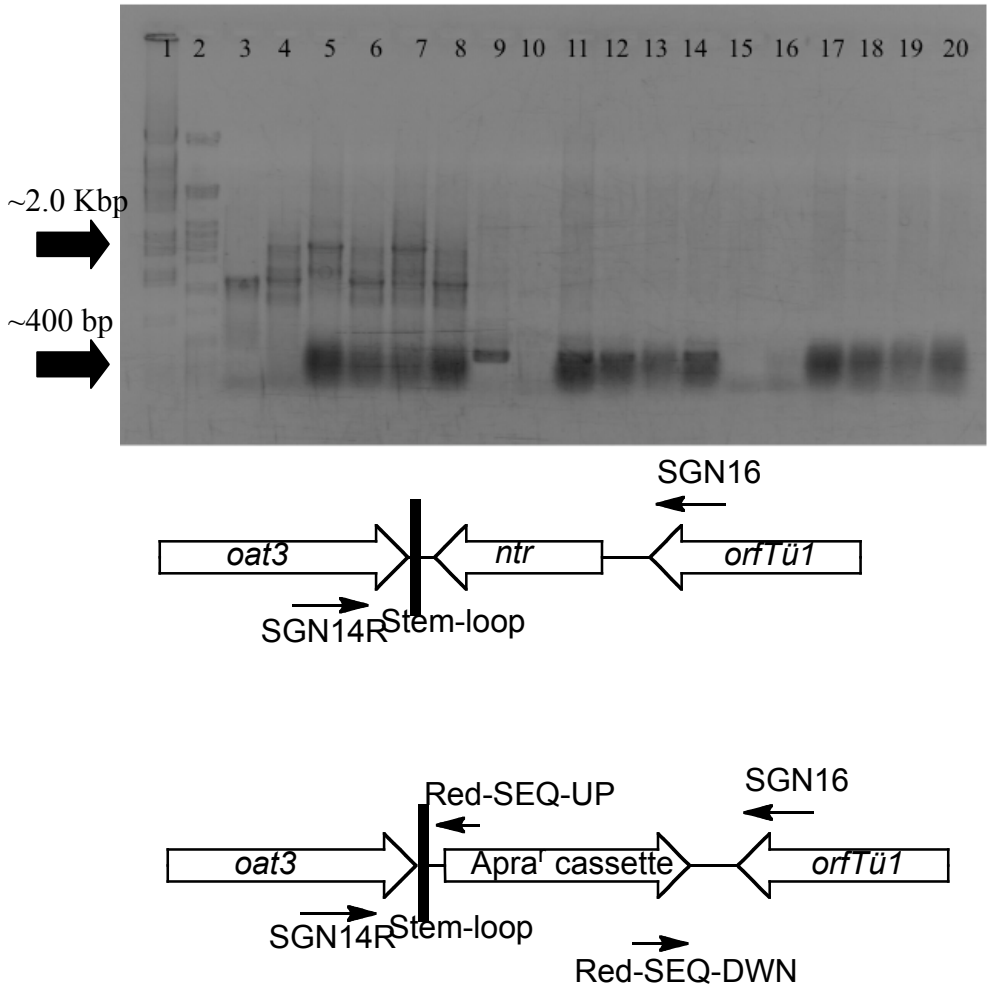


Figure III.18: PCR analysis of the *ntr* mutants. Lanes 1 and 2 show λBstEII and λPstI markers. Lane 3, PCR product obtained from the primers SGN14Rev and SGN16 using the *S. antibioticus* Tü1718 gDNA as a template. Lanes 4-8, PCR products of the same primers using the gDNA of the *ntr* mutants. Lane 10, wild type *S. antibioticus* Tü1718 gDNA amplified with primers Red-SEQ-DWN and SGN16. Lanes 9 and 11-14, PCR products of the same primers using the gDNA of the mutants. Lane 15, PCR product of the primers Red-SEQ-UP and SGN 14R using the gDNA of wild type *S. antibioticus* Tü1718. Lanes 16-20, PCR products of the mutants using the same primers.

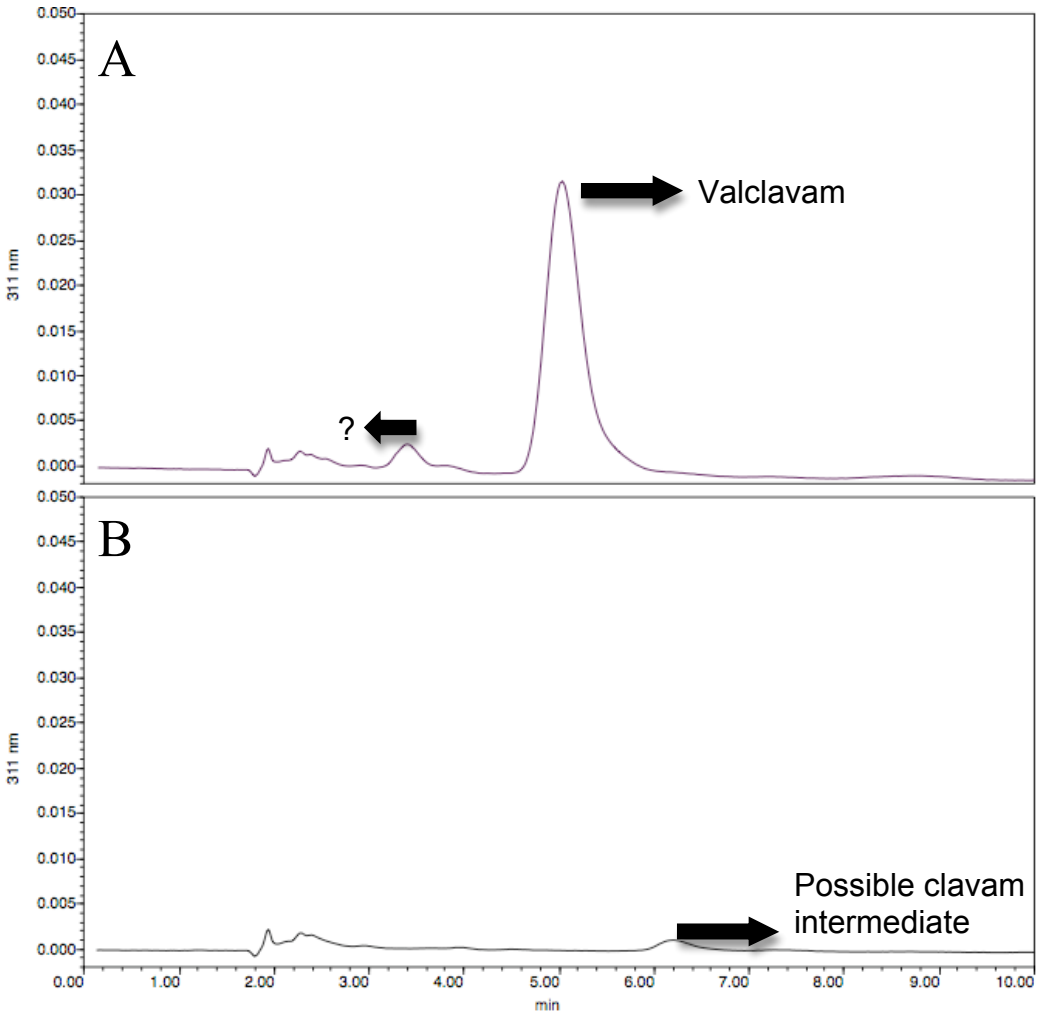


Figure III.19: HPLC analysis of the fermentation broths of wild type *S. antibioticus* Tü1718 (A) and the Δ*ntr* mutant (B). Cultures were grown in mannitol soy production medium for 48h at 28˚C.

IV. Discussion

IV.1. Discovery of the clavam gene cluster from *S. antibioticus* Tü1718

IV.1.1. The early genes

The biosynthetic pathway for clavam production in *S. clavuligerus* is comprised of early and late steps. The early steps end with the production of clavaminic acid. This compound is believed to be the branch point of 5*S* clavam and clavulanic acid pathways (Janc et al. 1993).

Using a PCR generated probe based on the regions of the *cas* genes that are highly conserved between *cas1* and *cas2* in *S. clavuligerus,* a *cas* orthologue was located in a cosmid library of genomic DNA fragments from *S. antibioticus* Tü1718. DNA sequence analysis of the regions flanking this *cas* gene, now called *cas3,* revealed the presence of other genes associated with the early steps of clavam biosynthesis.

As Figure I.7 shows, the clavam biosynthetic genes for the early steps in the clavulanic acid cluster are located in this order (from 5' to 3'): *ceaS2, bls2, pah2, cas2* and *oat2*. These are the genes that are required for the biosynthesis of clavaminic acid, in the shared pathway to clavulanic acid and the 5*S* clavams. These early genes of the clavam biosynthetic pathway in *S. antibioticus* Tü1718 (*ceaS3, bls3, pah3, cas3* and *oat3*) are in the same order and orientation as their orthologues in the clavulanic acid gene cluster in *S. clavuligerus*. This is shown in Figure III.15 A and B. These two microorganisms produce different clavams and yet the early genes are present in both of them with very high similarity. This observation confirms the role of these genes in the early steps of clavam

production in these two *Streptomyces* species and probably any clavam-producing microorganism and suggests the presence of these genes is likely to be observed in any other clavam producing species.

It has been shown that disrupting any of the early genes in *S. clavuligerus* reduces the production of clavulanic acid and *5S* clavams to a great extent (Bachmann et al. 1998; Jensen et al. 2000). The difference between *S. clavuligerus* and *S. antibioticus* Tü1718 is that the early genes in *S. clavuligerus* are present in two distinct and physically separated copies on the chromosome (mentioned in the Introduction). Because these genes are present in two copies, disruption of one of these copies would not eliminate clavam production completely.

It is suggested that the reason for the presence of the early genes in two copies is the participation of these genes in two pathways (clavulanic acid and 5*S* clavams) (Jensen and Paradkar 1999). Therefore given the fact that *S. antibioticus* Tü1718 produces only 5*S* clavams and no clavulanic acid, we expected to observe only one copy of these genes in this microorganism. This study showed that the early genes in *S. antibioticus* Tü1718 have only one copy with the exception of *pah,* which has a second copy of the gene in the cluster. This was surprising because the second copy was not in a separate cluster from the first one. In addition, this was the only early gene that had two copies.

The early steps for clavam biosynthesis and the function and characteristics of the different enzymes are explained in the Introduction. In order to confirm the

role of these genes in clavam production in *S. antibioticus* Tü1718 and show their necessity for the process, one of them, *cas3*, was selected to be subjected to PCR targeted mutagenesis. Also Janc et al., (1995) had previously shown Cas3 to be similar to Cas1 and Cas2 (more similar to Cas1). Mutation of *cas3* resulted in the complete loss of clavam production in *S. antibioticus* Tü1718. This confirms the role of *cas3* in the early steps of clavam production. In addition, it strongly suggests that *cas3* is the only copy of *cas* in *S. antibioticus* Tü1718, because if there were another copy there would still be some clavam production in the mutant.

