USING GENE SHUFFLING TO INCREASE GENETIC DIVERSITY IN GENES INVOLVED IN BETA-LACTAM BIOSYNTHESIS

THESIS

Submitted in fulfilment of the requirements for the degree of MASTER OF SCIENCE of Rhodes University

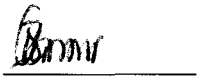
by

SHAHIDA TARR

January 2001

DECLARATION

In accordance with the conditions for the award of the degree of Master of Science, I declare that the work presented in this thesis is my own original research. Any assistance with experimental procedure or technique has been declared elsewhere when necessary. To the best of my knowledge, neither the whole or part of this thesis has been, is being or will be submitted before for any degree or examination to any other University.



Shahida Tarr

This work is dedicated to the following persons:

My parents, Amaanullah and Ashroon Hassim Tarr

My brothers Iqbal and Salim

My sister Sharifa

My nieces Ashroon and Ismat

My nephew Ziyaad

TABLE OF CONTENTS

TABLE OF CONTENTS ii

ABSTRACT vi

LIST OF FIGURES viii

LIST OF TABLES xi

LIST OF ABBREVIATIONS xii

ACKNOWLEDGEMENTS xiii

CHAPTER 1: LITERATURE REVIEW

1.1 Introduction 1

1.2 History of antibiotic discovery 3

1.3 Classification of β-lactam compounds 5

1.4 Biosynthesis of β-lactam compounds 8

1.4.1 Biosynthesis of Penicillin 9

1.4.2 Biosynthesis of Cephalosporins and Cephamycins 9

1.4.3 Common genes to all penicillin and cephalosporin producers (*pcbC*) gene 13

1.4.4 Isopenicillin N Synthase (*pcbC*) gene 14

1.5 Biosynthetic production of new antibiotics 15

1.6 DNA shuffling 17

1.6.1 Principle of DNA shuffling 18

1.7 Purpose of the study 20

CHAPTER 2: GENERAL MATERIALS AND METHODS

2.1 Introduction 21

2.2 Maintenance of β-lactam producing Streptomyces using different media 22

2.2.1 Culture preservation 23

2.2.2 Total DNA extraction 23

2.2.3 Quantification of DNA 25

2.2.4 Detection and analysis of DNA 26

2.2.4.1 Analysis of DNA 26

2.2.5 Bioassays 26

2.2.5.1 Agar Plug method 27

2.2.5.2 Cup plate method 27

2.3.7 Protein assay 27

2.3 Results 28

2.4 Discussion 32

CHAPTER 3: DETECTION OF IPNS GENE PRODUCTS

3.1 Introduction 34

3.1.1 Primer design 35

3.1.2 PCR programme 36

3.1.3 Cloning 36

3.1.4 Selection of positive transformed bacteria 37

3.1.5 pGEM-T-Cloning 37

3.2 Materials and methods 38

3.2.1 Oligonucleotide Design 38

3.2.2 Preparation of competent cells 41

3.2.3. Cloning of the PCR product 42

3.2.4. Long-term storage of clones 43

3.2.5 Plasmid DNA extraction 43

3.3 Results 45

3.4 Discussion 50

CHAPTER 4: CLONING AND EXPRESSION OF THE IPNS GENE

4.1 Introduction 51

4.1.1 TOPO TA Cloning 51

4.1.2 HPLC 53

4.2 Materials and methods 54

4.2.1 TOPO Cloning 54

4.2.2 Expression of IPNS protein in *E. coli* 54

4.2.3 Analysis of time point samples 55

4.2.4 Preparation of *Streptomyces* cell-free extracts 56

4.2.5 IPNS Assay Using HPLC 56

4.3 Results 57

4.4 Discussion 66

CHAPTER 5: GENE SHUFFLING

5.1 Introduction 68

5.2 Materials and Methods 69

5.2.1 Gene shuffling methods 69

5.2.2 Agarose gel electrophoresis 70

5.2.3 Restriction digestion of IPNS gene 71

5.3 Results 72

5.4 Discussion 80

CHAPTER 6: BIOINFORMATIC STUDIES

6.1 Introduction 82

6.1.1 Bioinformatic database 83

6.2 Materials and Methods 84

6.2.1 Preparation of samples for sequencing 84

6.3 Results 85

6.4 Discussion 112

CHAPTER 7: FINAL DISCUSSION 115

APPENDICES 118

REFERENCES 136

ABSTRACT

The actinomycetes are gram-positive bacteria that produce more than two-thirds of the known biologically active microbial natural products, including many commercially important antibiotics, anti-cancer agents, other pharmacologically useful agents, animal health products and agrochemicals.

The prevailing utilization of antibiotics continues to be the mainstay against microbial infections and a majority of the over six thousand antibiotics discovered thus far are from *Streptomyces spp.* One of the most well-characterized antibiotic biosynthetic pathway is the one involving the biosynthesis of the penicillins, cephalosporins and cephamycins. This pathway involves two initial steps which are common in filamentous fungi, lower eukaryotes and prokaryotes. The penam nucleus of penicillins and the cephem nucleus of both cephamycins and cephalosporins are formed by the condensation of the three precursor amino acids L-α-aminoadipic acid, L-cysteine and L-valine by a mechanism designated as 'non-ribosomal peptide synthesis', which involves activation and condensation of the three component amino acids and epimerization of the L- to D-valine to form a linear acyclic tripeptide called δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine (ACV) by the action of a peptide synthetase. ACV is then cyclized to form isopenicillin N, an intermediate that contains an L-α-aminoadipyl side-chain attached to the penem nucleus (Fig. 1.2) by isopenicilin N synthase (IPNS or Cyclase) and this encompasses the creation of the Beta-lactam and thiazolidine rings.

A broad range of β-lactam producing *Streptomyces spp* were grown, the DNA extraction procedure optimised and total chromosomal DNA isolated. A bioinformatics analysis of known

IPNS gene sequences allowed the synthesis of PCR primers for the iso-penicillin N synthase gene. IPNS genes and IPNS-like genes were successfully amplified from the total DNA of ten strains including two novel thermophilic strains, A and B. Sequencing was carried out on the genes from *S. hygroscopicus*, *S.tanashiensis* and the two thermophiles A and B. This allowed development of the conditions for gene shuffling of the IPNS gene which was carried out pairwise and resulted in the reconstitution of shuffled genes of the correct size. The resulting mixed gene sequences were cloned into the pTrcHis2-TOPO expression vector and the plasmid DNA screened and assayed for IPNS activity using HPLC which showed ten fold increase in IPNS activity as a result of the shuffling.

LIST OF FIGURES

Page

Figure 1.1. Chemical structures of β-lactam compounds. 7

Figure 1.2. Biosynthetic pathway of β-lactam antibiotics including the enzyme and gene designation. 12

Figure 1.3. Outline of the steps involved in DNA shuffling. 18

Figure 2.1. Agarose gel electrophoresis of total DNA using method adapted from mini-prep of bacterial genomic DNA. 28

Figure 2.2. Graph showing DNA concentrations obtained with the two methods. 29

Figure 2.3. Protein concentrations using Bradford's assay. 31

Figure 3.1. Circular vector map of p-GEM T Easy vector. 37

Figure 3.2. IPNS gene product of *Amycolatopsis spp.* 45

Figure 3.3. IPNS gene from novel Thermophilic A and B. 46

Figure 3.4. IPNS gene from *S. clavuligerus.* 47

Figure 3.5. Plasmid DNA using High Pure. 48

Figure 3.6. Plasmid DNA using CTAB mini-prep. 48

Figure 3.7. Plasmid DNA restricted with Eco RI. 49

Figure 3.8. Plasmid DNA restricted with Eco RI from thermophilic strains. 49

Figure 4.1. Circular vector map of pTrcHis2-TOPO. 52

Figure 4.2. SDS-PAGE gel of *S. griseus.* 57

Figure 4.3. SDS-PAGE gel of Thermophilic B. 58

Figure 4.4. SDS-PAGE gel of Thermophilic A. 59

Figure 4.5. SDS-PAGE gel of *S. tanashiensis.* 59

Figure 4.6. SDS-PAGE gel of *Amycolatopsis spp.* 60

Figure 4.7. SDS-PAGE of *S. cattleya*. 60

Figure 4.8. HPLC chromatogram of 6-APA. 61

Figure 4.9. HPLC chromatogram of ampicillin. 62

Figure 4.10. HPLC chromatogram of Pen VK. 63

Figure 4.11. HPLC chromatogram of Thermophilic A. 64

Figure 4.12. HPLC chromatogram of Thermophilic A after expression. 65

Figure 5.1. Agarose gel electrophoresis of shuffled PCR product. 72

Figure 5.2. Agarose gel electrophoresis of shuffled IPNS gene. 74

Figure 5.3. Change in base pairs of IPNS gene after shuffling. 75

Figure 5.4. Restriction digest of 8 strains before shuffling. 76

Figure 5.5. Restriction digest of 2 strains. 77

Figure 5.6. HPLC chromatogram of *S. cattleya* before and after shuffling. 78

Figure 5.7. HPLC retention times before and after cloning. 79

Figure 6.1. Electropherogram from *S. hygroscopicus*. 85

Figure 6.2. Nucleotide sequence of IPNS gene from *S. hygroscopicus*. 87

Figure 6.3. Percentage identity matrix of *S. hygroscopicus*. 87

Figure 6.4. Nucleotide sequence alignment of *S. hygroscopicus*. 90

Figure 6.5. Predicted Open reading frames of *S. hygroscopicus* 90

Figure 6.6. Nucleotide sequence of IPNS gene from Thermophilic A. 92

Figure 6.7. Percentage identity matrix for Thermophilic A. 92

Figure 6.8. Nucleotide alignment of Thermophilic A. 95

Figure 6.9. Open reading frames for Thermophilic A 96

Figure 6.10. Nucleotide sequence of IPNS gene from Thermophilic B. 97

Figure 6.11. Percentage identity matrix of Thermophilic B. 97

Figure 6.12. Nucleotide sequence alignment of Thermophilic B. 100

Figure 6.13. Electropherogram of *S. tanashiensis*. 101

Figure 6.14. Nucleotide sequence from *S. tanashiensis*. 102

Figure 6.15. Percentage identity matrix for *S. tanashiensis* 103

Figure 6.16. Nucleotide alignment sequence of *S. tanashiensis*. 106

Figure 6.17. Predicted open reading frames for *S. tanashiensis* 107

Figure 6.18. Nucleotide sequence of shuffled *S. jumonjinensis*. 109

Figure 6.19. Nucleotide sequence alignment of *S. jumonjinensis*. 111

LIST OF TABLES

Page

Table 2.1. Antibacterial activity against selected indicator organisms. 30

Table 3.1. Sequences of the primers used. 38

Table 3.2. Reagents and final concentrations used in all PCR reactions. 39

Table 3.3. PCR conditions used for IPNS amplification. 40

Table 3.4. Volumes of reagents used for PCR reactions. 40

Table 3.5. Ingredients of the ligation reaction used with P-GEM T Easy vector. 42

Table 3.6. Sizes of IPNS genes. 46

Table 4.1. Retention times of standard antibiotics. 63

Table 4.2. Retention times of the *Streptomyces spp*. 64

Table 5.1. Volumes of reagents for primer less PCR. 69

Table 5.2. Reassembly reagents. 70

Table 5.3. Restriction digest Cocktail. 71

Table 5.4. Change in size as a result of shuffling 73

Table 6.1. Percentage identity of the ORF against documented strains 108

LIST OF ABBREVIATIONS

Bp	= Base Pair
CTAB	= Cetyl dimethyl ethylammonium bromide
DNA	= Deoxy Ribo Nucleic Acid
Hrs	= Hours
IPTG	= Isopropyl-β-D-Thiogalactopyranoside
Min	= Minutes
nm	= Nano metre
OD	= Optical Density
ORF	= Open Reading Frame
PCR	= Polymerase Chain Reaction
Pen VK	= Phenoxymethylpenicillin
Sec	= Seconds
X Gal	= 5-Bromo-4-Chloro-3-Indolyl-β-D-Galactoside

ACKNOWLEDGEMENTS

My Sincere thanks to the following people:

My Supervisor Professor Ralph Kirby for his support, guidance, enthusiasm and above all for always being ready to trouble-shoot when necessary.

The World Health Organization for funding my studies.

The Ministry of Health Lesotho for granting me study leave.

Professor Greg Blatch for allowing free access to his computing facilities.

Mrs Anna Clark for being a colleague/ friend and all the help with sequencing.

Miss Manjusha Joseph, my best friend who supported me in everything and made many of those long working hours bearable.

Miss Carol Hartley for her advise, suggestions and helpful discussion whenever things were not working in the lab.

Miss Vicky Longshaw for her invaluable help with bioinformatic analysis and proofreading this thesis.

My Lab mates in 410 for making working in the lab fun.

The Staff and post-graduate students of the department of Biochemistry & Microbiology for their support throughout my studies.

The Technical staff for their role in making my work in the department possible.

Finally to my Parents whose moral, financial and emotional support made this thesis a reality.

CHAPTER 1

LITERATURE REVIEW

1.1 INTRODUCTION

Streptomyces are Gram-positive aerobic soil bacteria numbering about 10^6 to 10^8 germs per gram of soil, where they make up one of the dominant groups, owing to their optimum adaptation to their natural environment (Madduri *et al.* 1991; Aidoo *et al.* 1994). The *Streptomyces* exhibit a differential growth cycle, starting with the germination of spores and proceeding to the formation of substrate mycelia (Shima *et al.* 1996). The cycle closes with the formation of a multigenomic aerial hyphae, which impart a characteristic fuzzy appearance to the developing colonies and subsequently differentiate to form chains of exospores. These spores are relatively resistant to heat, dryness, cold and thus the bacteria's survival is guaranteed in periods when no nutrients are available (Hopwood 1978,1988; Salas and Mendez 1998; Walter *et al.* 1998). It should be noted that the spores of *Streptomyces* are not endospores and are thus not highly thermotolerant like the spores of *Bacillus spp.*

The sporulation in *Streptomyces* is characterized by the formation of these chains of uninucleoidal, unicellular spores by growth from multinucleoidal branched filamentous cells. The two distinctive features of the *Streptomyces* are its reproductive nature and the requirement for extensive growth during differentiation (Hao and Kendrick 1998). The development of aerial hyphae in *Streptomyces* has been defined genetically by the isolation of *bld* mutants that lack aerial mycelium and, therefore have a shiny 'bald' appearance. Unlike a second class of developmental mutations, *whi* mutations

which appear only to affect the differentiation of aerial hyphae into spores, *bld* mutations have pleiotropic effects that generally prevent antibiotic production and cause defects in carbon catabolite repression in addition to blocking differentiation (Keleman and Buttner 1998). On solid media, a spore germinates to grow as a multinucleoid substrate mycelium of branching hyphae. In response to nutrient limitation, older parts of the substrate mycelium produce aerial mycelium. After septa have been produced at regular intervals along the hyphae to form uninucleoid compartments, long chains of spores are formed (Hopwood 1988; Chhabra *et al.* 1993; Kudo *et al.* 1995).