IV.1.2. The late genes and the predicted biosynthetic pathway

The late steps for clavam production in *S. clavuligerus* are not yet fully understood. However some of the genes that function in the late steps have been identified. In *S. clavuligerus* the late genes for clavam biosynthesis are located in the clavam and paralogue gene clusters (Tahlan et al. 2007). The paralogue cluster contains the paralogues for the early genes, except *cas* (Tahlan et al. 2004b). The paralogue for *cas2* (*cas1*) is in the clavam cluster, flanked by some of the late genes (Mosher et al. 1999; Marsh et al. 1992). In addition to the paralogues of the early genes, the paralogue cluster also contains paralogues for *cvm6* and *cvm7,* two of the genes from the clavam cluster. Other genes responsible for clavam production in *S. clavuligerus* have been identified in the paralogue cluster. These were *orfA, orfB, orfC, orfD,* and they were shown to be required for the late steps in alanylclavam production in *S. clavuligerus.*

Analyzing the ORFs surrounding the early genes in *S. antibioticus* Tü1718 showed the presence of a number of ORFs that may encode enzymes for the late steps of the pathway (Figure III.15). Studying the ORFs found in *S. antibioticus* Tü1718 and comparing them with the corresponding genes from *S. clavuligerus* made it possible to predict a biosynthetic pathway for clavams in this microorganism. The proposed biosynthetic pathway for clavam biosynthesis is shown in Figure IV.1. As shown in the pathway, the first reaction after formation of clavaminic acid is believed to be a transamination. A potential candidate for this step is the enzyme Atr is produced by *atr*. This gene shows very high similarity to a 4-aminobutyrate aminotransferase from *Segniliparus rotundus*. A typical reaction catalyzed by an Atr is shown in Figure IV.2. According to Tahlan et al. (2007), in *S. clavuligerus* a transamination step is necessary after clavaminic acid. C6p, the product of *c6p* from the paralogue cluster, is proposed to catalyze this step in *S. clavuligerus*. The gene *atr* from *S. antibioticus* Tü1718 shows no similarity at the nucleotide or amino acid level to *c6p*. However, since they both have the same function, *atr* is thought to be a functional counterpart of *c6p*.

The next step of the pathway involves reduction of the double bond of the side chain and the candidate enzyme for this step is the aldo-keto reductase encoded by *akr*. The aldo-keto reductase (*akr*) is an orthologue of *cvm1* from the clavam gene cluster in *S. clavuligerus*. *akr* shows high similarity at the nucleotide and amino acid levels to *cvm1*. A study done by Mosher et al. showed that the disruption of *cvm1* abolishes the production of two of the 5*S* clavams in *S. clavuligerus*. However clavulanic acid and cephamycin C production were not

affected by this mutation (Mosher et al. 1999). This clearly indicates that Cvm1 is involved in the biosynthesis of 5*S* clavams only. It is not exactly clear which step of clavam biosynthesis in *S. clavuligerus* is catalyzed by Cvm1, however in our proposed biosynthetic pathway, it is suggested that Cvm1, like Akr, catalyzes the second step of the late steps, which is the reduction of the C-2—C-8 double bond.

The gene *adh* does not have any counterparts in *S. clavuligerus.* The sequence of this gene, which lies at the one end of the sequenced insert in cosmid 6J11, seems to be incomplete. The predicted biosynthetic pathway in *S. antibioticus* Tü1718 differs from that of *S. clavuligerus* in that it involves another reduction on the same carbon (C-9) after the action of Akr. This step would give rise to 2-hydroxyethylclavam, which is one of the end products of the biosynthetic pathway in *S. antibioticus* Tü1718, but not in *S. clavuligerus.*

An alternative step, which results in a branch in the pathway, is the reaction carried out by Soh. This enzyme is the product of *soh,* which shows very high similarity to *orfA* from *S. clavuligerus* (at both the nucleotide and amino acid levels). OrfA resembles a serine hydroxymethyltransferase and is proposed to add glycine to one of the clavam intermediates to produce 8-hydroxyalanylclavam (Zelyas et al. 2008). A typical reaction catalyzed by a serine hydroxymethyltransferase is shown in Figure IV.3. In *S. antibioticus* Tü1718, Soh is proposed to add a glycine to the product of Akr and give rise to a 9-hydroxyclavam intermediate, analogous to the 8-hydroxyalanylclavam intermediate seen in *S. clavuligerus*.

The final step required to convert the 9-OH clavam intermediate into valclavam is attachment of a valine residue via a peptide bond. There is no clear candidate for such a reaction. The ligase (Lig) is produced by *lig* and does not have any counterpart in *S. clavuligerus.* This enzyme is proposed to catalyze the addition of a valine molecule to the 9-hydroxyclavam intermediate and give rise to valclavam.

This enzyme also shows limited similarity to D-alanyl-D-alanine ligase. This enzyme belongs to the ATP-grasp fold superfamily. Some other proteins of this family are carbamoyl phosphate synthetase, biotin carboxylase, and glycinamide ribonucleotide synthetase (Fan et al. 1997; Galperin and Koonin 1997; Thoden et al. 1998; Thoden et al. 2000; Wang et al. 1998). It is believed that these proteins share an ATP-dependent carboxylate-amine or thiol ligation mechanism. As Figure IV.4 shows this reaction is a carboxylate-amine ligation, involving the activation of a carboxylate, in our case, valine, as an acyl-phosphate intermediate that reacts with a nucleophilic substrate, in our case, the 9-OH clavam intermediate (Galperin and Koonin 1997).

ORF17 produced by *orf17* in *S. clavuligerus* is one of the enzymes that is necessary for clavulanic acid production. This enzyme, *N*-glycyl-clavulanic acid synthetase, catalyzes the conversion of (3*S*, 5*S*)-clavaminic acid to *N*-glycyl-clavaminic acid Figure IV.5. This enzyme is also a member of the ATP-grasp fold superfamily and requires ATP, glycine, Mg^{2+}, K^{+} for its action (Arulanantham et al. 2006). It is believed that the reaction involving ORF17 proceeds through an enzyme-bound *O*-glycyl-phosphate intermediate, which is analogous to the D-

alanyl phosphate intermediate that is seen in the D-alanine-D-alanine ligase subfamily (Healy et al. 2000; Mullins et al. 1990). However, there is no overall similarity between Lig and ORF17 at either the nucleotide or amino acid level.

The first gene on the 5' end of the cluster is *psr.* The predicted product of this gene is similar to SanG from *Streptomyces ansochromogenes,* which is a pathway specific regulator in nikkomycin biosynthesis. Liu et al. (2005) showed that this gene was necessary for nikkomycin production and it is needed for the transcription of two genes, *sanO* and *sanN.* The N-terminus of the protein SanG shows similarity to the SARP family (Liu et al. 2005). According to Sheldon et al. (2002) the proteins from the SARP family contain an OmpR-like DNA binding domain, which binds to certain repeats in the DNA (Sheldon et al. 2002). The promoter of *sanG* contains an A-T rich region that shares highly conserved residues with some genes including *ccaR* from *S. clavuligerus* (Perez-Llarena et al. 1997). No studies were done on the gene *psr,* therefore the exact role of this gene in clavam biosynthesis is not known, but it may encode a pathway specific regulator which controls clavam biosynthesis.

The gene *ctr* is another gene from the clavam gene cluster that shows similarity to transcriptional regulators. This gene shows high similarity to *c7p* from *S. clavuligerus,* and according to Tahlan et al. 2007, disruption of *c7p* resulted in the loss of 5*S* clavam production in *S. clavuligerus,* while the level of clavulanic acid was as high as the wild type. On this basis, Ctr may represent another transcriptional regulator involved in the production of 5S clavams in *S. antibioticus* Tü1718.

There were two ORFs in the cluster (*orf*Tü1 and *orf*Tü2) that did not show similarity to any antibiotic production genes. They are considered to be within the cluster because they are flanked by genes with known or proposed roles in clavam biosynthesis. *orf*Tü1 was similar to a NUDIX hydrolase from *Streptomyces platensis*. The NUDIX hydrolases are a family of enzymes that are found in viruses, archaea, eukaryotes and prokaryotes. These enzymes are mainly pyrophosphohydrolyases and the general structure of their substrates is a nucleoside diphosphate linked to variable group, **X** (NDP-X). The reaction yields NMP and P-X (Bessman et al. 1996). Some examples include nucleotide sugars and alcohols, dNTPs, dinucleoside polyphosphates (Np_nN), capped RNAs and dinucleotide co-enzymes (Fisher et al. 2004; Hori et al. 2005; Ito et al. 2005; Xu et al. 2004). No role is obvious for this *orf* at present.

The other ORF, *orf*Tü2, resembles a major facilitator transporter from *Thermobispora bispora*. The major facilitator superfamily (MFS) is also called the uniporter-symporter-antiporter family (Baldwin 1993; Goswitz and Brooker 1995; Griffith et al. 1992; Henderson 1991; Marger and Saier 1993). These transporters are single-polypeptide secondary carriers and transfer small solutes in response to the chemiosmotic gradient. The MFS transporters are present in many organisms and have been classified into 17 groups (Pao et al. 1998). It is possible that the product of this gene would have a role in the resistance of *S. antibioticus* Tü1718 against the antibiotics it produces, or in transport of clavam products out of the cell.

IV.2. The surprising features of the cluster

IV.2.1. A second copy of *pah*

The genes mentioned in the previous sections of this chapter may have a function in the clavam production pathway. However, there are some genes in the cluster, which either do not have an obvious role in the biosynthesis of clavams in *S. antibioticus* Tü1718, or their presence in the cluster was surprising. One of these interesting features was a second copy of *pah* in the cluster. Normally in *Streptomyces* spp., genes for the biosynthesis of secondary metabolites are clustered together. *S. clavuligerus* is an exception in having the genes for clavam biosynthesis in three different clusters. These clusters are in distant locations on the chromosome of this microorganism and the paralogue gene cluster is located on a giant linear plasmid outside the chromosome (Tahlan et al. 2004; Medama et al. 2010). In *S. clavuligerus* all of the early genes have a paralogue. Our studies have shown that in *S. antibioticus* Tü1718 *pah* is the only gene that has a paralogue, and interestingly, it is not spatially distant from the other copy. According to Wu et al. (1995), Pah removes the guanidino group from guanidinoproclavaminate and produces proclavaminate (Figure IV.6). This is because Cas is not able to use guanidinoproclavaminate as a substrate until the guanidino group has been removed (Wu et al. 1995). Pah is therefore an important enzyme in the pathway. However, there seems to be no clear explanation why there should be two copies of this gene only.