Streptomyces are well known for producing a wide range of commercially important antibiotics as their secondary metabolites (Hopwood 1978; Bibb *et al.* 1980; Chater and Bruton 1983; Martin *et al.* 1989; Davies 1990; Forsman *et al.* 1990; Madduri *et al.* 1991; Kieser *et al.* 1992; Ludovice *et al.* 1992; Paradkar *et al.* 1996; Sanchez and Brana 1996), producing about two-thirds of naturally occurring antibiotics and a wide array of other secondary metabolites (Aceti and Champness 1998; Baltz 1998). The antibiotics produced include penicillin, chloramphenicol, streptomycin, tetracycline, erythromycin and many others (Fang and Demain 1995). Antibiotic production by these bacteria is regulated by a variety of physiological and nutritional conditions and is also linked to the morphological development of the organism including the formation of aerial mycelia and sporulation (Hopwood *et al.* 1987; Petrich *et al.* 1992; Cochet and Demain 1996; Paradkar *et al.* 1998). Antibiotics produced by actinomycetes are diverse both in terms of their chemical structure and their mode of action. With the current knowledge of actinomycete biology and the use of molecular techniques, there is still room for novel antibiotics to be discovered (Floss 1987; Kirby 1992).

1.2 History of Antibiotic Discovery

Antibiotics have been the cornerstone of antibacterial therapy worldwide; these agents are active against identified diseases in humans, animals and plants (Courvalin and Davies 1999). Probably the most important discovery in the history of therapeutic medicine started with the discovery of the efficacy of a beta-lactam compound, as clinical antibiotics thus becoming the staging point for large scale investment of the pharmaceutical industry into antibiotic research. In the early 1930's, Professor Domagk discovered at the Bayer's laboratories that particular sulphonomides had an antibacterial activity. His discovery, published in 1935 opened a new era in medicine. The clinical successes of the sulphonamide therapy drastically reduced death caused by infectious diseases. The well-known discovery of penicillin in 1929, when Alexander Fleming observed an inhibition of growth of *Staphylococcus aureus* on an agar plate contaminated with *Penicillium notatum* predated this, it was only three years later that it was actually shown that the growth inhibition was due to a chemical penicillin.

The first clinical trials with penicillin were undertaken much later in 1941 which culminated in an effort to provide penicillin in large amounts for the 2nd World War military medical services and this required that the structure of this compound be elucidated. Hodgkin and Low used X-ray crystallography to show that the compound was composed of a β-lactam structure. Penicillin was the first true antibiotic to be discovered and is still, seventy years later, one of the most important in terms of both therapeutic use and annual volume of production (Penalva *et al.* 1998). It is clear that even at seventy years of age, β-lactams are going strong and no one contemplates their early retirement (Demain 1991). The β-lactam antibiotics represented by many penicillins and the

structurally related cephalosporins are still by far the most frequently prescribed antimicrobial agents, despite the availability of an ever-growing number of alternative antimicrobials (Skatrud 1992).

In 1941, a bacteriologist named Giuseppe Brotzu isolated from a species of *Cephalosporium acremonium* (now called *Acremonium chrysogenum*) a substance with strong antibacterial activity, and was named Cephalosporin C. Later, it was established that this fungus produced several different antibiotics which are now called cephalosporins (Valegard *et al.* 1998).

Further studies which included both chemical and X-ray studies established that Cephalosporin C contained a β-lactam ring. This ring was found to be resistant to penicillinases (β-lactamases) and this class of antibiotics was named Cephalosporins. The discovery of Cephalosporin C generated a whole new group of clinically significant β-lactams.

The success and prevailing utilization of β-lactams in the treatment of diseases/bacterial infections continue to be the mainstay against microbial infections and they form a significant part of the majority of over six thousand antibiotics discovered thus far which are from *Streptomyces spp* (Loke *et. al.* 2000). They are the most widely used antimicrobials accounting for over 50% of the world market in antibiotics. The other factors that prompt the use of β-lactams are their specificity and low toxicity (Doran *et al.* 1990; Petrich *et al.* 1992; Durairaj and Jensen 1996; Brakhage 1998). The overriding interest in these compounds and in finding new β-lactams stems from the severe clinical problems that still exists in the treatment of infectious disease, and because of the unique mode of

action of many of these antibiotics as inhibitors of bacterial cell-wall biosynthesis which is carried out by preventing cross-linking of small peptide chains in peptidoglycan, the main wall polymer of bacteria. Pre-existing cells are unaffected, but all newly produced cells grow abnormally, unable to maintain their wall rigidity, and hence become susceptible to osmotic lysis (Cohen *et al.* 1990). Antibiotics may be static or bactericidal in action, although the reasons for different activities have not been completely defined. For example, why is chloramphenicol a bacteriostatic drug for most bacterial species but has a killing activity against *Haemophilus influenzae*. Many questions remain about antibiotics but their use remains unabated (Courvalin and Davies 1999).

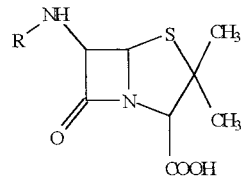
1.3 Classification of β-lactam compounds

The β-lactam antibiotics comprise several classes of compounds, among which the cephalosporins and the penicillins are most important. Both classes contain bulky side chain attached to the 7-aminocephalosporanic acid or 6-aminopenicillanic acid nuclei respectively (Niessen 1998).

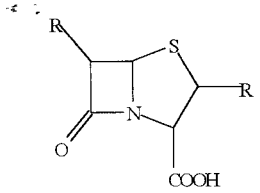
The classical β-lactam antibiotics can be divided into hydrophobic and hydrophilic members. The former include benzylpenicillin (Penicillin G) and phenoxymethylpenicillin (Penicillin V) which contain the non polar side chains, phenylacetate and phenoxyacetate respectively. The antibacterial spectrum of hydrophobic penicillins is essentially gram positive. The hydrophilic types are penicillin N, cephalosporins and 7-α-methoxycephalosporins (cephamycins) which are made amongst others by actinomycetes. They contain the polar side chain D-α-aminoadipate, and possess a broad-spectrum activity (Demain 1991; Ambedkar *et al.* 1997).

The two major groups of β-lactam compounds produced are the sulphur-containing β-lactam which includes the penicillins, cephalosporins and cephamycins and the oxygen-containing β-lactam compounds commonly referred to as clavams (McGowan *et al.* 1998). This group includes clavulanic acid as well as a number of structurally related compounds. Clavulanic acid has potent β-lactamase inhibitory properties and is used clinically in combination with β-lactam antibiotics to combat infections caused by β-lactamase producing microorganisms (Paradkar and Jensen 1995).

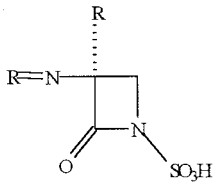
The β-lactam compounds can also be classified conveniently on the basis of the chemical structures into five groups viz, penicillin, cephalosporins, cephamycins, clavams, carbapenems and the monocyclic β-lactam (Figure 1.1) (Garcia-Dominguez *et al.* 1991).



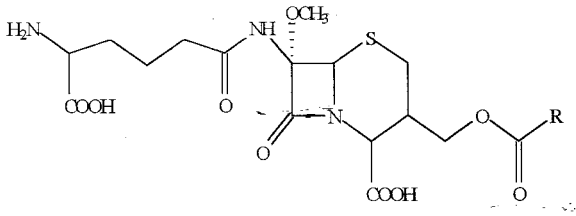
PENICILLINS



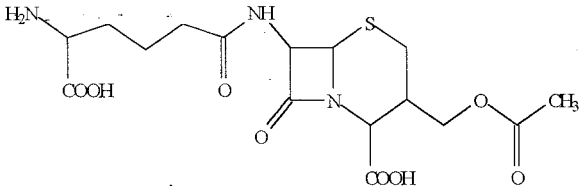
CARBAPENEMS



MONOBACTAMS



CEPHAMYCINS



CEPHALOSPORIN C

Figure 1.1. Chemical structures of β-lactam compounds.

These antibiotics are distinguished by the β-lactam ring plus either a five-membered thiazolidine ring (present in the penicillins) or the 'expanded' six-membered dihydrothiazine structure (found in cephalosporins). This is the main chemical difference between the core structures of these antibiotics (Kovacevic *et al.* 1990). The monolactams have a single ring whereas all the other β-lactam consist of a bicyclic ring system. Penicillins, cephalosporins and cephamycin each contain a sulphur atom in the fused ring and have an early biosynthetic pathway. They may be regarded as members of a single 'sulphur-containing class' of β-lactams. The other classes are characterized by either the presence within the fused ring of oxygen (clavams) or carbon (carbapenems) atoms or the absence of a fused ring (monobactams) (McGowan *et al.* 1998). The ability to synthesize β-lactams is widespread in nature, being found among the fungi, the Gram- positive bacteria and the Gram-negative bacteria.

1.4 Biosynthesis of β-lactam compounds

The biochemistry of penicillin and cephalosporins is now rather well understood and together with the use of recombinant techniques, it has now become possible to elucidate the molecular regulation of β-lactam biosynthesis which is controlled by a complex regulatory network. The two initial steps in the biosynthetic pathway are essentially alike in filamentous fungi, lower eukaryotes and prokaryotes (Romero *et al.* 1984; Aharonowitz *et al.* 1992; Khetan *et al.* 1996; Loke *et al.* 2000). This is probably because the pathway had a single evolutionary origin and spread to the wide range of producing micro-organisms by horizontal gene transfer (Kirby pers. comm 1999).

1.4.1 Biosynthesis of penicillin

Studies on biosynthesis of fungal β–lactams using cell-free systems led to the crucial observation that a cysteine-containing tripeptide, δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine (ACV), was the earliest direct precursor of these antibiotics (Fang *et al.* 1996).

ACV is then cyclized to form isopenicillin N, an intermediate which has an L-α-aminoadipyl side chain attached to the nucleus of penicillin (Ramos *et al.* 1985; Williamson *et al.* 1985; Madduri *et al.* 1989, 1991) (Figure 1.2). The α-aminoadipyl side chain is later exchanged for phenylacetic acid in penicillin producing microorganisms but not in cephalosporin producers.

1.4.2 Biosynthesis of cephalosporins and cephamycins

The biosynthetic pathways of cephalosporins and cephamycins resemble that of penicillin (Dylan and Jensen 1998). ACV and isopenicillin N are, in fact intermediates in the biosynthesis of both groups of β-lactams, isopenicillin N is formed by cyclization of the ACV tripeptide, as in penicillin biosynthesis. The bicyclic ring structure of this compound has been referred to as the "enchanted" ring because it possesses the ability to inhibit the transpeptidase function of enzymes involved in bacterial cell wall synthesis eventually leading to cell death by lysis. The cyclases of the actinomycetes, *S. clavuligerus*, *S. lipmanii* and *S. lactamdurans* (*Amycolatopsis spp*) have been found to require dithiothreitol and oxygen and were stimulated by ferrous ions and ascorbate. Isopenicillin N is converted to penicillin N in cephalosporin and cephamycin-producing microorganisms, but not in penicillin producers, by an epimerase that isomerizes the L-α-aminoadipyl side chain to the D configuration. Penicillin N is transformed into

deacetoxycephalosporin C by the deacetoxycephalosporin C synthetase which is referred to as ring-expanding enzyme or expandase. Expandase converts the five-membered thiazolidine ring of penicillins into the six-membered dihydrothiazine ring of cephalosporins and requires ferrous ions and α-ketoglutarate (Kovacevic and Miller 1991; Roach *et al.* 1995; Sim and Tan 1995). This enzyme expands the β-lactam ring of penicillin N but does not accept isomers of isopenicillin N such as penicillin G or 6-aminopenicillanic acid as substrates (Liu *et al.* 1994).

The β-lactam-thiazolidine ring nucleus is common to all penicillins whereas the β-lactam-dihydothiazine ring is common to cephalosporins and cephamycins and both these rings are formed by common pathway involving the condensation of L-cysteine and D-valine (Cortes *et al.* 1986; Demain 1990; Kovacevic and Miller 1991; Roach *et al.* 1995; Sim and Tan 1995; Loke *et al.* 2000). A third amino acid α-aminoadipic acid (α-AAA) is involved in the formation of a linear cysteine-containing tripeptide δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine (LLD-ACV) which is an intermediate in the biosynthesis of all penicillins and cephalosporins (Martin and Liras 1985; Martin and Gutierrez 1995; Brakhage 1998; Loke and Sim 1998; Schugerl and Seidal 1998) together with being a precursor of the β-lactam antibiotics (Castro *et al.* 1988; Baldwin *et al.* 1997). ACV is then cyclized to form isopenicillin N, an intermediate which has an L-α-aminoadipyl side chain attached to the nucleus of penicillin (Williamson *et al.* 1985; Ramos *et al.* 1985; Leskiw *et al.* 1988; Madduri *et al.* 1991). The reactions required for the formation of this tripeptide, are catalyzed by a single multi functional enzyme designated ACV synthase (ACVS). This enzyme is encoded by a single structural gene namely *pcbAB* (Durairaj *et al.* 1992, 1996; Perez-Llarena *et al.* 1997, 1998; Alexander and Jensen 1998; Martin 1998; Penalva *et al.* 1998).

Deacetoxycephalosporin C is hydroxylated to deacetylcephalosporin C by another α-ketoglutarate dependent dioxygenase. This enzyme catalyzes the incorporation of oxygen atom from O_2 into deacetoxycephalosporin C, and is also stimulated by α-ketoglutarate, ascorbate, dithiothreitol and Fe^{2+} (Martin and Gutierrez 1995). Deacetoxycephalosporin C synthase and Deacetoxycephalosporin C hydroxylase are separate enzymes in cephamycin producing actinomycetes (Jensen *et al.* 1986; Cortes *et al.* 1986). Further reactions are involved in the attachment of the C-7 methoxy group and the C-3` carbomoyl group during cephamycin biosynthesis in the actinomycetes (Jensen *et al.* 1986; Martin and Liras 1985; Garcia-Dominguez *et al.* 1991). All available evidence on β-lactam biosynthesis suggests that the β-lactam antibiotic specific genes and the enzymes they encode are related in all producers. However, clavulanic acid synthase has no detectable similarity at the nucleotide or protein level to IPNS and other synthases have not been cloned (Kirby pers. comm 2000).