IV.2.2. The stem-loops flanking *oat3*

One of the early genes in clavam production is *oat3*. This gene shows very high similarity at the nucleotide and amino acid levels to its counterparts in *S. clavuligerus*. Studies on *S. clavuligerus* have shown that *oat* has ornithine acetyltransferase activity (Kershaw et al. 2002) and transfers the acetyl group from *N*-acetylornithine to glutamate. This is one of the crucial steps in arginine biosynthesis (Udaka and Kinoshita 1958). It is suggested by Tahlan et al. (2004) that this activity is required to produce arginine, which is one of the precursors in clavam biosynthesis. The reaction is shown in Figure IV.7.

It was shown by Tahlan et al. (2004) that disruption of *oat1* in *S. clavuligerus* did not have a significant effect on clavam production. It was suggested that *oat2* and another gene with the same function (*argJ*) compensate for the loss of *oat1*. Double mutants of *oat1/oat2* produced lower levels of clavams in *S. clavuligerus*. The residual levels of clavam biosynthesis are believed to be due to the activity of ArgJ (Tahlan et al. 2004).

The exact role of Oat in clavam biosynthesis in *S. clavuligerus* is not very clear. It seems that even in the absence of both copies of this gene there is still clavam production taking place. Therefore *oat* cannot be considered one of the essential early genes like *bls, ceaS, cas,* and *pah.* In *S. clavuligerus* neither of the copies of *oat* is flanked by a significant secondary structure. In contrast, in *S. antibioticus* Tü1718, *oat3* is flanked by two strong stem-loops with free energies of -81 and -55.5 Kcal/mol. These structures may to be isolating the gene or have a regulatory function disconnecting the transcription of *oat3* from that of the rest of

the early genes. It might be possible that *oat3* is only transcribed when the medium is deprived of arginine and there is a need for this precursor.

IV.2.3. The nitroreductase gene

Another gene with an unclear function in the clavam cluster of *S. antibioticus* Tü1718 is nitroreductase (*ntr*). This gene encodes a putative protein very similar to a nitroreductase from *Thermomonospora curvata.* The whole genome of this microorganism has been sequenced recently. The family of nitroreductases comprises enzymes that are FMN- or FAD- and NAD(P)H-dependent and they are found in bacteria and some eukaryotes. Nitroreductases are not present in photosynthetic eukaryotes. The Ntr from *T. curvata* is believed to have an oxidoreductase activity. A schematic reaction catalyzed by nitroreductase is shown in Figure IV.8 (Yanto et al. 2010).

There is no obvious place for involvement of a nitroreductase in the clavam biosynthetic pathway in *S. antibioticus* Tü1718. However the gene is located in the cluster upstream of the second copy of *pah.* Therefore it was suggested that this gene might have a role in clavam production. *ntr* was subjected to PCR-targeted mutagenesis and the gene was replaced with an apramycin resistance cassette. Disruption of *ntr* resulted in the loss of clavam production in the mutant, which was surprising because there seems to be no obvious role for this enzyme in the pathway. Because *ntr* is in reverse orientation to the early genes downstream of it, the loss of clavam production also should not be due to a polar effect.

IV.3. Future directions

Our studies provided insight about the clavam biosynthetic pathway in *S. antibioticus* Tü1718, the different genes involved in this process, and their similarities to the genes found in *S. clavuligerus.* More thorough studies are required to obtain a complete understanding of clavam biosynthesis in different *Streptomyces* species. *cas3* was the only early gene that was subjected to mutagenesis. From the early genes, *pah3* and its second copy, *pah4*, would also be good candidates for mutagenesis. Single mutants of each gene could be prepared and compared to wild type. This would show if one copy could compensate for the other one or not. If they could, a *pah3/pah4* double mutant would be expected to be unable to produce clavams. It would also be interesting to determine if only one copy of *pah* (either *pah3* or *pah4*) could restore clavam production to wild type levels.

Studies could be done on *psr* and the transcription of the clavam genes could be examined. The promoter region of the genes that their transcription is regulated by *psr* could be analyzed and compared to those in *sanO* and *sanN* to see whether they have the same motifs or not.

Complementation studies on *cas3* mutant are still in progress. If the *cas3* gene could successfully be inserted back into the *cas3* mutant of *S. antibioticus* Tü1718 it could be examined whether clavam production could be restored to wild type levels. If clavam production was restored, this would confirm that disruption of *cas3* was solely responsible for the loss of clavam production (as is most likely). However, if production was not restored, the *cas3* gene disruption

might also be having polar effects on expression of downstream genes in the cluster. The attempts for this experiment were unsuccessful. Possibly, the resistance marker that is present on the vector carrying the complementing gene could not give resistance to *S. antibioticus* Tü1718. It is also possible that the vector could not survive and replicate in *S. antibioticus* Tü1718. Regardless, the disruption of *cas3* clearly showed that the 5*S* clavam gene cluster of *S. antibioticus* Tü1718 has been identified. Complementation experiments for *ntr* are less relevant, because this gene is oriented in the opposite direction compared to the other genes flanking it. Therefore polarity should not be an issue, and the loss of clavam production can be specifically attributed to the mutation in *ntr*.

Another early gene that would be interesting to study further is *oat3,* especially due to the transcriptional isolation of this gene by the two strong stem-loops that are flanking it. Considering the fact that the secondary structures made PCR across the region almost impossible, mutation of this gene by a PCR-based strategy might be more difficult than other genes because of the secondary structures. The levels of clavam production could be compared to *oat1/oat2* double mutants in *S. clavuligerus*. Point mutations in the stem-loops could be made and the effects of these mutations could be examined on antibiotic production and transcription levels of this gene.

Another study that can be done on the early genes would be complementing the *S. clavuligerus* mutants of the early genes with the early genes from *S. antibioticus* Tü1718 and to examine whether these genes could restore clavam production to wild type levels in *S. clavuligerus* or not. More particularly,

introduction of *S. antibioticus* Tü1718 clavam genes might alter the spectrum of 5*S* clavam products produced by *S. clavuligerus.* This is one reason that the cosmid pOJ436 was chosen for library construction, to facilitate heterologous expression studies.

The nitroreductase is another mysterious gene with no obvious role in clavam production, but which was shown to be important for the biosynthesis of these compounds. Purification of nitroreductase enzyme and more structural studies on the protein would give more insight about the substrate and activity of this enzyme. The role of nitroreductase could then be more evident.

In the HPLC analysis of the fermentation broth from the *ntr* mutant, there was a very small peak that appeared approximately 9 minutes after injection. This peak showed an absorption spectrum that indicated it was a clavam product, and it was not present in wild type fermentation broth. Purification and structural analysis of this product could potentially help to identify the substrate for the nitroreductase enzyme.

HPLC analysis of the fermentation broth of the wild type *S. antibioticus* revealed two clavam peaks, and *S. antibioticus* Tü1718 is reported to produce two clavam products, 2-hydroxyethylclavam and valclavam. Mass spectrometry and bioassay analysis allowed the larger of the two peaks to be identified as valclavam. The mass of other peak however was not consistent with it being 2-hydroxyethylclavam. This peak might be one of the intermediates in clavam biosynthesis, and if it could be isolated and purified in larger amounts for

structural analysis, this would give further insights into the details of the biosynthetic pathway.

The two ORFs that were found in the cluster (*orf* Tü1 and *orf* Tü2) are believed to be general housekeeping genes more likely to be involved in primary metabolism than specifically necessary for clavam biosynthesis. These ORFs could be mutated to determine whether or not they have a role in clavam biosynthesis.

These experiments and similar studies on other clavam producing species would help to provide a complete picture of the clavam production pathway. This could lead to a better understanding of how it can be manipulated in *S. clavuligerus* towards gaining more clavulanic acid and less 5*S* clavams, or possibly even novel forms of clavulanic acid.

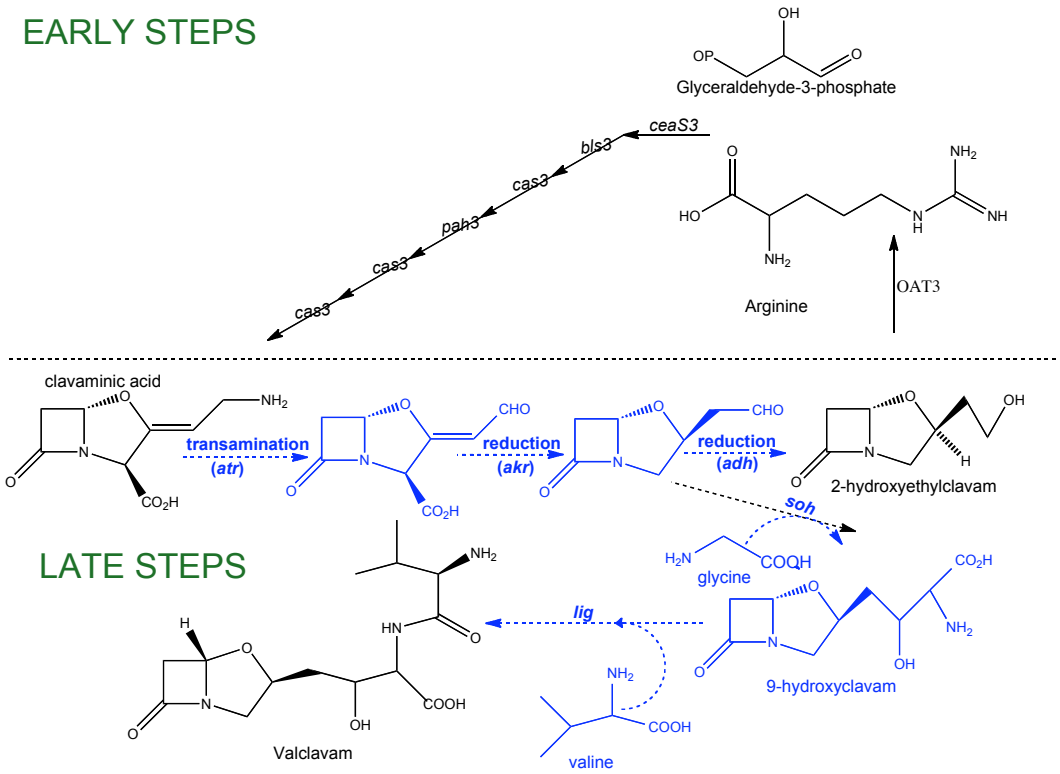


Figure IV.1: The proposed biosynthetic pathway for clavam production in *S. antibioticus* Tü1718. The early steps have not been shown in detail and only the order of the enzymes is shown. Each arrow represents a step, showing the gene producing the enzyme for that step on the top of the arrow. The proposed late steps are shown at the bottom (below the dashed line). The steps and products shown in blue are hypothetical. For details about the pathway refer to the text.