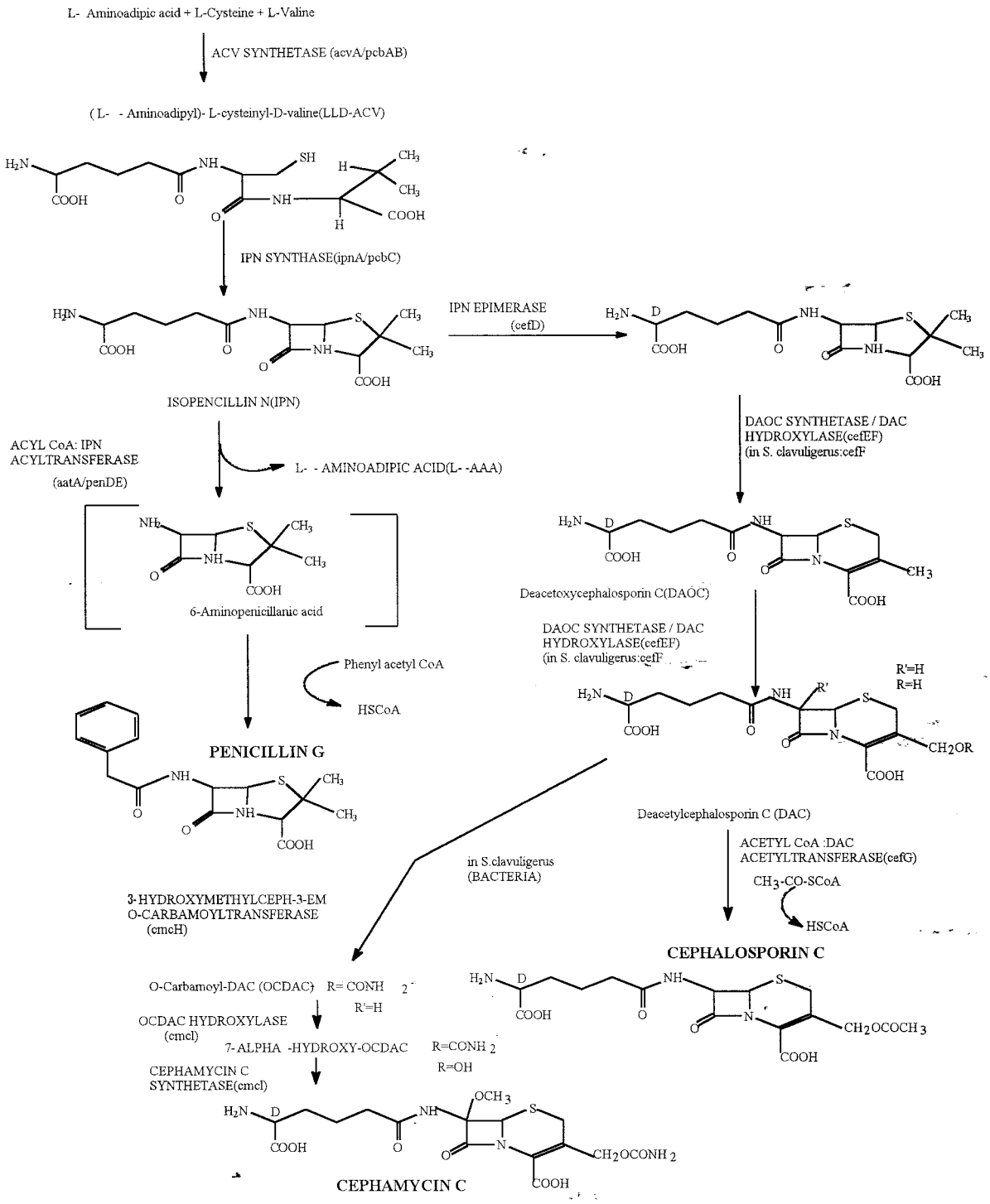


Figure 1.2. Biosynthetic pathway of β-lactam antibiotics including the enzyme and gene designation.

1.4.3 Common genes to all penicillin and cephalosporin producers (*pcb*)

Genes encoding enzymes involved in penicillin and cephalosporin (or cephamycin) biosynthesis are designated *pcb*. The genes *pcbAB* and *pcbC* are present in all prokaryotic and eukaryotic penicillin and cephalosporin producers (Samson *et al.* 1985; Kovacevic *et al.* 1989; Skatrud 1992; Yu *et al.* 1994; Loke and Sim 1998; Leitao *et al.* 1999). Synthesis of the tripeptide is carried out by a single multienzyme ACV synthetase (ACVS). ACVS catalyzes the ATP dependent activation of each of the L-amino acids, the binding of the activated amino acids as thioesters, the epimerization of L-valine to the D-configuration, and the polymerization of the amino acids to make ACV (Van Liempt *et al.* 1989). The gene encoding the ACVS was designated *pcbAB* because two genetic loci were supposed to be responsible for synthesis of the α-aminoadipyl-cysteine (AC) dipeptide and the ACV tripeptide.

However genetic evidence indicates that a single gene encodes all the activities needed to form ACV (Diez *et al.* 1997). A conspicuous feature of the ACVS protein sequence is the presence of three repeated regions or domains, each containing more than 500 amino acids. A multienzyme thiotemplate mechanism has been proposed for the ACVS reaction (Gutierrez *et al.* 1991; Coque *et al.* 1991; Aharonowitz *et al.* 1993; Diez *et al.* 1997). All domains of the bacterial and fungal ACVSs contain a phosphopanteteine binding sequence that is similar to the phosphopanteteine site of the polyketide and fatty acid synthetases (Martin and Gutierrez 1995). The *S. lactamdurans* and *S. clavuligerus pcbAB* and *pcbC* genes are closely linked and are transcribed in the same orientation in contrast to the divergent expression in the fungal β-lactam clusters (Coque *et al.* 1991; Leitao *et al.* 1999). In an early step of the pathway α-aminoadipic acid which is a specific precursor of β-

lactam antibiotics and is formed in actinomycetes by deamination of lysine by the enzyme Lysine 6-aminotransferase followed by condensation of the three amino acid precursors, viz L-α-aminoadipate, L-cysteine and L-valine which are L-isomers of the constituent amino acid to form δ-(L-α-aminoadipyl-L-cysteinyl-D-valine) (ACV), which is a common precursor of penicillin and cephalosporin is a linear tripeptide which is catalyzed by ACV synthetase a multifunctional enzyme, followed by subsequent conversion to isopenicillin N.

1.4.4 Isopencillin N synthase (*pcbC*) gene

The second step in the biosynthetic route to penicillins and cephalosporins is cyclization of the linear tripeptide ACV to isopenicillin N (IPN) which shows weak antibiotic activity. In this reaction by IPNS, four hydrogen atoms are removed with the subsequent reduction of one molecule of oxygen into H_2O (Tan and Sim 1996).

The IPNS reaction requires the presence of iron and ascorbate, indicating that ferrous ions are essential for its enzymatic activity (Ramos *et al.* 1985; Kriauciunas *et al.* 1991; Martin and Gutierrez 1995). A critical role for the iron indicate that it binds directly to ACV through its cysteinyl thiol (Barredo and Martin 1991; Petrich *et al.* 1992; Burzalaff *et al.* 1999). This reaction is carried out by the isopenicillin N synthase (IPNS) (Cohen *et al.* 1990; Burzalaff *et al.* 1999; Loke *et al.* 2000) (frequently referred to as cyclase) which is encoded by the gene pcbC. This enzyme is of particular interest because of its unusual mechanism of action, which results in the stepwise formation of the β-lactam and thiazolidine fused rings (Wolfe and Jokinen 1979; Castro *et al.* 1985; Jensen *et al.* 1986; Shiffman *et al.* 1988; Van Liempt *et al.* 1989; Coque *et al.* 1991; Madduri *et al.* 1991; Orford

et al. 1991; Brakhage *et al.* 1992; Aharonowitz *et al.* 1993; Cohen *et al.* 1993; Townsend 1993; Ward and Hodgson 1993; Yu *et al.* 1994; Valegard *et al.* 1998; Leitao *et al.* 1999), and because its broad substrate specificity has led to novel penicillins from ACV analogs (Wolfe *et al.* 1984; Castro *et al.* 1988).

The nature of the next enzymatic step of penicillin biosynthesis and the role of 6-APA as an intermediate in penicillin biosynthesis has remained obscure for many years (Martin and Liras 1985). A two step model for the conversion of IPN to benzylpenicillin require the involvement of two enzyme activities, IPN amidohydrolyase (6-APA forming), which would cleave IPN to 6-APA and α-aminoadipic acid, and acyl-Coa:6-APA acyltransferase (AAT). All these enzymes are encoded by the single *penDE* gene (Brakhage 1998).

1.5 Biosynthetic production of new antibiotics

In the current climate of increasing antibiotic resistance, there is a growing need for novel effective antimicrobials (McGowan *et al.* 1996). An approach used for the formation of antibiotics other than those that occur naturally is to divert the biosynthetic machinery of antibiotic producing organisms (Butler *et al.* 1996). The basis for the approach is that many enzymes do not have the substrate specificity, but can handle a variety of compounds resembling their normal substrate in the same way as the original. The earliest demonstration of this approach involved the feeding of false precursors (i.e analogs of normal precursors) to organisms which then produced analogs to their normal metabolites, (Floss 1987; Skatrud 1992; Pertursson and Baldwin 1998).

A number of refinement methods have been developed over the years, but they were all biochemical and relied on the biosynthetic machinery of a particular existing organism. These methods have a drawback from the manufacturing point of view in that they are potentially difficult, time consuming and very costly. However entirely new avenues for exploitation were opened up by the emergence of recombinant DNA technology that set the stage for more precise genetic manipulation of organisms. This technique has also been continuously revised and improved leading to an understanding of the genetic and biochemical basis for the biosynthesis of antibiotics (Skatrud 1992; Diez *et al.* 1997).

For many years genetic manipulation of micro-organisms was limited to improvement programs based on random mutations and selection, and even today these techniques are indispensable tools for the development of complex processes with a background of molecular knowledge (Bellgardt 1998). Random mutations and screening are an essential part of conventional antibiotic production resulting in strain improvement (Diez *et al.* 1997; Salas and Mendez 1998).

1.6 DNA SHUFFLING

DNA shuffling or sexual PCR is a practical powerful *in vitro* process for directed molecular evolution that uses recombination to accelerate the rate at which one can evolve genes. It is also used to recombine homologous DNA sequences during *in vitro* molecular evolution allowing many mutations for improved phenotypes to be evolved rapidly (Stemmer 1994; Zhao and Arnold 1997; Crameri *et al.* 1996, 1997 & 1998; Christians *et al.* 1999).

The technology was developed by Stemmer and has been enhanced in the past three years extending its range of applications to small molecule pharmaceuticals, pharmaceutical proteins, gene therapy vehicles, vaccines and evolved viruses for vaccines (Stemmer 1994; Hansson *et al.* 1999; Sun 1999).

Nature changes DNA by a variety of mechanisms that involve mutation, selection and recombination to evolve highly adapted individuals from the effectively infinite possibilities encoded implicitly in the genome; this is copied by DNA shuffling which mimics natural recombination by allowing the *in vitro* homologous recombination of DNA. It also allows introduction of point mutations at a very low controlled rate (Zhao and Arnold 1997) thus combining recombination, point mutation and selection techniques to create a general, parallel algorithm for evolving improved genes (Moore *et al.* 1997; Patten *et al.* 1997; Zhang *et al.* 1997).

DNA shuffling combined with focused selection pressure in the laboratory will allow one to rapidly evolve genes for a wide variety of industrial applications (Patten *et al.* 1997). It will play a central role in the development of applied molecular evolution technologies thus becoming indispensable

for bringing existing biological diversity into the service of human health care, agricultural, industrial and chemical needs.

1.6.1 PRINCIPLE OF DNA SHUFFLING

The principle of DNA shuffling involves use of closely related molecules with point mutations prepared either through error-prone PCR or other mutation techniques such as oligonucleotide-directed mutagenesis (Harayama 1998).

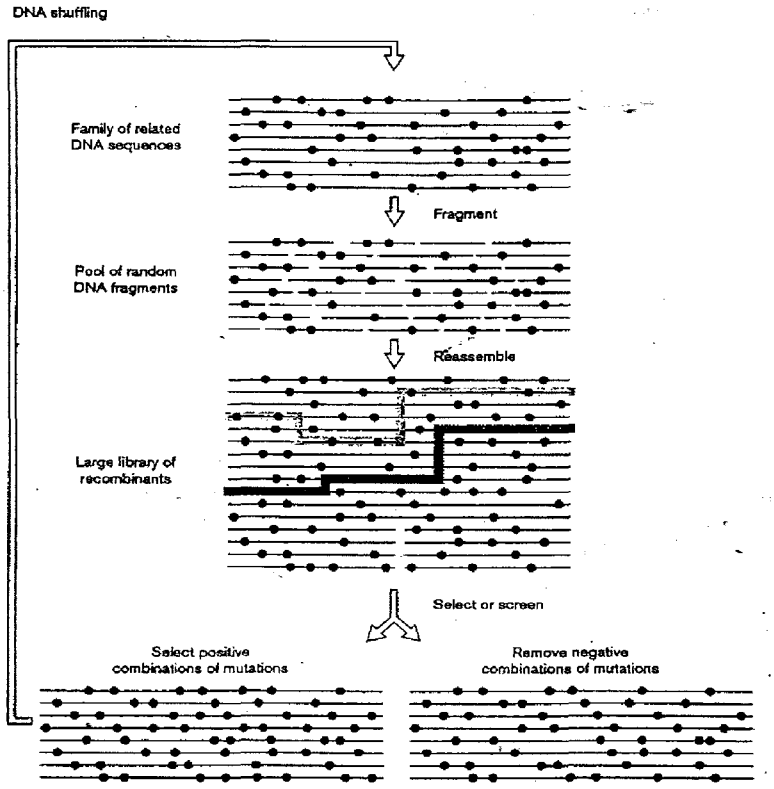


Figure 1.3. Outline of the steps involved in DNA shuffling (Stemmer 1994).

The first step in DNA shuffling involves breaking the molecule into random fragments using DNase I which can randomly create nicks along each strand of a DNA molecule. The fragments can then go through PCR without added primers, in three steps in the PCR process. The first step is denaturing by increasing the temperature so that double-stranded fragments are separated completely into single-stranded ones. The second step involves lowering the temperature so that single-stranded fragments anneal to other fragments overlapping by a certain number of bases that are complementary at the overlapping region. After the annealing step the homologous templates should prime each other to form 5' and 3' overhangs. The third step is polymerase extension involving increase in temperature to the level that allows DNA polymerase to extend the 5' overhangs using the other annealed strand as template. The 3' overhangs are not changed as DNA polymerase can only extend from a 5' end. The three steps of denaturation, annealing and polymerase extension are repeated for multiple cycles. The average fragment length is increased in each PCR cycle. After a number of cycles of PCR without primers molecules of the original size will be obtained. When a template from one molecule primes a fragment from another molecule, recombination occurs. The idea behind DNA shuffling is that by recombining beneficial mutations from different molecules, the product with increased function will be obtained (Sun 1999). The addition at some stage of specific primers allows the large scale amplification of the shuffled product.

As an applied molecular evolution, DNA shuffling will play a key role in biomolecular engineering. In contrast to the point mutation techniques, DNA shuffling exchanges large functional domains of sequences to search for the best candidate molecule, thus mimicking and accelerating the process of sexual recombination in the evolution of life.