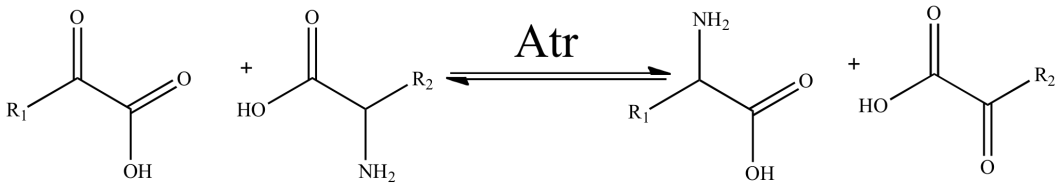


Figure IV.2: A typical reaction catalyzed by an aminotransferase (Atr). The amino group is removed from the amino acid to the oxo acid, making the oxo acid an amino acid and the amino acid an oxo acid.

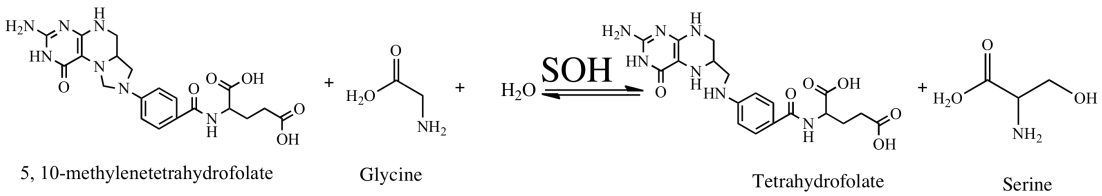


Figure IV.3: A typical reaction catalyzed by serine hydroxymethyltransferase (Soh). The methyl group from 5,10-methyltetrahydrofolate is transferred to glycine in the presence of water, giving tetrahydrofolate and serine.

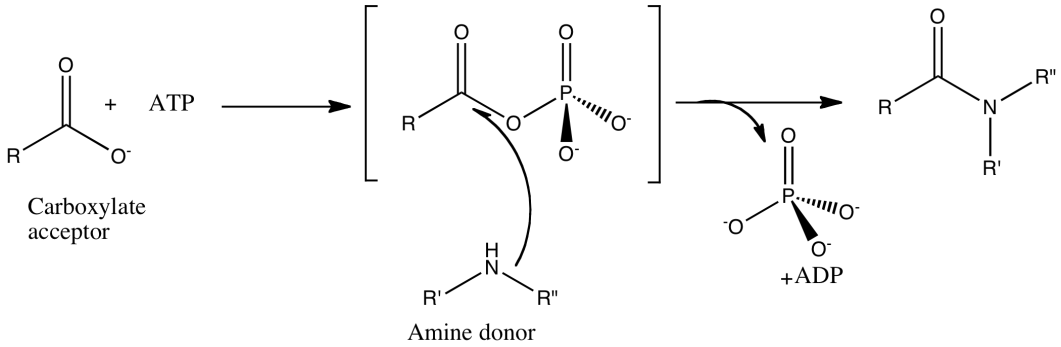
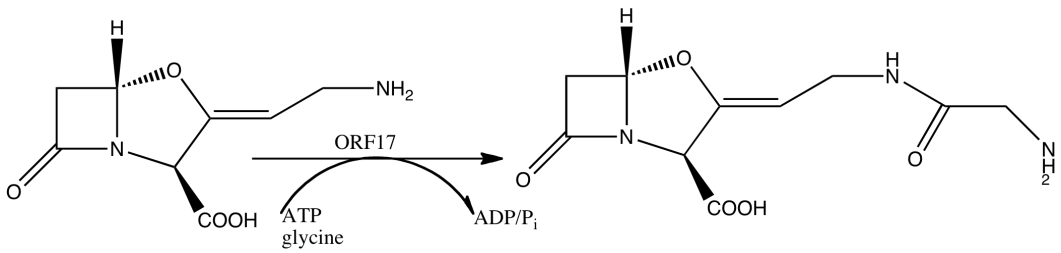


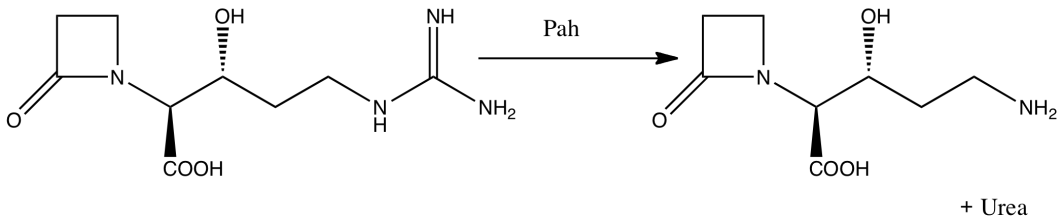
Figure IV.4: General mechanism of the ATP-grasp enzymes, a generic carboxylate-amine ligation reaction.



(3*S*, 5*S*)-clavaminic acid

N-glycyl-clavaminic acid

Figure IV.5: The action of ORF17 in the clavulanic acid pathway in *S. clavuligerus.* This enzyme catalyzes the conversion of (3*S,* 5*S*)-clavaminic acid to *N*-glycyl-clavaminic acid. The reaction involves the addition of glycine to the substrate and ATP hydrolysis.



Guanidinoproclavaminate Proclavaminate

Figure IV.6: The action of proclaminate amidinohydrolase (Pah). This enzyme is crucial in clavam biosynthesis and removes the guanidino group from guanidinoproclavaminate and gives proclavaminate.

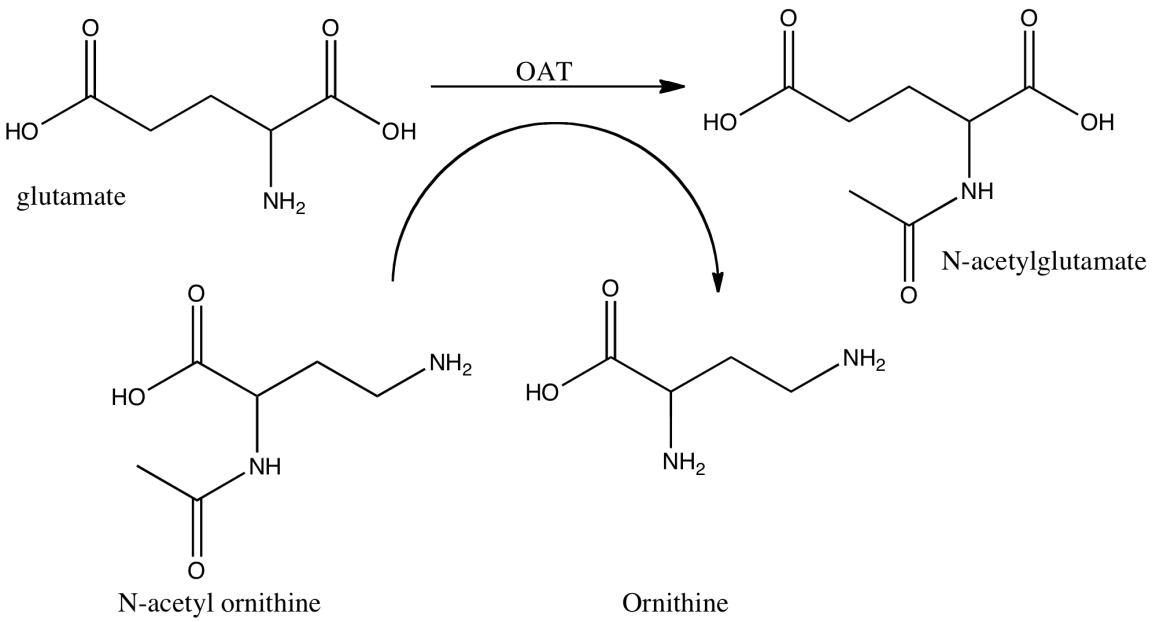
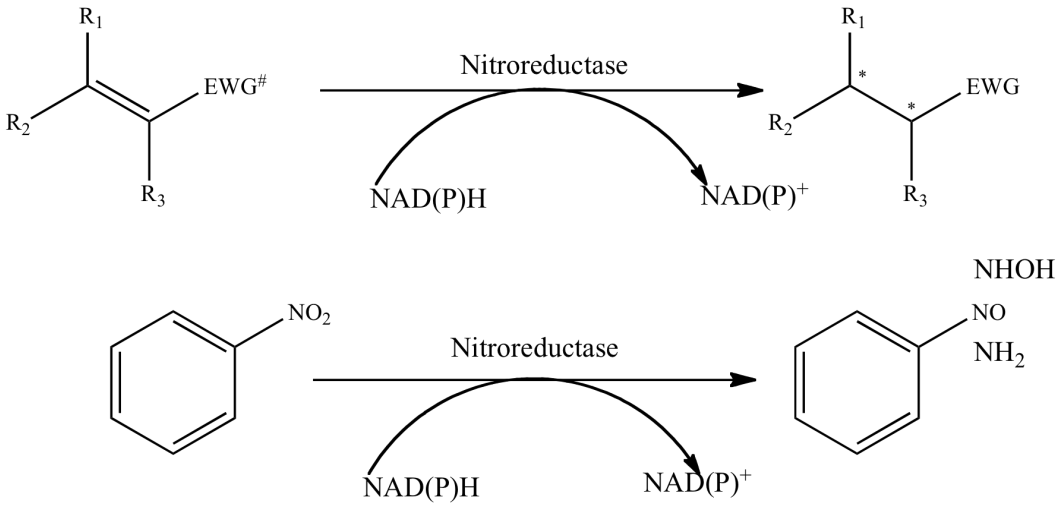


Figure IV.7: The activity of ornithine acetyltransferase (OAT). This enzyme is required for arginine biosynthesis. OAT catalyzes the transfer of an acetyl group from N-acetylornithine to glutamate (Udaka and Kinoshita 1958).



#= Electron Withdrawing Group

Figure IV.8: The reactions catalyzed by nitroreductase as shown by Yanto et al. (2010).

V. References

Aidoo, K.A., Wong, A., Alexander, D.C., Rittammer, R.A.R., and Jensen, S.E. 1994. Cloning, sequencing and disruption of a gene from *Streptomyces clavuligerus* involved in clavulanic acid biosynthesis. Gene **147**(1): 41-46.

Alexander, D.C., and Jensen, S.E. 1998. Investigation of the *Streptomyces* clavuligerus cephamycin C gene cluster and its regulation by the CcaR protein. J. Bacteriol. **180**(16): 4068-4079.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Res. **25**(17): 3389-3402.

Ando, N., Matsumori, N., Sakuda, S., Beppu, T., and Horinouchi, S. 1997. Involvement of AfsA in A-factor biosynthesis as a key enzyme. J. Antibiot. **50**(10): 847-852.

Antón, N., Mendes, M.V., Martín, J.F., and Aparicio, J.F. 2004. Identification of PimR as a Positive Regulator of Pimaricin Biosynthesis in *Streptomyces natalensis.* J. Bacteriol. **186**(9): 2567-2575.

Arulanantham, H., Kershaw, N.J., Hewitson, K.S., Hughes, C.E., Thirkettle, J.E., and Schofield, C.J. 2006. ORF17 from the clavulanic acid biosynthesis gene cluster catalyzes the ATP-dependent formation of N-glycyl-clavaminic acid. J. Biol. Chem. **281**(1): 279-287.

Bachmann, B.O., Li, R., and Townsend, C.A. 1998. β-lactam synthetase: A new biosynthetic enzyme. Proc. Natl. Acad. Sci. **95**(16): 9082-9086.

Baggaley, K.H., Brown, A.G., and Schofield, C.J. 1997. Chemistry and biosynthesis of clavulanic acid and other clavams. Nat. Prod. Rep. **14**(4): 309-333.

Baldwin, J.E., Lloyd, M.D., Wha-Son, B., Schofield, C.J., Elson, S.W., Baggaley, K.H., and Nicholson, N.H. 1993. A substrate analogue study on clavaminic acid synthase: Possible clues to the biosynthetic origin of proclavamic acid. J. Chem. Soc. D. (issue 6): 500-502.

Baldwin, S.A. 1993. Mammalian passive glucose transporters: Members of an ubiquitous family of active and passive transport proteins. Biochim. Biophys. Acta, Rev. Biomembr. **1154**(1): 17-49.

Bentley, S.D., Chater, K.F., Cerdeño-Tárraga, A.-., Challis, G.L., Thomson, N.R.,

James, K.D., Harris, D.E., Quail, M.A., Kieser, H., Harper, D., Bateman, A., Brown, S., Chandra, G., Chen, C.W., Collins, M., Cronin, A., Fraser, A., Goble, A., Hidalgo, J., Hornsby, T., Howarth, S., Huang, C.-., Kieser, T., Larke, L., Murphy, L., Oliver, K., O'Neil, S., Rabbinowitsch, E., Rajandream, M.-., Rutherford, K., Rutter, S., Seeger, K., Saunders, D., Sharp, S., Squares, R., Squares, S., Taylor, K., Warren, T., Wietzorrek, A., Woodward, J., Barrell, B.G., Parkhill, J., and Hopwood, D.A. 2002. Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). Nature **417**(6885): 141-147.

Bessman, M.J., Frick, D.N., and O'Handley, S.F. 1996. The MutT proteins or 'Nudix' hydrolases, a family of versatile, widely distributed, 'housecleaning' enzymes. J. Biol. Chem. **271**(41): 25059-25062.

Bierman, M., Logan, R., O'Brien, K., Seno, E.T., Nagaraja Rao, R., and Schoner, B.E. 1992. Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp. Gene **116**(1): 43-49.

Bignell, D.R.D., Lau, L.H., Colvin, K.R., and Leskiw, B.K. 2003. The putative anti-anti-sigma factor BldG is post-translationally modified by phosphorylation in *Streptomyces coelicolor*. FEMS Microbiol. Lett. **225**(1): 93-99.

Bignell, D.R.D., Tahlan, K., Colvin, K.R., Jensen, S.E., and Leskiw, B.K. 2005. Expression of ccaR, encoding the positive activator of cephamycin C and clavulanic acid production in *Streptomyces clavuligerus,* is dependent on bldG. Antimicrob. Agents Chemother. **49**(4): 1529-1541.

Bignell, D.R.D., Warawa, J.L., Strap, J.L., Chater, K.F., and Leskiw, B.K. 2000. Study of the bldG locus suggests that an anti-anti-sigma factor and an anti-sigma factor may be involved in *Streptomyces coelicolor* antibiotic production and sporulation. Microbiol. **146**(9): 2161-2173.

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

Box, S.J. 1977. Clavulanic acid and its salts by fermentation. Chem.Abstr. **87**.

Brown A.G., Butterworth D., Cole M., Hauscomb G., Hodd J.D., Reading C. 1976. Naturally-occurring β-lactamase inhibitors with antibacterial activity. J. Antibiot. (Tokyo) 29:668-669.

Bush, K. 1989. Classification of β-lactamases: Groups 1, 2a, 2b, and 2b'. Antimicrob. Agents Chemother. **33**(3): 264-270.

Bush, K. 1989. Classification of β-lactamases: Groups 2c, 2d, 2e, 3, and 4. Antimicrob. Agents Chemother. **33**(3): 271-276.

Butterworth D. 1984. Clavulanic acid: properties, biochemistry and fermentation.

In Biotechnology of industrial antibiotics. *Edited by* Vandam E. J. Marcel Dekker Inc., New York. pp. 226-234.

Chater K. F. 1989. Aspects of multicellular differentiation in *Streptomyces coelicolor*. *In* Genetics and molecular biology of industrial microorganisms. *Edited by* Herschberber C. L., Queener S. W., Hegeman G. American society for microbiology, Washington D. C. pp. 99-107.

De La Fuente, J.L., Martín, J.F., and Liras, P. 1996. New type of hexameric ornithine carbamoyltransferase with arginase activity in the cephamycin producers *Streptomyces clavuligerus* and *Nocardia lactamdurans*. Biochem. J. **320**(1): 173-179.

Demain, A.L., and Elander, R.P. 1999. The β-lactam antibiotics: Past, present, and future. Anton. Leeuw. Int. J. Gen. Mol. Microbiol. **75**(1-2): 5-19.

Egan, L.A., Busby, R.W., Iwata-Reuyl, D., and Townsend, C.A. 1997. Probable role of clavaminic acid as the terminal intermediate in the common pathway to clavulanic acid and the antipodal clavam metabolites. J. Am. Chem. Soc. **119**(10): 2348-2355.

Elkins, J.M., Clifton, I.J., Hernandez, H., Doan, L.X., Robinson, C.V., Schofield, C.J., and Hewitson, K.S. 2002. Oligomeric structure of proclavaminic acid amidino hydrolase: evolution of a hydrolytic enzyme in clavulanic acid biosynthesis. Biochem. J. **366**:423-434.

Elson, S.W., Gillett, J., Nicholson, N.H., and Tyler, J.W. 1988. N-acyl derivatives of clavaminic acid produced by a mutant of *Streptomyces clavuligerus*. J. Chem. Soc. D. (14): 979-980.

Elson, S.W., Baggaley, K.H., Fulston, M., Nicholson, N.H., Tyler, J.W., Edwards, J., Holms, H., Hamilton, I., and Mousdale, D.M. 1993. Two novel arginine derivatives from a mutant of *Streptomyces clavuligerus*. J. Chem. Soc. D. (15): 1211-1212.

Enos-Berlage, J.L., Langendorf, M.J., and Downs, D.M. 1998. Complex metabolic phenotypes caused by a mutation in yjgF, encoding a member of the highly conserved YER057c/YjgF family of proteins. J. Bacteriol. **180**(24): 6519-6528.

Fan, C., Park, I., Walsh, C.T., and Knox, J.R. 1997. D-alanine:D-alanine ligase: Phosphonate and phosphinate intermediates with wild type and the Y216F mutant. Biochemistry (N. Y.) **36**(9): 2531-2538.

Fisher, D.I., Cartwright, J.L., Harashima, H., Kamiya, H., and McLennan, A.G. 2004. Characterization of a Nudix hydrolase from *Deinococcus radiodurans* with a marked specificity for (deoxy)ribonucleoside 5′-diphosphates. BMC

Biochem. **5**(1): 1-24.

Fleming, A. 1929. On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of *B. influenzae*. Brit. J. Exp. Path. **10**: 226-236.

Folcher, M., Gaillard, H., Nguyen, L.T., Nguyen, K.T., Lacroix, P., Bamas-Jacques, N., Rinkel, M., and Thompson, C.J. 2001. Pleiotropic Functions of a *Streptomyces pristinaespiralis* Autoregulator Receptor in Development, Antibiotic Biosynthesis, and Expression of a Superoxide Dismutase. J. Biol. Chem. **276**(47): 44297-44306.

Foulstone, M., and Reading, C. 1982. Assay of amoxicillin and clavulanic acid, the components of augmentin, in biological fluids with high-performance liquid chromatography. Antimicrob. Agents Chemother. **22**(5): 753-762.

Gallagher, D.T., Gilliland, G.L., Xiao, G., Zondlo, J., Fisher, K.E., Chinchilla, D., and Eisenstein, E. 1998. Structure and control of pyridoxal phosphate dependent allosteric threonine deaminase. Structure **6**(4): 465-475.

Galperin, M.Y., and Koonin, E.V. 1997. A diverse superfamily of enzymes with ATP-dependent carboxylate-amine/thiol ligase activity. Protein Sci. **6**(12): 2639-2643.

Gomez-Escribano, J.P., Liras, P., Pisabarro, A., and Martín, J.F. 2006. An rplKΔ29-PALG-32 mutation leads to reduced expression of the regulatory genes *ccaR* and *claR* and very low transcription of the *ceaS2* gene for clavulanic acid biosynthesis in *Streptomyces clavuligerus*. Mol. Microbiol. **61**(3): 758-770.

Goswitz, V.C., and Brooker, R.J. 1995. Structural features of the uniporter/symporter/antiporter superfamily. Protein Sci. **4**(3): 534-537.

Griffith, J.K., Baker, M.E., Rouch, D.A., Page, M.G.P., Skurray, R.A., Paulsen, I.T., Chater, K.F., Baldwin, S.A., and Henderson, P.J.F. 1992. Membrane transport proteins: Implications of sequence comparisons. Curr. Opin. Cell Biol. **4**(4): 684-695.

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. **100**(4): 1541-1546.

Hara, O., and Beppu, T. 1982. Mutants blocked in streptomycin production in *Streptomyces griseus* - the role of A-factor. J. Antibiot. **35**(3): 349-358.

Hardisson, C., and Manzanal, M.B. 1976. Ultrastructural studies of sporulation in *Streptomyces.* J. Bacteriol. **127**(3): 1443-1454.

Hardisson, C., Manzanal, M.B., Salas, J.A., and Suarez, J.E. 1978. Fine structure, physiology and biochemistry of arthrospore germination in *Streptomyces antibioticus*. J. Gen. Microbiol. **105**(2): 203-214.

Hashimoto, K., Nihira, T., and Yamada, Y. 1992. Distribution of virginiae butanolides and IM-2 in the genus *Streptomyces*. J. Ferment. Bioeng. **73**(1): 61-65.