This system has so far been employed in biotechnology for the screening of combinatorial antibody libraries, but its application will be extended further into the science of protein-receptor or protein-ligand interaction. The bioinformatics for genome and protein analyses will contribute greatly towards accelerated advances in pharmaceutical industry.

Biomolecular engineering will no doubt become one of the most important scientific disciplines, because it will enable systematic and comprehensive analyses of gene expression patterns in both normal and diseased cells, as well as the discovery of many new high-value molecules. New metabolites, including novel antibiotics that are active against resistance strains, will also be produced soon by recombinant organisms having *de novo* engineered biosynthetic pathway enzyme systems. The biomolecular engineering era is here, and many benefits will be derived from this field of scientific research for years to come if we are willing to put it to good use (Ryu and Nam 2000).

PURPOSE OF THE STUDY

The purpose of the study was to amplify the IPNS gene from the eight mesophilic β-lactam antibiotic producing *Streptomyces* and two novel thermophilic strains using PCR. Sequencing of the unknown IPNS genes from *S. hygroscopicus*, *S. tanashiensis* and the two novel thermophilic strains will then be carried out in order to enable documentation of this gene sequence in The World IPNS Gene Bank. The IPNS genes were to be used for gene shuffling, cloning and expression of the gene in an *E. coli* system to detect changes, increase/decrease as a result of gene shuffling followed by subsequent molecular analysis. .

CHAPTER 2

GENERAL MATERIALS AND METHODS

2.1 INTRODUCTION

Streptomyces is a genus represented by a large number of species and varieties having a GC base ratio of between 60% and 73% (Kieser *et al.* 1992). Growth in *Streptomyces* occurs at the tips of filaments and is often accompanied by branching so that the vegetative phase consists of a complex, tightly woven matrix, resulting in a compact convoluted colony. As the colony ages, characteristic aerial filaments are formed, which project above the surface of the colony and give rise to spores (Walter *et al.* 1998). *Streptomyces* are strict aerobes, whose growth in liquid culture is usually markedly stimulated by forced aeration. Sporulation usually takes place not in liquid culture but only when the organism is growing on the surface of agar or another solid substrate; it can also occur when an organism forms a pellicle on a surface of an unshaken liquid culture or in a very few strains under the liquid (Shima *et al.* 1996; Aceti and Champness 1998; Kelemen and Buttner 1998).

Nutritionally, the *Streptomyces* are quite versatile. Growth-factor requirements are rare, and a wide variety of carbon sources, such as sugars, alcohols, organic acids, amino acids and some aromatic compounds can be utilized. Most isolates produce extracellular hydrolytic enzymes that permit utilization of polysaccharides (starch, cellulose, hemicellulose), proteins and fats, and some strains can use hydrocarbons, lignin, tannin or even rubber (Hao and Kendrick 1998).

The most striking property of the *Streptomyces* is the extent to which they produce antibiotics (Kieser *et al.* 1992; Ludovice *et al.* 1992). Antibiotic production is often seen on the agar plates

used in the initial isolation of *Streptomyces*, adjacent colonies of other bacteria show zones of inhibition (Shima *et al.* 1996; Aceti and Champness 1998; Kirby Pers. Comm. 1998).

2.2 MAINTENANCE OF β-LACTAM PRODUCING STREPTOMYCES USING DIFFERENT MEDIA

The following strains from culture collections were used in this study: *Streptomyces clavuligerus* (DSMZ 40751); *Streptomyces jumonjinensis* (DSMZ 747); *Streptomyces griseus* (DSMZ 40759); *Streptomyces lipmanii* (DSMZ 40752); *Amycolatopsis spp* (DSMZ 40930) previously known as *Streptomyces lactamdurans* after being named *Norcadia lactamdurans*: The other strains were viable cultures obtained from microbiology culture collection (Rhodes), and these were *Streptomyces cattleya*, *Streptomyces tanashiensis* and *Streptomyces hygroscopicus* (ATCC 31039) and two unknown thermophiles named in this study as Thermophilic A and B.

The DMSZ strains and the two thermophilic strains were routinely maintained on *Streptomyces* media M65 agar (Appendix A) whereas the Rhodes culture collection were grown on ISP2 agar and subcultured monthly using a four-way streak method. The petri dishes with organisms fully grown were sealed with parafilm and stored at 4^0C. These cultures had been incubated at 28^0C for 5 days for all the others and 7 days for *Streptomyces hygroscopicus* prior to being kept at 4^0C whereas the thermophiles were incubated at 55^0C prior to storage.

The other media used was sporulation media (Appendix A) which was found to result in confluent growth of the above strains within 2-3 days incubation at 28^0C with the thermophiles growing confluently in 24 hrs at 55^0C. This media was used for providing cells for various tests and not for maintenance of cultures. M3 media was used to prepare agar slants which were inoculated with spores from the strains and incubated at room temperature for a week.

2.2.1 Culture preservation

All plate cultures were found to retain viability beyond 4-6 months even though they were re-streaked every month, to ensure viability. The slants were found to retain their viability throughout the study period, being kept at room temperature. The strains that were stored at -20^{0}C in 20% v/v glycerol as spore suspensions were found to be viable throughout the 2 year study period.

2.2.2 Total DNA Extraction

The *Streptomyces* used in this study were grown in M3 broth from a spore inoculum (Appendix A) except *S. lipmanii* which was grown in M65 broth. The broth cultures were grown in 250ml conical flasks with a medium volume of 50ml. They were incubated at 28^{0}C in an orbital shaker at a speed of 200 rpm. Sterile glass beads were used in flasks to prevent the mycelium growing as a large compacted pellet that would make DNA extraction very difficult. The cultures were incubated for 3 - 5 days.

Numerous attempts were made to extract total genomic DNA using methods that had been adapted for *Streptomyces* (Krallis MSc Thesis) without any success. The DNA yield was very low with these mini-prep methods. A number of parameters were investigated which were thought to have contributed to the low poor quality DNA obtained without success. A method that eventually proved very successful was adapted from the basic mini-prep protocol for bacterial genomic DNA and was carried out as follows.

1.5ml broth culture was placed in a top bench microcentrifuge and spun for 2 minutes resulting in compact pellet and the supernatant being discarded. The pellet was resuspended in 570µl TE buffer (Appendix B) by repeated careful pipetting. 30µl of 10% SDS, 3µl (20mg/ml Proteinase K) and 10µl of RNase (4mg/ml) were added giving a final concentration of 100µg/ml proteinase K in 0.5% SDS. At this point the solution became viscous due to the detergent (SDS) lysing the bacterial cell wall. The solution was mixed by inverting the eppendorf tube three times and incubating the solution for 1 hr in a water-bath set at 37^0C. The 1 hr incubation was followed by addition of 100µl of 5M NaCl and 80µl CTAB/NaCl solution (Appendix B). The solution was mixed thoroughly by inverting the tube several times before incubating at 65^0C for 10 min. This step was very important because a CTAB-nucleic precipitate forms if the salt concentration drops below 0.5M at room temperature. The aim being to remove the cell wall debris, denatured protein and polysaccharides complexed to CTAB while retaining the nucleic acids in solution.

An approximately equal volume 800µl of chloroform/isoamyl alcohol was added and the solution was spun for 3 min in a microcentrifuge. This step removes CTAB-protein /polysaccharide complexes, after centrifuging, a white interphase became visible. A widened tip was used to remove aqueous, viscous supernatant to a clean sterile eppendorf tube leaving the interphase behind. An equal volume of phenol/chloroform/isoamyl alcohol was added and the solution was spun for 3 min in a microcentrifuge.

The supernatant was carefully transferred to a sterile eppendorf where 0.6 volume isopropanol was added to precipitate the nucleic acids. This was achieved by shaking the tube back and forth gently a few times. The stringy white DNA was precipitated by a short spin in a microfuge. The supernatant was discarded and the remaining pellet was washed with 70% ethanol to remove the residual CTAB by spinning in a microfuge for 5 min at 10,000 rpm. The supernatant was

discarded and the pellet was air-dried under the fume hood for 3-5 hrs. The pellet obtained was resuspended in 100µl TE buffer. Aliquots of the DNA obtained were made to enable storing some of the DNA at -20°C and 20µl aliquots that were to be used immediately were stored at 4°C.

Total DNA extractions using method 2 adapted from miniprep protocol of bacterial genomic DNA yielded about 10 fold more DNA than the protocol that had been optimized for *Streptomyces*, and was already in use in our laboratory. The other problem/factor noted was the high viscosity which hindered further manipulations. Some strains such as *S. tanashiensis*, *S. clavuligerus*, *S. griseus* and the two Thermophilic strains A and B formed compact mycelial pellet in broth media making attempts at DNA extraction procedures unsuccessful. Therefore it was imperative that formation of the mycelial pellet be avoided by using sterile glass beads and shaking cultures on an orbital shaker at higher speeds than normal. The speed that was found to be optimal was 220 rpm.

2.2.3 Quantification of DNA

Spectrophotometric quantitation was carried out using a Unicam spectrophotometer by diluting DNA to 10^{-3} in TE buffer. Spectrophotometric readings (absorbances) were converted to concentrations using the equation below.

1OD = 50µg/ml

Spectrophotometric measurements should generally lie between 0.1 and 1.0 to ensure reliability.

2.2.4 Detection and Analysis of DNA

The simplest way to check for the presence of DNA, its quality and possibly its concentration is to load a sample along with the appropriate molecular weight marker, onto an agarose gel which contains 0.8% ethidium bromide. Agarose gel analysis enables quick and easy quantitation of DNA, especially for small fragments such as PCR products or small amounts of DNA which are difficult to quantify spectrophotometrically. DNA bands on the gel can be visualized, under ultraviolet trans-illumination. By comparing DNA bands to be analyzed with those of the appropriate molecular marker with known amounts of DNA one is able to identify the expected DNA band by comparison of the band intensity with the standards either visually or using a scanner.

2.2.4.1 Analysis of DNA

Agarose gel analysis was found to be a quick way of quantifying DNA especially one that was extracted using method 1 since this method gave minute amounts of DNA which could not be detected spectrophotometrically especially when they had been RNase treated since the resultant DNA was free of RNA whose presence results in elevated spectrophotometric readings due to the machine's inability to differentiate between the various types of nucleic acids.

2.2.5 Bioassays

The strains used were tested for antibiotic production seen on agar plates as zones of inhibition observed around the test bacteria. Two methods were used *viz* Agar plug method and Cup plate method which were carried out as follows.

2.2.5.1 Agar plug method

This was carried out by cutting out about 10mm diameter using a sterile pipette tip from a plate containing *Streptomyces spp* grown on sporulation agar. The plugs were placed on M3 plates that had been spread with 100µl of a 10^{-1} dilution of an overnight test organism such as *E. coli* or *M. luteus*. Plates were incubated at 37^0C and zones of inhibition were measured and taken to indicate potential antimicrobial activity.

2.2.5.2 Cup plate method

This method was carried out by cutting out wells of about 10mm diameter from M3 agar plates that had been seeded with *E. coli*. The wells were filled with *Streptomyces* broth cultures and were incubated as above. The test organisms used were prepared by diluting overnight culture to an optical density equivalent to an absorbance of 0.3 at 600nm. 1ml of the dilute suspension was used per 100ml of assay agar.

2.2.6 Protein Assay

Protein concentrations were performed on the cell-free extracts by the method of Bradford assay using bovine serum albumin (Sigma) as a reference standard and Coomasie brilliant blue (G-250; Fluka chemica, Switzerland) was used as a binding dye. The reaction mixtures were then quantified spectrophotometrically using Shimadzu 160-A (Japan) spectrophotometer and the absorbance was read at 595nm.

2.3 RESULTS

Total DNA was successfully extracted using a method that was adapted for *Streptomyces spp* from the mini-prep protocol of bacterial genomic DNA and visualized on an agarose gel (Figure 2.1) and recorded using a Kodak Digital Science camera and software.

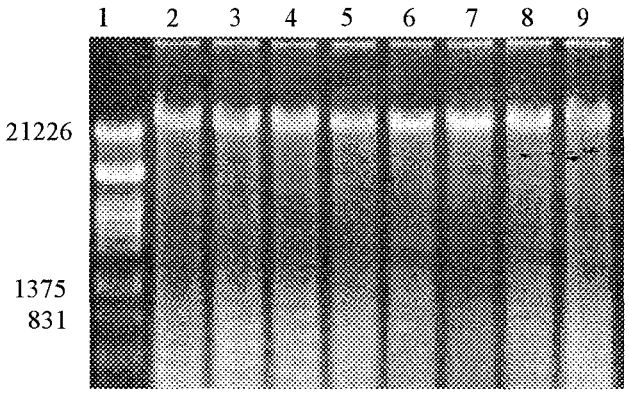


Figure 2.1. Agarose gel electrophoresis of total DNA obtained using method adapted from miniprep of bacterial genomic DNA. Lane 1 Molecular weight marker III, Lanes 2-8 Total DNA from: *Amycolatopsis* spp, *S. lipmanii S. cattleya, S. clavuligerus, S. griseus, S. hygroscopicus, S. jumonjinensis* and *S. tanashiensis* respectively.

The adapted method was compared to the method in use in our laboratory for extraction of total DNA from *Streptomyces spp*. The method in use was unsuccessful, despite numerous attempts that were made to optimize this method. The main constraint of this method was the low yield of DNA compared to method 2 which showed almost ten times more yield (Figure2.2) as well as sufficient purity to allow further molecular analysis.

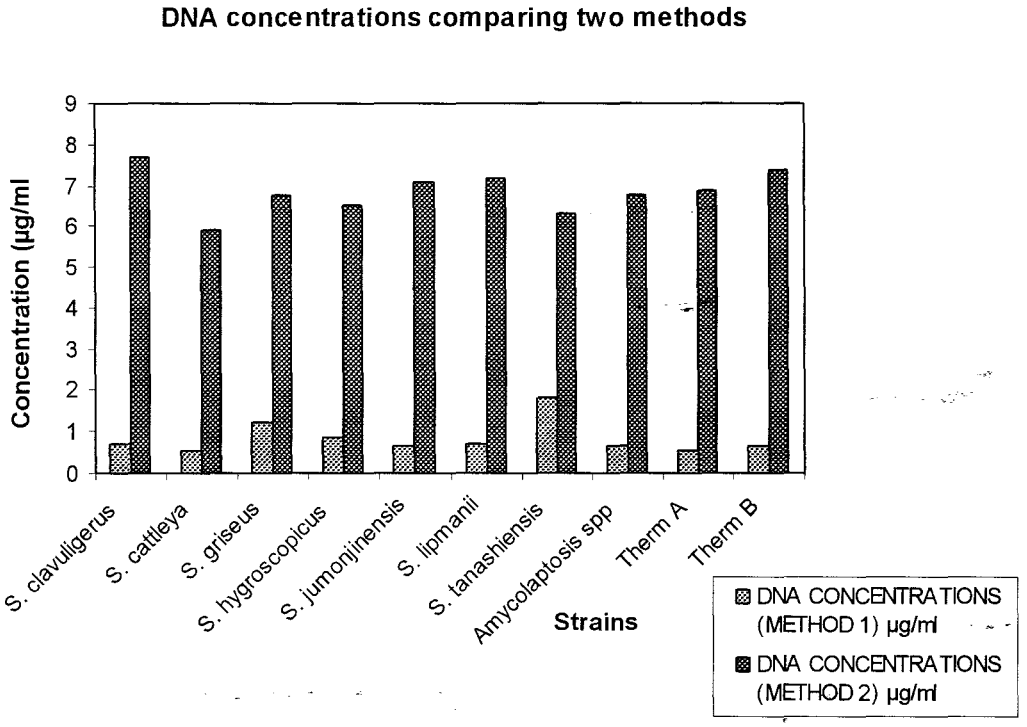


Figure 2.2. This graph shows the differences in yields of DNA obtained using method 1 and 2.