Hayakawa, T., Tanaka, T., and Sakaguchi, K. 1979. A linear plasmid-like DNA in *Streptomyces* sp. producing lankacidin group antibiotics. J. Gen. Appl. Microbiol. **25**(4): 255-260.

Healy, V.L., Mullins, L.S., Li, X., Hall, S.E., Raushel, F.M., and Walsh, C.T. 2000. D-Ala-D-X ligases: Evaluation of D-alanyl phosphate intermediate by MIX, PIX and rapid quench studies. Chem. Biol. **7**(7): 505-514.

Henderson, P.J.F. 1991. Sugar transport proteins. Curr. Opin. Struct. Biol. **1**(4): 590-601.

Higgens, C.E., and Kastner, R.E. 1971. *Streptomyces clavuligerus* sp. nov., a β-lactam antibiotic producer. Int. J. Syst. Bacteriol. **21**: 326-331.

Hodgson, J.E., Fosberry, A.P., Rawlinson, N.S., Ross, H.N.M., Neal, R.J., Amell, I.C., Earl, A.J., and Lawlor, E.J. 1995. Clavulanic acid biosyntbesis in *Streptomyces clavuligerus*: Gene cloning and characterization. Gene **166**(1): 49-55.

Hori, M., Fujikawa, K., Kasai, H., Harashima, H., and Kamiya, H. 2005. Dual hydrolysis of diphosphate and triphosphate derivatives of oxidized deoxyadenosine by Orf17 (NtpA), a MutT-type enzyme. DNA Repair **4**(1): 33-39.

Horinouchi, S., Suzuki, H., Nishiyama, M., and Beppu, T. 1989. Nucleotide sequence and transcriptional analysis of the Streptomyces griseus gene (*afsA*) responsible for A-factor biosynthesis. J. Bacteriol. **171**(2): 1206-1210.

Hutchings, M.I., Hoskisson, P.A., Chandra, G., and Buttner, M.J. 2004. Sensing and responding to diverse extracellular signals? Analysis of the sensor kinases and response regulators of *Streptomyces coelicolor* A3(2). Microbiology **150**(9): 2795-2806.

Hyun, S.K., Yong, J.L., Chang, K.L., Sun, U.C., Yeo, S., Yong, I.H., Tae, S.Y., Kinoshita, H., and Nihira, T. 2004. Cloning and characterization of a gene encoding the γ-butyrolactone autoregulator receptor from *Streptomyces clavuligerus*. Arch. Microbiol. **182**(1): 44-50.

Ikeda, H., Ishikawa, J., Hanamoto, A., Shinose, M., Kikuchi, H., Shiba, T., Sakaki, Y., Hattori, M., and Omura, S. 2003. Complete genome sequence and

comparative analysis of the industrial microorganism *Streptomyces avermitilis*. Nat. Biotechnol. **21**(5): 526-531.

Ishikawa, J., and Hotta, K. 1999. FramePlot: A new implementation of the Frame analysis for predicting protein-coding regions in bacterial DNA with a high G+C content. FEMS Microbiol. Lett. **174**(2): 251-253.

Ito, R., Hayakawa, H., Sekiguchi, M., and Ishibashi, T. 2005. Multiple enzyme activities of *Escherichia coli* MutT protein for sanitization of DNA and RNA precursor pools. Biochem. (N. Y.) **44**(17): 6670-6674.

Janc, J.W., Egan, L.A., and Townsend, C.A. 1995. Purification and characterization of clavaminate synthase from *Streptomyces antibioticus*. A multifunctional enzyme of clavam biosynthesis. J. Biol. Chem. **270**(10): 5399-5404.

Janc, J.W., Egan, L.A., and Townsend, C.A. 1993. Emerging evidence for a shared biosynthetic pathway among clavulanic acid and the structurally diverse clavam metabolites. Bioorg. Med. Chem. Lett. **3**(11): 2313-2316.

Jensen, R.A., and Gu, W. 1996. Evolutionary recruitment of biochemically specialized subdivisions of family I within the protein superfamily of aminotransferases. J. Bacteriol. **178**(8): 2161-2171.

Jensen, S.E., and Paradkar, A.S. 1999. Biosynthesis and molecular genetics of clavulanic acid. Anton. Leeuw. Int. J. Gen. Mol. Microbiol. **75**(1-2): 125-133.

Jensen, S.E., Wong, A., Griffin, A., and Barton, B. 2004. *Streptomyces clavuligerus* has a second copy of the proclavaminate amidinohydrolase gene. Antimicrob. Agents Chemother. **48**(2): 514-520.

Jensen, S.E., Elder, K.J., Aidoo, K.A., and Paradkar, A.S. 2000. Enzymes catalyzing the early steps of clavulanic acid biosynthesis are encoded by two sets of paralogous genes in *Streptomyces clavuligerus*. Antimicrob. Agents Chemother. **44**(3): 720-726.

Jensen, S.E., Paradkar, A.S., Mosher, R.H., Anders, C., Beatty, P.H., Brumlik, M.J., Griffin, A., and Barton, B. 2004. Five Additional Genes Are Involved in Clavulanic Acid Biosynthesis in *Streptomyces clavuligerus*. Antimicrob. Agents Chemother. **48**(1): 192-202.

Jin, W., Ryu, Y.G., Kang, S.G., Kim, S.K., Saito, N., Ochi, K., Lee, S.H., and Lee, K.J. 2004. Two relAlspoT homologous genes are involved in the morphological and physiological differentiation of *Streptomyces clavuligerus*. Microbiol. **150**(5): 1485-1493.

Jones, C., Thompson, A., and England, R. 1997. Guanosine 5'-diphosphate 3'-diphosphate (ppGpp) and clavulanic acid production in *Streptomyces*

clavuligerus. World J. Microbiol. Biotechnol. **13**(6): 633-636.

Kamitori, S., Okamoto, A., Hirotsu, K., Higuchi, T., Kuramitsu, S., Kagamiyama, H., Matsuura, Y., and Katsube, Y. 1990. Three-dimensional structures of aspartate aminotransferase from *Escherichia coli* and its mutant enzyme at 2.5 Å resolutio n. J. Biochem. **108**(2): 175-184.

Kelly, K.S., Ochi, K., and Jones, G.H. 1991. Pleiotropic effects of a *relC* mutation in *Streptomyces antibioticus*. J. Bacteriol. **173**(7): 2297-2300.

Kenig, M., and Reading, C. 1979. Holomycin and an antibiotic (MM 19290) related to tunicamycin, metabolites of *Streptomyces clavuligerus*. J. Antibiot. **32**(6): 549-554.

Kershaw, N.J., Caines, M.E.C., Sleeman, M.C., and Schofield, C.J. 2005. The enzymology of clavam and carbapenem biosynthesis. Chem. Comm. (34): 4251-4263.

Kershaw, N.J., McNaughton, H.J., Hewitson, K.S., Hernández, H., Griffin, J., Hughes, C., Greaves, P., Barton, B., Robinson, C.V., and Schofield, C.J. 2002. ORF6 from the clavulanic acid gene cluster of *Streptomyces clavuligerus* has ornithine acetyltransferase activity. Eur. J. Biochem. **269**(8): 2052-2059.

Khaleeli, N., Li, R., and Townsend, C.A. 1999. Origin of the β-lactam carbons in clavulanic acid from an unusual thiamine pyrophosphate-mediated reaction [3]. J. Am. Chem. Soc. **121**(39): 9223-9224.

Khokhlov, A.S., Tovarova, I.I., Borisova, L.N., Pliner, S.A., Shevchenko, L.N., Kornitskaia, E.I., Ivkina, N.S., and Rapoport, I.A. 1967. The A-factor, responsible for streptomycin biosynthesis by mutant strains of *Actinomyces streptomycini*. Dokl. Akad. Nauk SSSR **177**(1): 232-235.

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

Kieser, H.M., Kieser, T., and Hopwood, D.A. 1992. A combined genetic and physical map of the *Streptomyces coelicolor* A3(2) chromosome. J. Bacteriol. **174**:5496-5507.

Kim, J., Yoshikawa, H., and Shirahige, K. 2001. A member of the YER057 c/yjgf/Uk114 family links isoleucine biosynthesis and intact mitochondria maintenance in *Saccharomyces cerevisiae*. Genes to Cells **6**(6): 507-517.

King, H.D., Langharig, J., and Sanglier, J.J. 1986. Clavamycins, new clavam antibiotics from two variants of *Streptomyces hygroscopicus*. I. Taxonomy of the producing organisms, fermentation, and biological activities. J. Antibiot. **39**(4): 510-515.

Kinoshita, H., Tsuji, T., Ipposhi, H., Nihira, T., and Yamada, Y. 1999. Characterization of binding sequences for butyrolactone autoregulator receptors in *Streptomycetes*. J. Bacteriol. **181**(16): 5075-5080.

Kitano, K., Kintaka, K., and Katamoto, K. 1979. Clavulanic acid production by *Streptomyces katsurahamanus*. Chem. Abstr. **90:** 119758b.

Lawlor, E.J., Baylis, H.A., and Chater, K.F. 1987. Pleiotropic morphological and antibiotic deficiencies result from mutations in a gene encoding a tRNA-like product in *Streptomyces coelicolor* A3(2). Genes Dev. **1**(10): 1305-1310.

Leblond, P., Redenbach, M., and Cullum, J. 1993. Physical map of the *Streptomyces lividans* 66 genome and comparison with that of the related strain *Streptomyces coelicolor* A3(2). J. Bacteriol. **175**(11): 3422-3429.

Leblond, P., Francou, F.X., Simonet, J., and Decaris, B. 1990. Pulsed-field gel electrophoresis analysis of the genome of *Streptomyces ambofaciens* strains. FEMS Microbiol. Lett. **72**(1-2): 79-88.

Leskiw, B.K., Bibb, M.J., and Chater, K.F. 1991a. The use of a rare codon specifically during development? Mol. Microbiol. **5**(12): 2861-2867.

Leskiw, B.K., Mah, R., Lawlor, E.J., and Chater, K.F. 1993. Accumulation of bldA-specified tRNA is temporally regulated in *Streptomyces coelicolor* A3(2). J. Bacteriol. **175**(7): 1995-2005.

Leskiw, B.K., Lawlor, E.J., Fernandez-Abalos, J.M., and Chater, K.F. 1991b. TTA codons in some genes prevent their expression in a class of developmental, antibiotic-negative, *Streptomyces* mutants. Proc. Natl. Acad. Sci. **88**(6): 2461-2465.

Lezhava, A., Mizukami, T., Kajitani, T., Kameoka, D., Redenbach, M., Shinkawa, H., Nimi, O., and Kinashi, H. 1995. Physical map of the linear chromosome of *Streptomyces griseus*. J. Bacteriol. **177**(22): 6492-6498.