Antibiotic assays of the *Streptomyces spp* against the test organisms *M. luteus*, *S. aureus* and *E. coli* showed differences in the degree of antibiotic activity amongst the strains (Table 2.1) and the two novel thermophiles, Thermophilic A and Thermophilic B showing no antibiotic activity since they were unable to inhibit growth of the test strains.

Antibacterial activity against selected indicator organisms.

Scale: (-) = No inhibition to (+3) = Strong inhibition (Zone size: 15-25 mm)

Inhibition scored as (±) indicates an indistinct halo observed

Table 2.1. The observed antibiotic activity by the *Streptomyces* strains used.

	Degree of inhibition									
Indicator organisms	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
M. luteus	3	2	3	3	3	2	2	3	±	±
S. aureus	1	-	1	1	-	-	-	1	-	-
E. coli	2	1	1	1	1	1	1	2	-	-

1 - *S. clavuligerus*

2 - *S. cattleya*

3 - *S. hygroscopicus*

4 - *S. tanashiensis*

5 - *S. griseus*

6 - *S. lipmanii*

7 - *S. jumonjinensis*

8 - *Amycolatopsis* spp

9 - Thermophilic A

10- Thermophilic B

Cell free extracts were assayed for protein concentrations. Protein determination were carried out on cell free extracts according to the method of Bradford and the concentrations were found to be different amongst the strains used and they ranged from 3.2 to 11.7 µg/ml (Figure 2.3) with *S. lipmanii* showing the highest and Thermophilic A showing the least amount of protein.

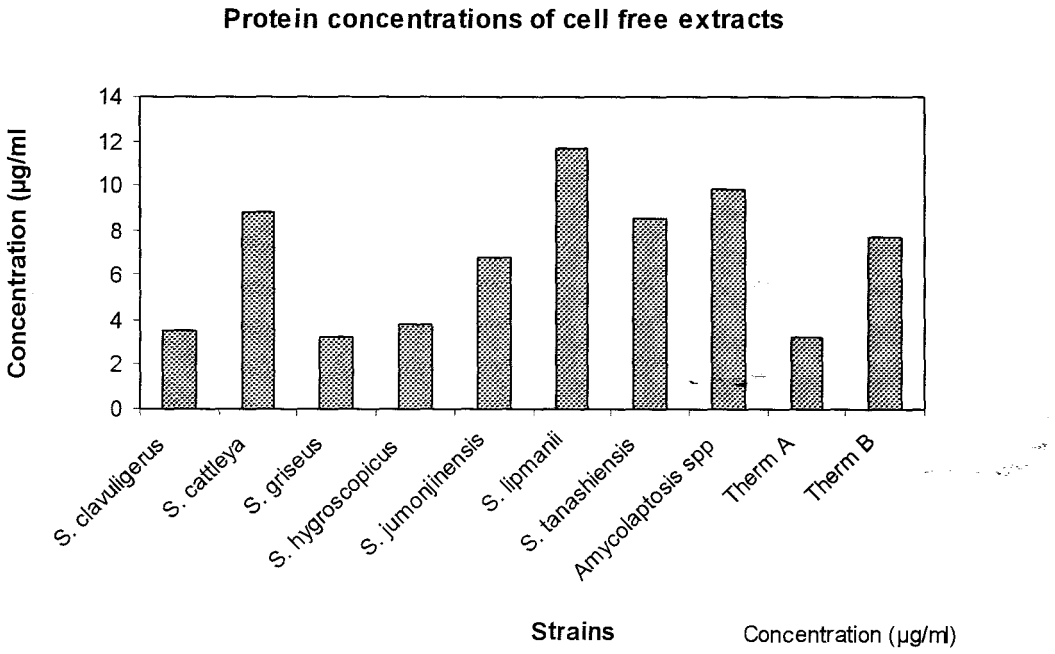


Figure 2.3. The above graph show concentrations of protein obtained by Bradford assay.

2.5 DISCUSSION

Preservation of cultures on plates for all the strains used was a practical and reliable method of maintaining the cultures, however, *S. lipmanii* seemed unable to withstand the monthly subculturing that was done, instead it could only be subcultured once, thereafter it was never viable, thus, new cultures were regularly obtained from the frozen glycerol stocks. The other finding pertaining to *S. lipmanii* was that, for maintenance of the culture, we could only use M65 and not ISP2 media. This suggests that $CaCO_3$ was essential for it to remain viable since that was the ingredient that was absent in ISP2 or M3 media. *S lipmanii* is thought to be sensitive to even small changes in pH. Although frequent subculturing might result in loss of desirable genetic functions, this was not the case with the IPNS gene which was the main focus of our study. The success of maintaining the cultures could have been attributed to the fact that sporulation media was used whenever antibiotic bioassays were carried out since this media was found to enable sporulation within 2 to 3 days as opposed to the 5 to 7 days required for optimal growth in *Streptomyces* since sporulation is linked to antibiotic production as a secondary metabolite. Storage of cultures at 4^0C was adequate for facilitating maintenance of the cultures without any detectable loss of the properties of interest for the 2 years that the strains were used.

Conventional procedures for isolation of *Streptomyces* DNA were not successful despite countless attempts using well-established protocols for *Streptomyces* (www.cbs.umn.edu/asirc) that provided the *Streptomyces* laboratory manual called 'Practical *Streptomyces* Genetics' (Hopwood *et al.* 1996). The quality of DNA obtained was very poor for most of the strains and yields very low except for *S. cattleya* which yielded good quality DNA though the yield was low. The success with the modified protocol from preparation of genomic DNA from bacteria resulted in high yields of good quality DNA (Figure 2.1. & Figure 2.2.) probably also because all the

technical procedures that are required when working with *Streptomyces* were incorporated into the revised method because of the sensitivity and nature of the DNA from *Streptomyces* and including the use of a widened pipette tip and other procedures to avoid shearing the DNA. The other factor that favored this method was that small volumes of 1.5ml were required which enabled extractions from all the strains simultaneously. The spectrophotometric DNA concentration estimates tend to give artificially high concentrations due to contaminating RNA and proteins when RNase treatment was not carried out but this way of estimating the concentration is preferred due to its speed and ease.

The use of test organisms against which the *Streptomyces spp* antibiotic producing strains were challenged does not really give a true picture of the kind of antibiotics being produced because the inhibition of growth and zones of clearing confirm antimicrobial activity without indicating the whole spectrum or array of antibiotics that the *Streptomyces* are able to produce and also the number of test organisms were too few to have a meaningful picture of the kind of antibiotics produced. However, Table 2.1 shows that *S. tanashiensis* produces antibiotics to a significant level while Thermophilic A and B show a low level of production possibly due to the IPNS gene being either silent, defective or amounts produced are in an inactive form.

CHAPTER 3

DETECTION OF IPNS GENE PRODUCTS

3.1 INTRODUCTION

Polymerase chain reaction (PCR) is an elegant common *in vitro* technique which is used to enzymatically amplify the number of copies of a specific region of DNA, in order to produce enough DNA for subsequent molecular protocols. This technique can be used to identify with a very high probability, disease-causing viruses and/or bacteria which previously was not possible due to insufficient nucleic acid. PCR has rapidly become one of the most widely used techniques in molecular biology and for good reason because it is rapid, inexpensive and simple means of producing relatively large numbers of copies of DNA molecules from minute quantities of source DNA material even when the source (template) DNA is of relatively poor quality. PCR is very versatile and so pervasive in molecular biology that it is difficult to think of life without it. Many types of samples can be analyzed for nucleic acids. Most PCR used DNA as a target rather than RNA, because of the stability of the DNA molecule and the ease with which DNA can be isolated. The essential criteria for any DNA sample are that it contain at least one intact DNA strand encompassing the region to be amplified and that any impurities are sufficiently diluted so as not to inhibit the polymerization step of the PCR reaction (Roche Molecular Biochemicals 1999; Veilleux 1999).

PCR reaction uses two oligonucleotide primers that hybridize to opposite strands and flank the target DNA sequence that is to be amplified. The elongation of the primers is catalyzed by a heat-stable DNA polymerase such as Taq DNA polymerase. A repetitive series of cycles involving template denaturation, primer annealing, and extension of the annealed primers by the

polymerase results in exponential accumulation of a specific DNA fragment (Roche Molecular Biochemicals 1999).

The ends of the fragment are defined by the 5' ends of the two primers. Primer extension products synthesized in a given cycle can serve as a template in the next cycle, the number of target DNA copies approximately doubles every cycle, thus 20 cycles of PCR yield about a million copies (2^{20}) of the target DNA.

The application of PCR is continually being updated increasing the usefulness and scope of the technique. Examples include use of primers containing sequences that were not completely complementary to the template turning PCR into a tool for *in vitro* mutagenesis. Another innovative approach is the use of 'Hot Start' method which minimizes the formation of primer-dimers during PCR (Veilleux 1999).

3.1.1 Primer Design

A primer is a short segment of nucleotides which is complementary to a section of the DNA. Primers are usually specific and should be between 18-24 bases in length. Primer design should take into account such aspects as primer length, Tm and complements of the primers to each other. For example, if the Tm difference is high, the lower Tm can be increased by increasing the length of that primer at either the 3' or the 5' end which also results in a constant size of the amplified locus (Veilleux 1999).

Primer pairs should be tested for primer-primer interaction using various softwares available such

as 'Cprimer' available at ftp.bio.indiana.edu. The other point to take into account is adjustment of cycling conditions and buffer concentrations for each primer pair so that amplification of the desired locus is specific, with no secondary products.

3.1.2 PCR Programme

The requirement of an optimal reaction is to amplify a specific locus without any unspecific by-products. Therefore annealing needs to take place at a sufficiently high temperature to allow only the perfect DNA-DNA matches to occur in any reaction. For a given primer pair, the PCR programme can be selected based on the GC content of the primers and the length of the expected PCR product. In most cases template DNA to be amplified is relatively small ranging from 0.1 to 2.5 Kb. The activity of the Taq DNA polymerase is about 2000 nucleotides/min at optimal temperature (72^0C - 78^0C) and the reaction can be calculated accordingly (Veilleux 1999).

3.1.3 Cloning

Gene cloning is one of the most powerful techniques of modern biotechnology, which results in production of protein of interest in high yield. There were three options available to researchers before the advent of genetic engineering. They included screening of microbes from soil and water samples to find a new high yielding isolate, medium and process parameters optimization to increase yield in known cultures and finally random mutagenesis and selection of the new overproducing mutants. The tools of genetic engineering allows a researcher to clone a gene of an already known biocatalyst and over-express it or modify it for the desired characteristics by site-directed mutagenesis after studying its primary and possible three-dimensional structure.

3.1.4 Selection of Positive Transformed Bacteria

Very often blue/white screening is used for identification of transformed bacteria that contain recombinant plasmids. This approach allows color discrimination of non-recombinants from recombinants.

3.1.5 pGEM-T Cloning

The pGEM-T easy vector system is a convenient system for the cloning of the PCR products. The vector is prepared by cutting Promega's pGEM ^{R}-T easy vector with EcoRV and adding a 3` terminal thymidine to both ends, thus greatly improving the efficiency of ligation of a PCR product into the plasmid by preventing recircularization of the vector and providing a compatible overhang for PCR products. pGEM^{R}-T easy vector has a high copy number that contain T7 and SP6 RNA polymerase promoters flanking a multiple cloning site (MCS) within the α-peptide coding region of the enzyme β-galactosidase. There are multiple restriction sites within the MCS which allow for the release of the insert by digestion with a single restriction enzyme (www.promega.com).

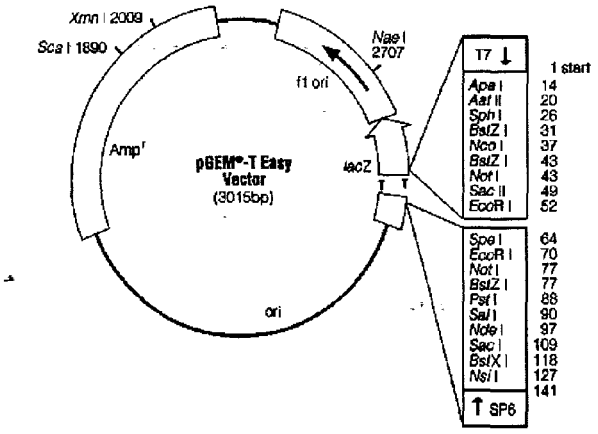


Figure 3.1. Circular vector map of p-GEM T Easy vector.

3.2 MATERIALS AND METHODS

3.2.1 Oligonucleotide Design (-20 mer)

The primers were designed after carrying out a bioinformatic database search that enabled accessing sequences of the available *Streptomyces* IPNS gene from GenBank. These sequences with their respective accession numbers were from *S. clavuligerus* (M19421), *S. cattleya* (D78166), *S. lipmanii* (M22081), *S. jumonjinensis* (M36687), *S. griseus* (X54609) and *S. lactamdurans* (*Amycolatopsis spp*, X57310). The forward primer was selected corresponding to the first 20 bp of their respective IPNS genes whereas the reverse primer was the end 20 bp. These selected sequences were sent to Ransom Hill Bioscience (California) where they were synthesized and sent as ethanol precipitated oligos.

TABLE 3.1. Sequences of the primers used.

Streptomyces spp	**Forward** 5' 3'	**Reverse** 5' 3'
S. jumonjinensis	1 a c g g c g c g c g g a c g g g t c a g	1026 a g t t c t t g c c g g t c t g g a c t
S. lipmanii	1 c g t t c t g t g g c g t g c g g t a c	1103 g t g t c g c a c c a c g g g c c g c c
S. cattleya	1 g t t g c t c g c a g a c c a g t g g c	1329 g g c c g g c c c g c a g g g c g c a c
S. griseus	1 t a c g g g t a a g g g t a c g a c g g	980 g c t t c t t g c c g g t c t g g a c t
S. lactamdurans	11030 t t t t a c g g c a g t c g c c t t c a	11984 a c t t c t t g c c g g t c t g g a c t
S. clavuligerus	3390 c g g c t g g c c a c g c c t t a c a c	980 a g t t c t t g c c g g t c t g g a c t

TABLE 3.2. Reagents and final concentrations used in all PCR reactions.

REAGENT	FINAL CONCENTRATION
Buffer 10X	1X
dNTP 10mM	0.2mM each
$MgCl_2$	1.5 - 2.0mM
Forward & Reverse primers	0.5μM
Template DNA	10ng
Taq DNA polymerase	2.5 units

For all IPNS amplification protocols the ingredients shown in Table 3.2 were used to obtain the respective final concentration. A Master mix was prepared for multiple reactions to minimize reagent loss and to enable accurate pipetting especially since small volumes were used.

The initial PCR reaction volume used was 25μl since conditions were being optimized both in terms of concentrations required and also the amplification conditions which involved numerous attempts at different annealing conditions until optimal conditions were established. The following PCR conditions were found to be ideal for most strains and were carried out in the HYBAID OmniGene.

TABLE 3.3. PCR conditions used for IPNS amplification.

CONDITION	TEMPERATURE	DURATION
Initial denaturation	94^{0}C	2 min
Denaturation	94^{0}C	45 sec
Annealing	55^{0}C	30 sec
Extension	72^{0}C	60 sec
Final extension	72^{0}C	3 min

The above conditions were used with a modification that involved a 'Hot Start' procedure that meant putting in the Taq DNA polymerase after the initial denaturation and allowing the temperature to decrease until it reached 80^{0}C where it was held to allow the addition of Taq DNA polymerase. The strains used required different amplification conditions that differed only in the annealing temperature. For example, *S. jumonjinensis* required 54^{0}C, *S. lipmanii*, *Amycolatopsis spp* 53^{0}C and *S. cattleya* 60^{0}C.

TABLE 3.4. Volumes of reagents used for PCR reactions.

REAGENT USED	VOLUME (µL)
Sterile dH_2O	36.0
Buffer 10X	6.0
dNTPs	1.5
MgCl2	1.5
Primer F	1.0
Primer R	1.0
Acetamide (4%)	1.0
Taq	1.0
DNA	1.0

35µl of mineral oil was used to overlay the reaction mixture in order to avoid evaporation of the reagent components.

3.2.2 Preparation of Competent Cells

Competent cells were prepared using *E. coli* DH5α as follows. 5ml of LB broth was dispensed into a sterile test-tube, inoculated with a colony from a culture plate of *E.coli* DH5α and incubated at 37^{0}C overnight on a horizontal Labcon shaker. Four sterile flasks were labelled 1-4, to each 100ml of LB broth was poured, followed by addition of 1.5ml, 1.0ml, 0.7ml and 0.3ml of the overnight culture in the respective tubes. The flasks were incubated at 37^{0}C on a Labcon shaker with speed set at 150 rpm. Samples were taken from the flask inoculated with 1.5ml culture to check for the desired OD of 0.8 which was obtained after 2.5 hrs incubation and read at 600nm. When the desired OD was achieved, the flasks were placed on ice for 10 min at which point the contents were transferred to pre-autoclaved centrifuge bottles (Beckman).

The cells were centrifuged in a Beckman centrifuge using a JA 14 rotor at 5,000 rpm for 10 min with the temperature set at 4^{0}C. The supernatant was discarded and the pellet was carefully resuspended in 50ml RF_1 solution (Appendix E) and incubated on ice for 20 min. The solution was centrifuged as described above in the previous centrifugation step. The supernatant was decanted and each of the four pellets was resuspended in 4ml RF_2 solution (Appendix E). The flask contents were pooled and aliquots of 450ml were dispensed into sterile pre-cooled eppendorf tubes. The cells were stored at -70^{0}C bio-freezer until required having been tested for their competency and transformation efficiency.

3.2.3 Cloning of the PCR Product

The PCR product was used for cloning using the pGEM-T easy vector kit using the following volumes.

Table 3.5. The ingredients of the ligation reaction used for the p-GEM T easy vector kit.

Component	Std reaction (µl)	+ve control (µl)	Background control (µl)
2X Rapid ligation buffer T_4 DNA ligase	2.5	2.5	2.5
p-GEM T easy vector(50ng)	0.5	0.5	0.5
PCR product	1.5	-	-
Control insert DNA	-	1.0	-
T4 DNA ligase (3 weiss units/µl)	0.5	0.5	0.5
dH_2O to 5µl	-	0.5	1.5

The ligation reactions were incubated overnight at 4^0C since the maximum number of transformants were required. 2µl of each ligation reaction was transferred to a sterile eppendorf together with another tube of about 0.1ng uncut plasmid for determination of transformation efficiency of the competent cells. This cells were removed from -70^0C freezer and allowed to thaw on ice for 5 min. 50µl of the competent cells were carefully transferred to each tube containing 2µl of ligation reaction whereas 100µl was used for the tube containing the uncut plasmid, the tubes were gently flicked and placed on ice for 20 min. This was followed by heat shock of the cells for 45 sec at 42^0C in a water-bath followed by incubation on ice for 2 min.

950µl room temperature SOC medium was added to the tubes containing cells transformed with the ligation reaction whereas 900µl was used for the uncut plasmid. The tubes were incubated for 1.5 hrs at 37^{0}C with shaking at 150 rpm. Since a higher number of colonies was desired, the cells were pelleted by centrifugation at 1000 x g for 10 min, resuspended in 200µl SOC followed by plating 100µl of the transformed cells on two LB/amp/IPTG/XGal plates. The plates were incubated at 37^{0}C overnight. Blue/White screening of colonies was used to identify colonies that contained the gene of interest.

3.2.4 Long-term storage of clones

A 96 well cell culture cluster plates (Costar) were used for the storage of the clones. 100µl of LB broth with ampicillin (100µgml^{-1}) containing glycerol to a final concentration of 7% was transferred to each of the wells. Individual colonies were picked with sterile toothpicks and inoculated into the broth in the wells. The plates were then incubated at 37^{0}C for 12 - 16 hrs with gentle rocking on a labcon horizontal shaker followed by sealing the plates with parafilm before being stored at -78^{0}C.

3.2.5 Plasmid DNA Extraction

The white colonies were re-streaked and incubated overnight. Individual colonies were used to inoculate LB broth with ampicillin (100µl ml^{-1}) overnight followed by plasmid DNA isolation using either CTAB method (mini-prep plasmid DNA isolation) or High Pure Plasmid isolation kit following the manufacturer's protocol. CTAB method was carried out as follows. Bacterial cells were pelleted from a 1.5ml volume, supernatant discarded, and the pellet was resuspended in 200µl STET with the addition of 5µl lysozyme and RNaseI (4µgml^{-1}), incubated for 5 min at room temperature followed by boiling for 45 sec. The tubes were spun for 10 min at 13,000 rpm

(Heraeus biofuge^{R}). The pellet obtained was removed using a 200µl pipette. 8µl CTAB was added to the remaining solution and spun for 5 min at 13,000 rpm. The pellet was resuspended in 300µl NaCl and 750µl of 100% ethanol was added. The solution was spun for 10 min at 13,000 rpm, the supernatant was discarded and the pellet was washed with 800µl of 70% ethanol, dried by placing the eppendorf at room temperature inverted on a sterile absorbent paper for 45 min. The dry pellet was resuspended in 20µl of TE buffer and stored at -20^{0}C.

3.3 RESULTS

3.3.1 IPNS Amplification

The PCR conditions were successfully optimized for all the strains and the IPNS gene was amplified from all species. The annealing conditions differed amongst the strains used *S. clavuligerus*, *S. hygroscopicus*, *S. tanashiensis*, *S. griseus* and Thermophilic A and B annealed at 55^{0}C (Figure 3.3) whereas *S. cattleya* annealed at 60^{0}C using a hot start protocol, *S. jumonjinensis* amplified at 53^{0}C, *S. lipmanii* at 62^{0}C and finally *Amycolatopsis* spp at 57^{0}C (Figure 3.2) and visualized on an agarose gel.

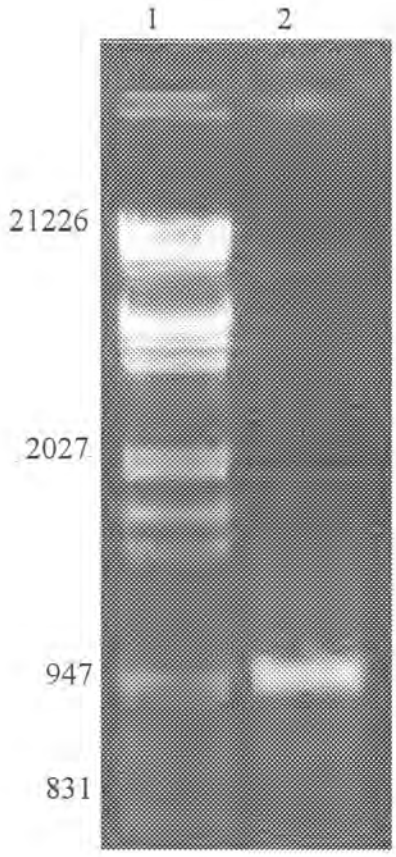


Figure 3.2. Agarose gel electrophoresis showing PCR product corresponding to the IPNS gene, Lane 1 Molecular weight marker III; Lane 2 *Amycolatopsis spp*.

The following table show the sizes of the amplified IPNS gene estimated against the marker used.

Table 3.6. Sizes of the IPNS genes estimated from the agarose gel migration patterns.

Strain	Estimated size (bp)
S. clavuligerus	980
S. cattleya	980
S. griseus	960
S. hygroscopicus	840
S. jumonjinensis	1100
S. lipmanii	980
S. tanashiensis	840

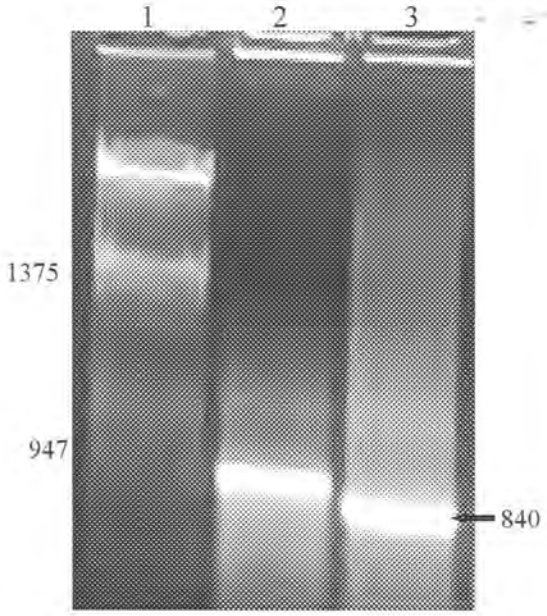


Figure 3.3. Agarose gel electrophoresis of the amplified IPNS gene by PCR from the two thermophiles used. Lane 1 Molecular weight marker III; Lane 2 Thermophilic A and Lane 3 Thermophilic B.

The IPNS gene was amplified using PCR for *S. clavuligerus* and the PCR product was visualized on an agarose gel and recorded using a Kodak Digital Science camera and software (Figure 3.4). The molecular weight of the amplified band was estimated to be 940bp which is in line with the documented IPNS gene in *S. clavuligerus*.

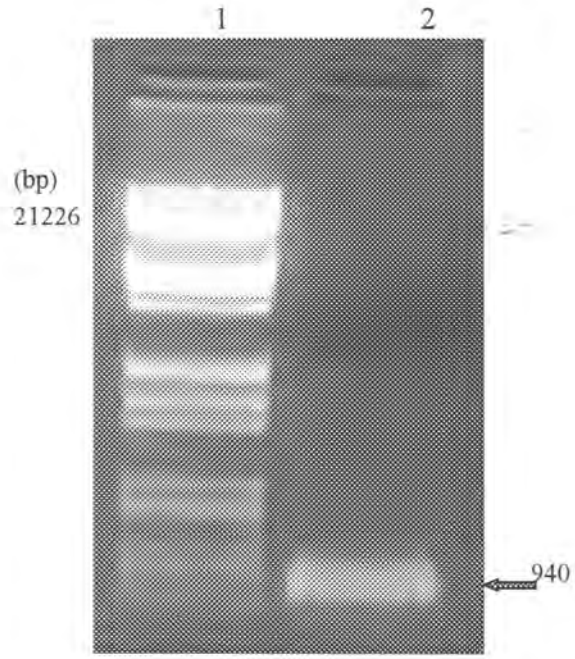


Figure 3.4. Agarose gel electrophoresis of the IPNS gene amplified from *S. clavuligerus*.

The gel photographs show an example of Plasmid DNA extracted from IPNS gene cloned into *E. coli* using a commercial kit (Figure 3.5) whereas (Figure 3.6) was obtained using CTAB mini-prep method. Whilst the two methods resulted in good yields, the commercial kit Plasmid DNA was much better.

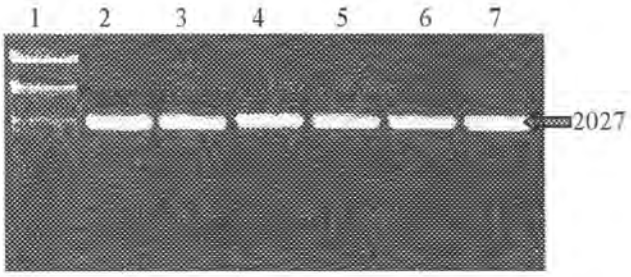


Figure 3.5. Agarose gel electrophoresis of Plasmid DNA obtained using High Pure Plasmid Isolation Kit (B & M), Lanes 1 & 2 *S. tanashiensis*; Lanes 3 & 4 *S. hygroscopicus*; Lane 5 Thermophilic A and Lane 6 Thermophilic B.

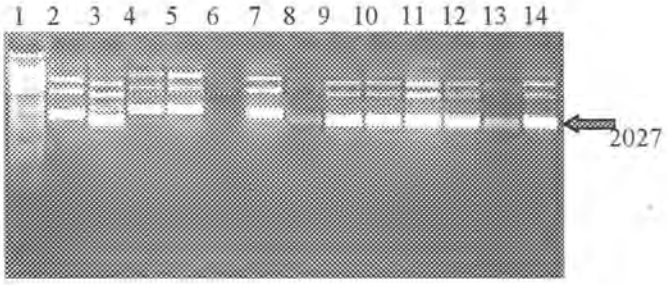


Figure 3.6. Agarose gel electrophoresis of Plasmid DNA obtained using CTAB mini-prep method, Lane 1 Molecular weight marker III; Lane 2 *Amycolatopsis spp*; Lane 3 *S. clavuligerus*; Lane 4 *S. cattleya*; Lane 5 *S. griseus*; Lane 6 & 7 *S.hygroscopicus*; Lane 8 *S. jumonjinensis*; Lane 9 *S. lipmanii*; Lane 10 & 11 *S. tanashiensis*; Lane 12 & 13 Thermophilic A; Lane 14 Thermophilic B.