Li, R., Khaleeli, N., and Townsend, C.A. 2000. Expansion of the clavulanic acid gene cluster: Identification and in vivo functional analysis of three new genes required for biosynthesis of clavulanic acid by *Streptomyces clavuligerus*. J. Bacteriol. **182**(14): 4087-4095.

Lin, Y., Kieser, H.M., Hopwood, D.A., and Chen, C.W. 1993. The chromosomal DNA of *Streptomyces lividans* 66 is linear. Mol. Microbiol. **10**(5): 923-933.

Liras, P., and Rodriguez-Garcia, A. 2000. Clavulanic acid, a β-lactamase inhibitor: Biosynthesis and molecular genetics. Appl. Microbiol. Biotechnol. **54**(4): 467-475.

Liu, G., Tian, Y., Yang, H., and Tan, H. 2005. A pathway-specific transcriptional regulatory gene for nikkomycin biosynthesis in Streptomyces ansochromogenes that also influences colony development. Mol. Microbiol. **55**(6): 1855-1866.

Lorenzana, L.M., Pérez-Redondo, R., Santamarta, I., Martín, J.F., and Liras, P. 2004. Two oligopeptide-permease-encoding genes in the clavulanic acid cluster of *Streptomyces clavuligerus* are essential for production of the β-lactamase inhibitor. J. Bacteriol. **186**(11): 3431-3438.

Madigan M. T., Martinko J. M., Parker J.,. 2000. Brock Biology of Microorganisms. Prentice Hall, Upper Saddle River, NJ 07458.

Marck, C. 1988. 'DNA Strider': a 'C' program for the fast analysis of DNA and protein sequences on the Apple Macintosh family of computers. Nucleic Acids Res. **16**(5 PART A): 1829-1836.

Marger, M.D., and Saier Jr., M.H. 1993. A major superfamily of transmembrane facilitators that catalyse uniport, symport and antiport. Trends Biochem. Sci. **18**(1): 13-20.

Marsh, E.N., Chang, M.D., and Townsend, C.A. 1992. Two isozymes of clavaminate synthase central to clavulanic acid formation: Cloning and sequencing of both genes from *Streptomyces clavuligerus*. Biochem. (N. Y.) **31**(50): 12648-12657.

McLennan, A.G. 2006. The Nudix hydrolase superfamily. Cell. Mol. Life Sci. **63**(2): 123-143.

Medema, M.H., Trefzer, A., Kovalchuk, A., van den Berg, M., Müller, U., Heijne, W., Wu, L., Alam, M.T., Ronning, C.M., Nierman, W.C., Bovenberg, R.A.L., Breitling, R., and Takano, E. 2010. The sequence of a 1.8-Mb bacterial linear plasmid reveals a rich evolutionary reservoir of secondary metabolic pathways. Genome Biol. Evol. **2**: 212-224.

Mehta, P.K., Hale, T.I., and Christen, P. 1993. Aminotransferases: Demonstration of homology and division into evolutionary subgroups. Eur. J. Biochem. **214**(2): 549-561.

Mellado, E., Lorenzana, L.M., Rodríguez-Sáiz, M., Díez, B., Liras, P., and Barredo, J.L. 2002. The clavulanic acid biosynthetic cluster of *Streptomyces clavuligerus*: Genetic organization of the region upstream of the *car* gene. Microbiol. **148**(5): 1427-1438.

Mendez, C., Brana, A.F., Manzanal, M.B., and Hardisson, C. 1985. Role of substrate mycelium in colony development in *Streptomyces*. Can. J. Microbiol. **31**(5): 446-450.

Merrick, M.J. 1976. A morphological and genetic mapping study of bald colony mutants of *Streptomyces coelicolor*. J. Gen. Microbiol. **96**(2): 299-315.

Miguelez, E.M., Garcia, M., Hardisson, C., and Manzanal, M.B. 1994. Autoradiographic study of hyphal growth during aerial mycelium development in *Streptomyces antibioticus*. J. Bacteriol. **176**(7): 2105-2107.

Miller, I.M., Stapley, E.O., and Chaiet, L. 1962. Production of synnematin B by a member of the genus *Streptomyces*. Bacteriol.Proc. **49**: 32.

Morishita, R., Kawagoshi, A., Sawasaki, T., Madin, K., Ogasawara, T., Oka, T., and Endo, Y. 1999. Ribonuclease activity of rat liver perchloric acid-soluble protein, a potent inhibitor of protein synthesis. J. Biol. Chem. **274**(29): 20688-20692.

Mosher, R.H., Paradkar, A.S., Anders, C., Barton, B., and Jensen, S.E. 1999. Genes specific for the biosynthesis of clavam metabolites antipodal to clavulanic acid are clustered with the gene for clavaminate synthase 1 in *Streptomyces clavuligerus*. Antimicrob. Agents Chemother. **43**(5): 1215-1224.

Mullins, L.S., Zawadzke, L.E., Walsh, C.T., and Raushel, F.M. 1990. Kinetic evidence for the formation of D-alanyl phosphate in the mechanism of D-alanyl-D-alanine ligase. J. Biol. Chem. **265**(16): 8993-8998.

Nagarajan R (1972) β-lactam antibiotics from *Streptomyces. In* Cephalosporins and penicillins, chemistry and biology. *Edited by* Flynn E. H. Academic, New York, pp 636-661

Nicholson, N.H., Baggaley, K.H., Cassels, R., Davison, M., Elson, S.W., Fulston, M., Tyler, J.W., and Woroniecki, S.R. 1994. Evidence that the immediate biosynthetic precursor of clavulanic acid is its N-aldehyde analogue. J. Chem. Soci. D. (11): 1281-1282.

Ochi, K. 1986. Occurrence of the stringent response in Streptomyces sp. and its significance for the initiation of morphological and physiological differentiation. J. Gen. Microbiol. **132**(9): 2621-2631.

Ohnishi, Y., Kameyama, S., Onaka, H., and Horinouchi, S. 1999. The A-factor regulatory cascade leading to streptomycin biosynthesis in *Streptomyces griseus*: Identification of a target gene of the A-factor receptor. Mol. Microbiol. **34**(1): 102-111.

Okamoto, A., Nakai, Y., Hayashi, H., Hirotsu, K., and Kagamiyama, H. 1998. Crystal structures of *Paracoccus denitrificans* aromatic amino acid aminotransferase: A substrate recognition site constructed by rearrangement of hydrogen bond network. J. Mol. Biol. **280**(3): 443-461.

Onaka, H., and Horinouchi, S. 1997. DNA-binding activity of the A-factor

receptor protein and its recognition DNA sequences. Mol. Microbiol. **24**(5): 991-1000.

Onaka, H., Anido, N., Nihara, T., Yamada, Y., Beppu, T., and Horinouchi, S. 1995. Cloning and characterization of the A-factor receptor gene from *Streptomyces griseus*. J. Bacteriol. **177**:6083-6092.

Oxelmark, E., Marchini, A., Malanchi, I., Magherini, F., Jaquet, L., Hajibagheri, M.A.N., Blight, K.J., Jauniaux, J.-., and Tommasino, M. 2000. Mmf1p, a novel yeast mitochondrial protein conserved throughout evolution and involved in maintenance of the mitochondrial genome. Mol. Cell. Biol. **20**(20): 7784-7797.

Pao, S.S., Paulsen, I.T., and Saier Jr., M.H. 1998. Major facilitator superfamily. Microbiol. Mol. Biol. Rev. **62**(1): 1-34.

Paradkar, A.S., Aidoo, K.A., and Jensen, S.E. 1998. A pathway-specific transcriptional activator regulates late steps of clavulanic acid biosynthesis in *Streptomyces clavuligerus*. Mol. Microbiol. **27**:831-843.

Paradkar, A.S., and Jensen, S.E. 1995. Functional analysis of the gene encoding the clavaminate synthase 2 isoenzyme involved in clavulanic acid biosynthesis in Streptomyces clavuligerus. J. Bacteriol. **177**(5): 1307-1314.

Pérez-Llarena, F.J., Liras, P., Rodríguez-García, A., and Martin, J.F. 1997. A regulatory gene (*ccaR*) required for cephamycin and clavulanic acid production in *Streptomyces clavuligerus*: Amplification results in overproduction of both β-lactam compounds. J. Bacteriol. **179**(6): 2053-2059.

Pérez-Redondo, R., Rodríguez-García, A., Martín, J.F., and Liras, P. 1998. The claR gene of *Streptomyces clavuligerus*, encoding a LysR-type regulatory protein controlling clavulanic acid biosynthesis, is linked to the clavulanate-9-aldehyde reductase (*car*) gene. Gene **211**(2): 311-321.

Pitlik, J., and Townsend, C.A. 1997. The fate of [2,3,3-2H3, 1,2-13C2]-D, L-glycerate in clavulanic acid biosynthesis. Chem. Comm. (2): 225-226.

Pyatkin K. 1967. Microbiology. Mir, Moscow.

Rappu, P., Shin, B.S., Zalkin, H., and Mäntsälä, P. 1999. A role for a highly conserved protein of unknown function in regulation of *Bacillus subtilis purA* by the purine repressor. J. Bacteriol. **181**(12): 3810-3815.

Reading, C., and Cole, M. 1977. Clavulanic acid: a beta lactamase inhibiting beta lactam from *Streptomyces clavuligerus*. Antimicrob. Agents Chemother. **11**(5): 852-857.

Retzlaff, L., and Distler, J. 1995. The regulator of streptomycin gene expression, StrR, of *Streptomyces griseus* is a DNA binding activator protein with multiple

recognition sites. Mol. Microbiol. **18**(1): 151-162.

Rhee, S., Silva, M.M., Hyde, C.C., Rogers, P.H., Metzler, C.M., Metzler, D.E., and Arnone, A. 1997. Refinement and comparisons of the crystal structures of pig cytosolic aspartate aminotransferase and its complex with 2-methylaspartate. J. Biol. Chem. **272**(28): 17293-17302.

Roach, P.L., Clifton, I.J., Fulop, V., Harlos, K., Barton, G.J., Hajdu, J., Andersson, I., Schofield, C.J., and Baldwin, J.E. 1995. Crystal structure of isopenicillin N synthase is the first from a new structural family of enzymes. Nature **375**(6533): 700-704.

Röhl, F., Rabenhorst, J., and Zahner, H. 1987. Biological properties and mode of action of clavams. Arch. Microbiol. **147**(4): 315-320.

Romero, J., Liras, P., and Martin, J.F. 1986. Utilization of ornithine and arginine as specific precursosrs of clavulanic acid. Appl. Environ. Microbiol. **52**(4): 892-897.

Salowe, S.P., Marsh, E.N., and Townsend, C.A. 1990. Purification and characterization of clavaminate synthase from *Streptomyces clavuligerus*: An unusual oxidative enzyme in natural product biosynthesis. Biochem. (N. Y.) **29**(27): 6499-6508.