The presence of the correct inserts in the clones was verified by restriction digests of Plasmid DNA with Eco RI to enable sizing both the vector and the inserts (Figure 3.7 and 3.8).

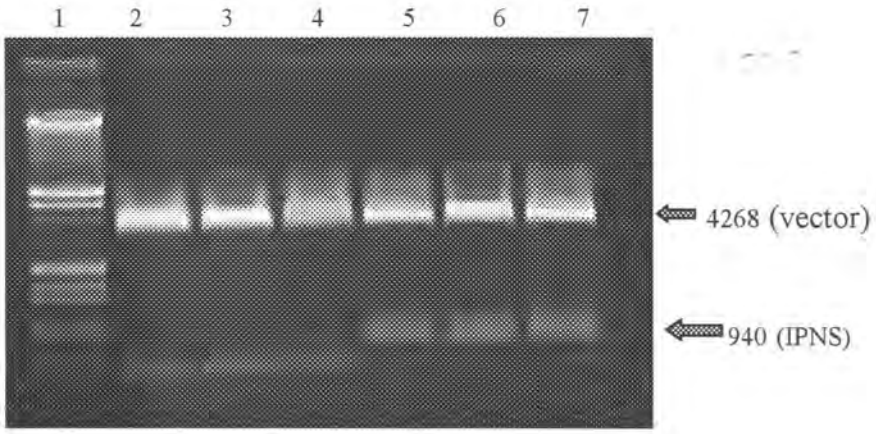


Figure 3.7. Agarose gel electrophoresis of Plasmid DNA restricted with EcoRI to verify the presence of the correct inserts, Lane 1 Molecular weight marker III; Lane 2 - 4 *S. hygroscopicus*; Lane 5 - 7 *S. tanashiensis*.

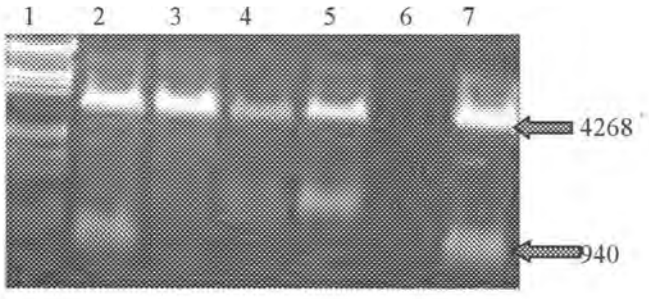


Figure 3.8. Agarose gel electrophoresis of Plasmid DNA restricted with EcoRI from clone of Thermophilic A and B, Lane 1 Molecular weight marker III; Lane 2- 4 Thermophilic A clones 1-3 respectively; Lane 5 - 7 Thermophilic B clones 1-3 respectively.

3.3. DISCUSSION

The polymerase chain reaction was found to be a rapid and reliable method of amplifying the IPNS gene. Although theoretically quite simple, the successful amplification involved optimizing a number of variables, including the quality and quantity of DNA, PCR conditions and reaction volumes and concentrations of the components.

The IPNS gene amplification by PCR using primers designed according to the Bioinformatics information from GenBank on some of the β-lactam producing strains (*S. clavuligerus*, *S. cattleya*, *S. griseus*, *S. jumonjinensis*, *S. lipmanii* and *Amycolatopsis spp*) from their gene sequences enabled successful amplification of the IPNS gene (Figure 3.2, 3.3 and 3.4) for all the strains used, as well as amplification of the IPNS-like gene from two novel strains, Thermophilic A and B that were isolated in our lab. The success in using these designed primers for *S. hygroscopicus*, *S.tanashiensis* and the two thermophiles suggest that the primers were specific enough for PCR-amplification yet sufficiently degenerate for use on all ten strains. Again, the IPNS gene amplification success may be due to the gene being conserved to a certain degree which is to be expected for *Streptomyces spp*. The use of PCR for IPNS amplification allows for a faster and easier method for detecting β-lactam producing organisms, as noted for the two thermophiles.

The IPNS gene was successfully cloned into *E. coli* using p-GEM T easy vector and Plasmid DNA was successfully extracted using a commercial kit and CTAB mini-pre method (Figure 3.5 and 3.6) respectively. The presence of the correct insert was verified by restriction digests of Plasmid DNA with Eco RI that gave two bands which corresponded to the expected vector and IPNS bands respectively (Figure 3.7 and 3.8).

CHAPTER 4

CLONING AND EXPRESSION OF THE IPNS GENE

4.1 INTRODUCTION

Cloning of amplified DNA is often a difficult step in analyzing the products of a polymerase chain reaction. Several techniques have been developed to facilitate cloning of PCR products. A method for cloning is chosen considering the following, the purpose of the cloning experiment and the length of the PCR product. If a DNA polymerase with proofreading activity such as Pwo DNA polymerase is used for site-directed mutagenesis, then the suitable cloning procedure is blunt-end. On the other hand, a method that allows directional cloning is suitable for protein expression (Gusek and Kinsella 1992; Hidaka *et al.* 1992).

4.1.1 TOPO TA Cloning

TA cloning using pTrcHis2 TOPO TA cloning^{R} kit is one of the simplest and most efficient methods providing a highly efficient, quick cloning strategy which allows for direct insertion of Taq polymerase-amplified PCR product into a plasmid vector for expression in *E. coli*. The procedure exploits the terminal transferase activity of certain thermophilic DNA polymerase, including Thermus aquaticus (Taq) polymerase. Taq polymerase has non-template dependent activity which preferentially adds a single adenosine to the 3'-ends of a double stranded DNA molecule and thus most of PCR amplified products by Taq polymerase possess 3'-A overhangs. The use of a linear 'T-vector' which has a single 3'-T overhangs on both ends allow direct, high efficiency cloning of PCR products and is facilitated by complementarity between the PCR product 3'-A overhangs and vector 3'-T overhangs. With a single T-vector at hand any DNA

fragment can be cloned without compromising the cloning efficiency. Ligation, post-PCR procedures and PCR primers with specific sequences are not required in this cloning procedure. However, the PCR primers should enable one to obtain the recombinant protein required, which in this study was the IPNS protein. PCR product was used immediately in TOPO cloning since long-term storage results in removal of 3' A-overhangs from the PCR product and hence decreased cloning efficiency (http://www.invitrogen.com).

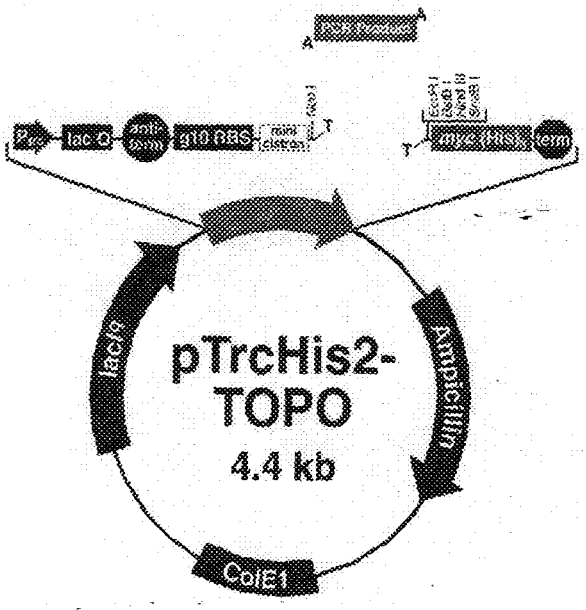


Figure 4.1. Circular vector map of pTrcHis2-TOPO (http://www.invitrogen.com).

4.1.2 High Performance Liquid Chromatography

High performance liquid chromatography (HPLC) is one of the most important and popular analytical techniques for the determination of the concentration of non-volatile and temperature-sensitive compounds like antibiotics. It is easy to learn and use and is not limited by the volatility or stability of the sample compounds, and its many applications include separation, identification, purification and quantification of various compounds. Because of the method of the production of most important β-lactam antibiotics, such as penicillin G and V as well as cephalosporin C, aseptic sampling of the cultivation medium is used. A careful preconditioning is necessary, because of the presence of several matrix components which impair analysis (Jensen *et al.* 1982; Schugerl and Seidal 1998; http://kerouac.pharm.uky.edu/ASRG/HPLC/hplcmytry.html).

Antibiotic analysis using HPLC requires the use of cell-free extracts due to the fact that most antibiotics that are produced by microorganisms are transported actively or passively out of the cells into the aqueous phase. Analysis of β-lactam compounds from broth cultures is carried out on samples that are completely deproteinated by methanol and microfiltered. The β-lactam antibiotics have measurable absorbance in the UV wavelength range, therefore their concentration can be measured by UV detection. Pre or post column derivatization can be used to improve the detection sensitivity by the formation of fluorophore products (Schugerl and Seidal 1998).

The analytical column being the heart of the HPLC system needs to be selected appropriately in order to obtain a good separation of the compounds in the broth cultures. For antibiotics reversed-phase HPLC columns are most often used, for example Shodex C_{18} 5A or similar packing materials. The particle size is usually 5μm and the sample flow 1.0-1.5ml min^{-1}. The

mobile phases consist of various mixtures of acetonitrile and phosphate buffer (Schugerl and Seidal 1998).

4.2 MATERIALS AND METHODS

4.2.1 TOPO Cloning

TOPO cloning and transformation was carried out by using 4µl of the amplified IPNS PCR product with an average length of 850 base pairs to give a proper insert : vector ratio for TOPO^R. 1µl of the TOPO^R vector was added to the PCR product, mixed gently and incubated at room temperature for 5 min, taking care not exceed this incubation time as that would decrease the transformation efficiency. 2µl of the TOPO^R cloning reaction was added into a vial of the supplied one Shot^R competent cells and mixed gently. The reaction mixture was incubated on ice for 30 min, followed by heat shock of the cells for 30 sec at 42^{0}C and immediately placed on ice. 250µl of room temperature SOC medium was added, tubes were capped tightly and shaken horizontally on a Labcon shaker at 37^{0}C for 30 min. Colonies obtained were used to inoculate LB broth containing 50µg ml^{-1} ampicillin and 0.5% glucose followed by overnight incubation to facilitate Plasmid DNA isolation, using High Pure Plasmid Isolation Kit according to the manufacturer's protocol.

4.2.2 Expression of IPNS Protein in *E. coli*

For each of the strains used, a single recombinant *E.coli* colony was used to inoculate 2ml of LB broth containing 50µgml^{-1} and incubated at 37^{0}C overnight with shaking at a speed of 200 rpm. The next day, 10ml of the LB broth with ampicillin (50µgml^{-1}) was inoculated with 0.2ml of the overnight culture and this was grown with shaking for about 2 hrs to an OD_{600} =0.6

corresponding to when the cells had reached mid log phase. 1ml aliquot of cells were removed, placed in eppendorf tubes and centrifuged at maximum speed in a micro-centrifuge for 30 sec. The supernatant was discarded and the pellet obtained was stored at -20^{0}C and taken to represent zero time point sample. IPTG was added to the broth cultures to give a final concentration of 1mM and grown at 37^{0}C with shaking. 1ml samples were collected every hour for a period of 5 hrs, followed by another sample that was taken after 12 hrs and the cell pellets were collected as previously described for the zero time point sample. The tubes were labeled to correspond with the number of hours post-induction.

4.2.3 Analysis of Time Point Samples

SDS-PAGE gels were prepared using Tall Mighty Small vertical slab units and applied as described by Laemmli (1970) with a stacking gel of 4% and a resolving gel of 10% (Appendix E). When all the time points were collected, each of the pellet was resuspended in 100µl of 1X SDS-PAGE sample buffer, boiled for 5 min and centrifuged using a top-bench micro-centrifuge. When the solution were viscous, they were sonicated for 5 sec and centrifuged again. Each sample was analyzed by loading 5µl into the wells of the SDS-PAGE gel and ran at 80V for 2.5 - 5 hrs. After electrophoresis, the top stacking gel was cut off and also the bottom plug leaving the resolving gel which was stained with Coomasie blue (Appendix E) for 3 hrs and de-stained overnight with three changes of De-stain solution (Appendix E). The positive control was included in the analysis of the samples to confirm that the growth and induction had been done properly whereas the negative control was used to distinguish recombinant proteins from background proteins. The positive control was TOP 10 cells containing pTrchis2-TOPO^R having been transformed into TOP10 cells whereas the negative control was the TOP 10 cells without

any vector.

4.2.4 Preparation of *Streptomyces* Cell-free Extracts

The *Streptomyces* used in this study were grown in M3 broth for 72 hrs. An equal volume of IPNS wash buffer (100ml) (Appendix C) was added to the broth cultures and the solutions were centrifuged for 10 min at 10,000 rpm. The cells were resuspended in 10ml original culture volume of IPNS assay buffer (Appendix C) and sonicated whilst on ice in a cooler box for 15 sec, and this was done twice at maximum intensity. Broken cell suspensions were centrifuged for 45 min at 10,000 rpm using a JA 14 rotor in a Beckman J2-21 centrifuge. Aliquots of cell-free extracts were placed in eppendorf tubes and stored at -20^{0}C until required. ACV (Bachem AG, Switzerland) was added just prior to analysis to obtain 0.05μgml^{-1} for each sample, incubated at room temperature for 10 min and analyzed using HPLC as described below.

4.2.5 IPNS Assay Using HPLC

HPLC analysis was carried out using HPLC model Lachrom L-7400 (Merck, Hitachi, Germany), equipped with a variable UV/VIS detector and a 80 position autosampler/autoinjector (20µl). Separation was achieved on a 4.6 x 250mm internal diameter Spherisorb S5ODS1 waters (U.S.A.) C18 reverse phase column. A guard column was positioned just before the analytical column. A column temperature of 25^{0}C and detector wavelength set at 220nm, flow rate of 1.0ml min^{-1}, run time of 20 min and a mobile phase of methanol and 0.05M potassium phosphate buffer p.H 7.0 (10:90) respectively that was degassed and filtered before use. Methanol was added to the samples that were analyzed to give a final concentration of 0.4% v/v to inactivate the reaction mixtures and precipitate the protein. They were centrifuged at 10,000 x g for 5 min to remove the precipitated proteins followed by filter sterilization using a 0.22µm durapore filter (Millipore,

Harrow U.K.) into autoclaved sample vial bottles.

4.3 RESULTS

4.3.1 Expression of the IPNS protein in *E. coli*

Broth cultures that were induced with IPTG as explained in Methods and Materials and samples collected over time resulted in protein profiles observed in SDS-PAGE gels. The expected molecular weight of the IPNS is 39.0 kDa (Figure 4.2).