Salowe, S.P., Krol, W.J., Iwata-Reuyl, D., and Townsend, C.A. 1991. Elucidation of the order of oxidations and identification of an intermediate in the multistep clavaminate synthase reaction. Biochem. (N. Y.) **30**(8): 2281-2292.

Sambrook, J., Fritsch, E.F., and Maniatas, T. 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Santamarta, I., Pérez-Redondo, R., Lorenzana, L.M., Martín, J.F., and Liras, P. 2005. Different proteins bind to the butyrolactone receptor protein ARE sequence located upstream of the regulatory *ccaR* gene of *Streptomyces clavuligerus*. Mol. Microbiol. **56**(3): 824-835.

Santamarta, I., Rodríguez-García, A., Pérez-Redondo, R., Martín, J.F., and Liras, P. 2002. CcaR is an autoregulatory protein that binds to the *ccaR* and *cefD-cmcI* promoters of the cephamycin C-clavulanic acid cluster in *Streptomyces clavuligerus*. J. Bacteriol. **184**(11): 3106-3113.

Scarsdale, J.N., Radaev, S., Kazanina, G., Schirch, V., and Wright, H.T. 2000. Crystal structure at 2.4 Å resolution of *Escherichia coli* serine hydroxymethyltransferase in complex with glycine substrate and 5-formyl tetrahydrofolate. J. Mol. Biol. **296**(1): 155-168.

Schmitz, G., and Downs, D.M. 2004. Reduced Transaminase B (IlvE) Activity

Caused by the Lack of *yjgF* Is Dependent on the Status of Threonine Deaminase (IlvA) in *Salmonella enterica* Serovar *Typhimurium*. J. Bacteriol. **186**(3): 803-810.

Schrempf, H., Bujard, H., Hopwood, D.A., and Goebel, W. 1975. Isolation of covalently closed circular deoxyribonucleic acid from *Streptomyces coelicolor* A3(2). J. Bacteriol. **121**(2): 416-421.

Sheldon, P.J., Busarow, S.B., and Hutchinson, C.R. 2002. Mapping the DNA-binding domain and target sequences of the Streptomyces peucetius daunorubicin biosynthesis regulatory protein, DnrI. Mol. Microbiol. **44**(2): 449-460.

Simanshu, D.K., Savithri, H.S., and Murthy, M.R.N. 2006. Crystal structures of *Salmonella typhimurium* biodegradative threonine deaminase and its complex with CMP provide structural insights into ligand-induced oligomerization and enzyme activation. J. Biol. Chem. **281**(51): 39630-39641.

Song, J.Y., Kim, E.S., Kim, D.W., Jensen, S.E., and Lee, K.J. 2009. A gene located downstream of the clavulanic acid gene cluster in *Streptomyces clavuligerus* ATCC 27064 encodes a putative response regulator that affects clavulanic acid production. J. Ind. Microbiol. Biotechnol. **36**(2): 301-311.

Stackebrandt, E., Rainey, F.A., and Ward-Rainey, N.L. 1997. Proposal for a new hierarchic classification system, *Actinobacteria classis* nov. Int. J. Sys. Bacteriol. **47**(2): 479-491.

Stanley T., Williams M., Sharpe E., Holt J. G. 1983. Bergey's manual of systematic bacteriology. Williams and Wilkins.

Sykes R. B., Bush K. 1982. Physiology, Biochemistry, and inactivation of β-lactamases *In* Chemistry and biology of β-lactam antibiotics. *Edited by* Morin R. B., Gorman M. Academic Press Inc., New York, pp: 155-208.

Tahlan, K. 2005. Characterization of the carboxyethylarginine synthase genes involved in clavulanic acid and 5S clavam metabolite biosynthesis in Streptomyces clavuligerus. Ph.D. thesis, University of Alberta, Edmonton, Alta.

Tahlan, K., Anders, C., and Jensen, S.E. 2004b. The paralogous pairs of genes involved in clavulanic acid and clavam metabolite biosynthesis are differently regulated in *Streptomyces clavuligerus*. J. Bacteriol. **186**:6286-6297

Tahlan, K., Park, H.U., and Jensen, S.E. 2004a. Three unlinked gene clusters are involved in clavam metabolite biosynthesis in *Streptomyces clavuligerus*. Can. J. Microbiol. **50**(10): 803-810.

Tahlan, K., Anders, C., Wong, A., Mosher, R.H., Beatty, P.H., Brumlik, M.J.,

Griffin, A., Hughes, C., Griffin, J., Barton, B., and Jensen, S.E. 2007. 5*S* Clavam Biosynthetic Genes Are Located in Both the Clavam and Paralog Gene Clusters in *Streptomyces clavuligerus*. Chem. Biol. **14**(2): 131-142.

Thirkettle, J.E., Baldwin, J.E., Edwards, J., Griffin, J.P., and Schofield, C.J. 1997. The origin of the β-lactam carbons of clavulanic acid. Chem. Comm. (11): 1025-1026.

Thoden, J.B., Blanchard, C.Z., Holden, H.M., and Waldrop, G.L. 2000. Movement of the biotin carboxylase B-domain as a result of ATP binding. J. Biol. Chem. **275**(21): 16183-16190.

Thoden, J.B., Miran, S.G., Phillips, J.C., Howard, A.J., Raushel, F.M., and Holden, H.M. 1998. Carbamoyl phosphate synthetase: Caught in the act of glutamine hydrolysis. Biochemistry (N. Y.) **37**(25): 8825-8831.

Tipper, D.J., and Strominger, J.L. 1965. Mechanism of action of penicillins: a proposal based on their structural similarity to acyl-D-alanyl-D-alanine. Proc. Natl. Acad. Sci. **54**(4): 1133-1141.

Townsend, C.A., and Krol, W.J. 1988. The role of molecular oxygen in clavulanic acid biosynthesis: evidence for a bacterial oxidative deamination. J. Chem. Soc. D. (18): 1234-1236.

Townsend, C.A., and Ho, M. 1985. Biosynthesis of clavulanic acid: Origin of the C5 unit. J. Am. Chem. Soc. **107**(4): 1065-1066.

Trepanier, N.K., Jensen, S.E., Alexander, D.C., and Leskiw, B.K. 2002. The positive activator of cephamycin C and clavulanic acid production in *Streptomyces clavuligerus* is mistranslated in a *bldA* mutant. Microbiology **148**(3): 643-656.

Udaka, S., and Kinoshita, S. 1958. Studies on L-ornithine fermentation. I. The biosynthetic pathway of L-ornithine in *Micrococcus glutamicus*. J. Gen. Appl. Microbiol. **4**: 272-282.

Valentine, B.P., Bailey, C.R., Doherty, A., Morris, J., Elson, S.W., Baggaley, K.H., and Nicholson, N.H. 1993. Evidence that arginine is a later metabolic intermediate than ornithine in the biosynthesis of clavulanic acid by *Streptomyces clavuligerus.* J. Chem. Soc. D. (15): 1210-1211.

Vujaklija, D., Horinouchi, S., and Beppu, T. 1993. Detection of an A-factor-responsive protein that binds to the upstream activation sequence of *strR*, a regulatory gene for streptomycin biosynthesis in *Streptomyces griseus*. J. Bacteriol. **175**(9): 2652-2661.

Waksman, S.A., and Henrici, A.T. 1943. The nomenclature and classification of the actinomycetes. J. Bacteriol. **46**: 337-341.

Wang, W., Kappock, T.J., Stubbe, J., and Ealick, S.E. 1998. X-ray crystal structure of glycinamide ribonucleotide synthetase from *Escherichia coli*. Biochem. (N. Y.) **37**(45): 15647-15662.

Ward, J.M., and Hodgson, J.E. 1993. The biosynthetic genes for clavulanic acid and cephamycin production occur as a 'super-cluster' in three *Streptomyces*. FEMS Microbiol. Lett. **110**(2): 239-242.

Wietzorrek, A., and Bibb, M. 1997. A novel family of proteins that regulates antibiotic production in streptomycetes appears to contain an OmpR-like DNA-binding fold [3]. Mol. Microbiol. **25**(6): 1181-1184.

Wildermuth, H. 1970. Development and organization of the aerial mycelium in *Streptomyces coelicolor*. J. Gen. Microbiol. **60**(1): 43-50.

Wise Jr., E.M., and Park, J.T. 1965. Penicillin: its basic site of action as an inhibitor of a peptide cross-linking reaction in cell wall mucopeptide synthesis. Proc. Natl. Acad. Sci. **54**(1): 75-81.

Wu, T., Busby, R.W., Houston, T.A., McIlwaine, D.B., Egan, L.A., and Townsend, C.A. 1995. Identification, cloning, sequencing, and overexpression of the gene encoding proclavaminate amidino hydrolase and characterization of protein function in clavulanic acid biosynthesis. J. Bacteriol. **177**(13): 3714-3720.

Xu, W., Jones, C.R., Dunn, C.A., and Bessman, M.J. 2004. Gene *ytkD* of Bacillus subtilis encodes an atypical nucleoside triphosphatase member of the Nudix hydrolase superfamily. J. Bacteriol. **186**(24): 8380-8384.

Yamada, T., Komoto, J., Takata, Y., Ogawa, H., Pitot, H.C., and Takusagawa, F. 2003. Crystal Structure of Serine Dehydratase from Rat Liver. Biochemistry (N. Y.) **42**(44): 12854-12865.

Yamazaki, H., Tomono, A., Ohnishi, Y., and Horinouchi, S. 2004. DNA-binding specificity of AdpA, a transcriptional activator in the A-factor regulatory cascade in *Streptomyces griseus*. Mol. Microbiol. **53**(2): 555-572.

Yamazaki, H., Takano, Y., Ohnishi, Y., and Horinouchi, S. 2003. *amfR*, an essential gene for aerial mycelium formation, is a member of the AdpA regulon in the A-factor regulatory cascade in *Streptomyces griseus*. Mol. Microbiol. **50**(4): 1173-1187.

Yanto, Y., Hall, M., and Bommarius, A.S. 2010. Nitroreductase from *Salmonella typhimurium*: Characterization and catalytic activity. Organic and Biomolecular Chemistry **8**(8): 1826-1832.

Zelyas, N.J., 2007. Identification and analysis of genes involved in 5*S* clavam biosynthesis in *Streptomyces clavuligerus*. M.Sc. thesis, University of Alberta,

Edmonton, Alta.

Zelyas, N.J., Cai, H., Kwong, T., and Jensen, S.E. 2008. Alanylclavam biosynthetic genes are clustered together with one group of clavulanic acid biosynthetic genes in *Streptomyces clavuligerus*. J. Bacteriol. **190**(24): 7957-7965.