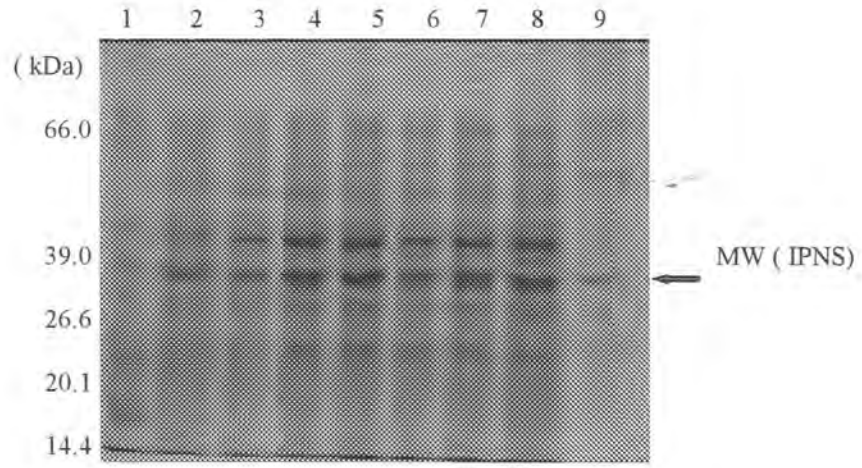


Figure 4.2. SDS-polyacrylamide gel electrophoresis of proteins stained with Coomassie Brilliant Blue obtained from *S. griseus* IPNS gene cloned into *E. coli* (TOP 10') Lane 1 Protein marker 14.4 - 94.4 kDa; Lane 2 pre-induction sample; Lanes 3 - 8 Post-induction samples collected after 1, 2, 3, 4, 5 & 12 hrs respectively.

These are examples of SDS-PAGE gels that were recorded to show the result obtained when expressing the IPNS gene in *E. coli*. The novel Thermophilic B showed very strong bands corresponding to the IPNS protein at 39 kDa (Figure 4.3) suggesting that the IPNS gene was successfully expressed in *E. coli*.

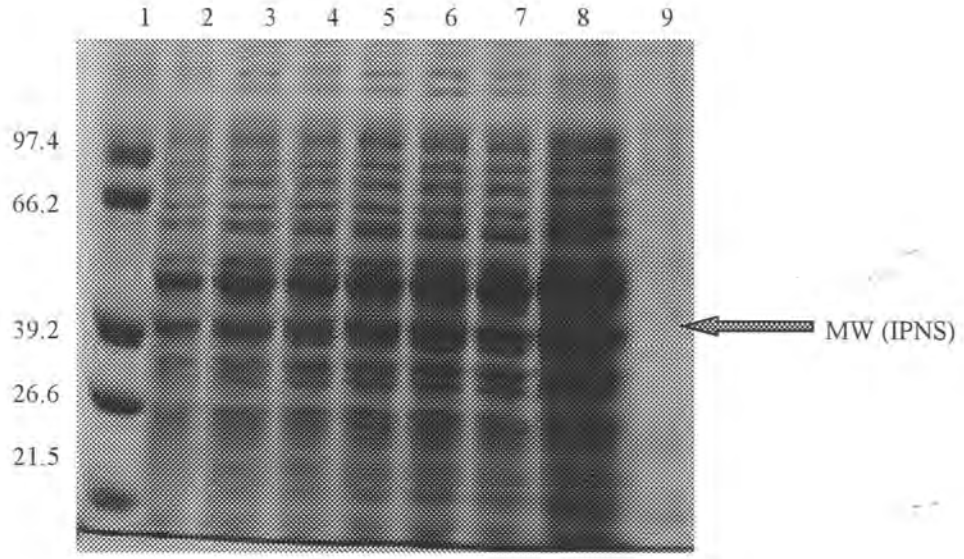


Figure 4.3. SDS-PAGE gel electrophoresis of proteins stained with Coomassie Brilliant Blue from Thermophilic B IPNS gene cloned into *E. coli* (TOP 10'), Lane 1 Protein marker 14.4 - 94.4 kDa; Lane 2 Pre-induction sample; Lanes 3 - 8 Post-induction samples after 1, 2, 3, 4, 5, and 12hrs respectively; Lane 9 Protein color marker.

Thermophilic A and *S.tanashiensis* IPNS genes when expressed in *E. coli* (Figure 4.4 and 4.5) respectively also show the presence of IPNS protein despite the degree of expression being lower than observed in Thermophilic B (Figure 4.3).

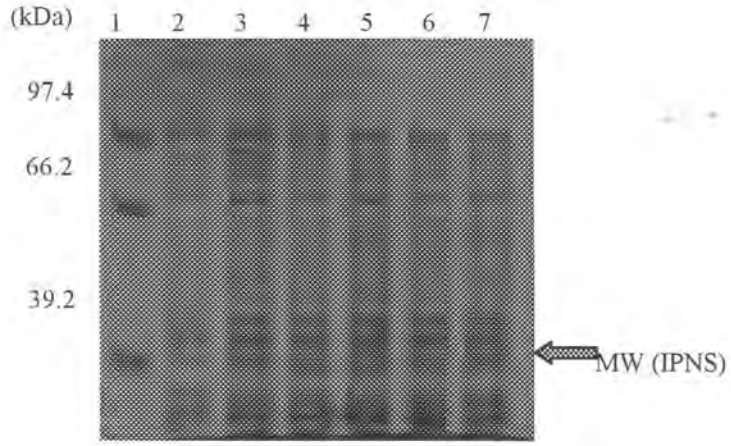


Figure 4.4. SDS-PAGE gel electrophoresis of proteins stained with Coomassie Brilliant Blue from Thermophilic A cloned into *E. coli* (TOP 10'), Lane 1 Protein marker 14.4 - 94.4 kDa; Lane 2 Pre-induction sample (0hrs); Lane 3 - 7 Post-induction sample 1, 2, 3, 4 and 5 hrs respectively.

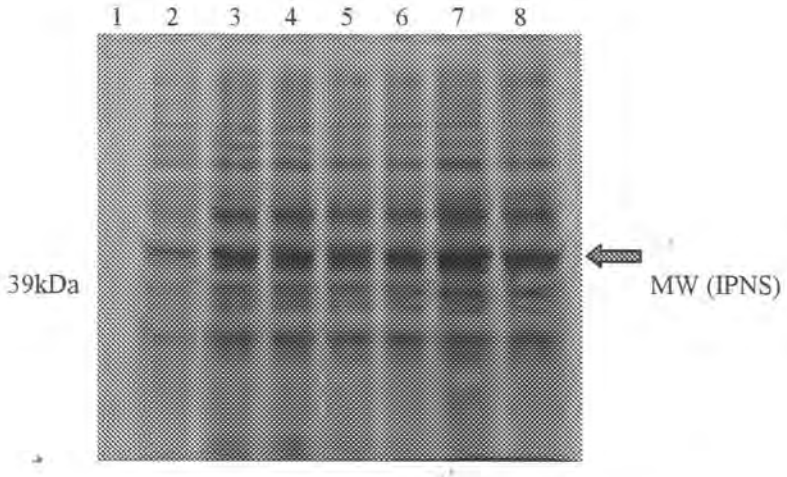


Figure 4.5. SDS-PAGE gel electrophoresis of proteins stained with Coomassie Brilliant Blue from *S. tanashiensis* cloned into *E. coli* (TOP 10'), Lane 1 Protein color marker; Lane 2 Pre-induction sample; Lanes 3 - 8 1, 2, 3, 4, 5 and 6 hrs Post-induction.

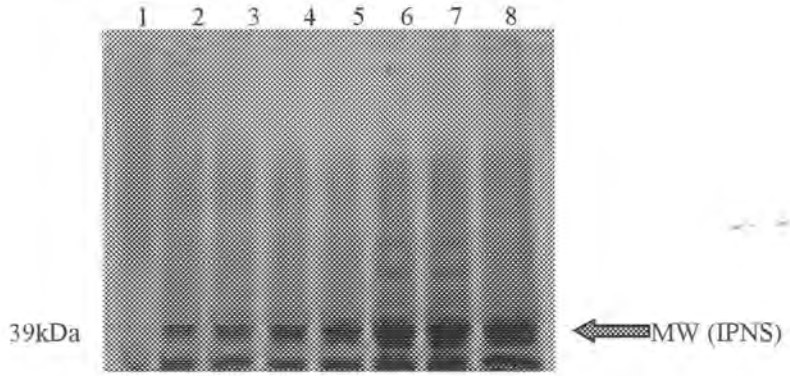


Figure 4.6. SDS-PAGE gel electrophoresis of proteins stained with Coomassie Brilliant Blue from *Amycolatopsis spp* cloned into *E. coli* (TOP 10'), Lane 1 Protein color marker; Lane 2 Pre-induction sample; Lanes 3 - 8 Post-induction samples after 1, 2, 3, 4, 5 and 12hrs.

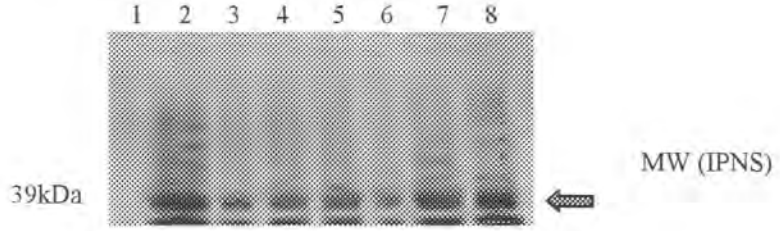


Figure 4.7. SDS-PAGE gel showing protein profiles stained with Coomassie Brilliant Blue from *S. cattleya* cloned into *E. coli* (TOP 10'), Lane 1 Protein color marker; Lane 2 4hr post-induction; Lane 3 1hr; Lane 4 2hr; Lane 5 3hr; Lane 6 Pre-induction sample (n.b the sample mix-up when loading since in lane 2 had intended to load the sample in Lane 6); Lane 7 5hr; Lane 8 12hr post-induction.

HPLC analysis was used as a method for determining the presence of antibiotics in the cell-free extracts of the *Streptomyces* strains. 6-aminopenicillanic acid was used as a control (Figure 4.8) to determine which biosynthetic route in the pathway the organism is likely to follow (Figure 1.2). Ampicillin and Pen VK were also used as controls to establish the production of penicillin by the organism (Figure 4.9 and 4.10).

The cell-free extracts from the *Streptomyces* strains showing the activity of IPNS (Table 4.2). However the presence of an IPNS gene does not necessarily correlate with the production of Pen VK *in vivo*.

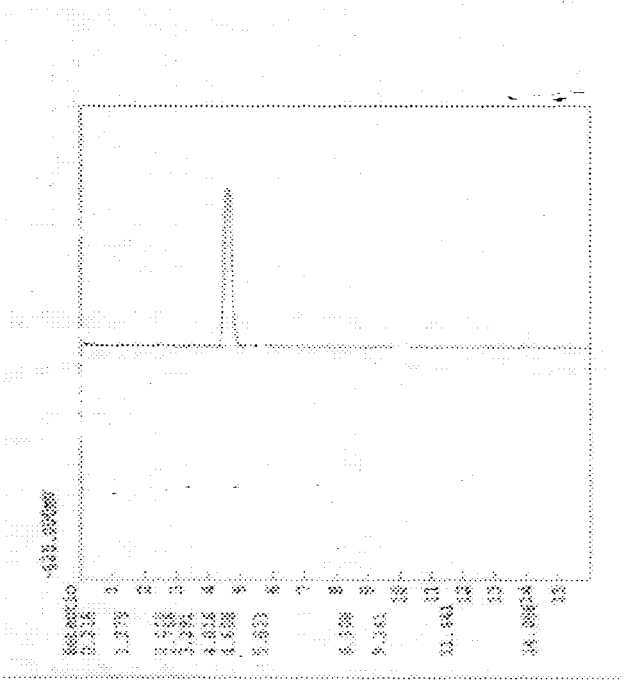


Figure 4.8. HPLC chromatogram of 6 aminopenicillanic acid showing a retention time of 4.6.

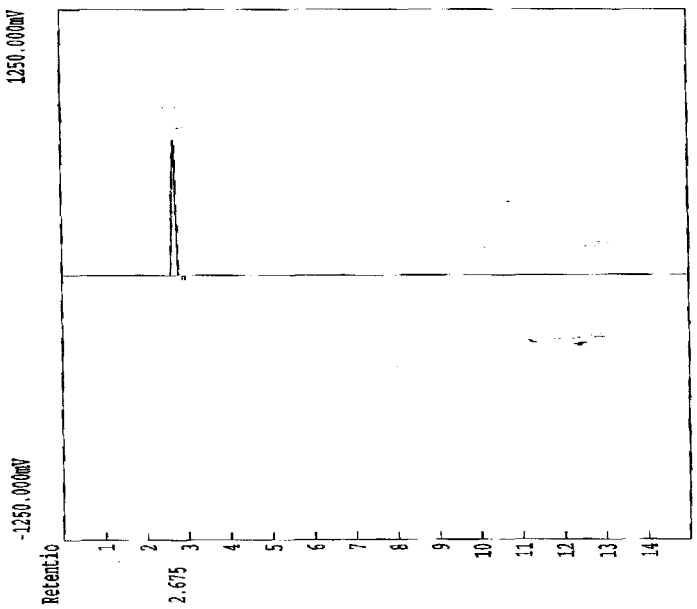


Figure 4.9. HPLC chromatogram of ampicillin showing a retention time of 2.675.

Table 4.1. Retention times obtained for the other standards and solutions used.

Antibiotic	Retention time
Penicillin V K	2.6
Ampicillin	2.6
6-aminopenicillanic acid	4.6
DTT	14
ACV	7.2

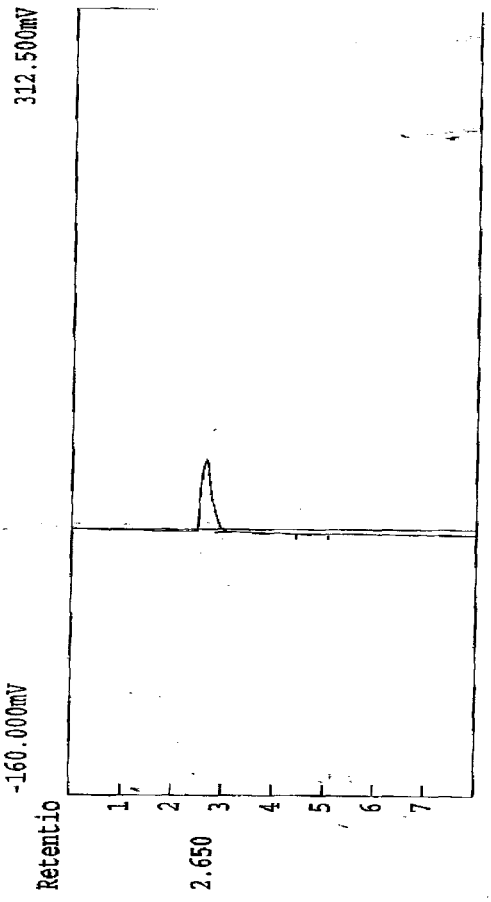


Figure 4.10. HPLC chromatogram of Pen VK showing a retention time of 2.65.

Table 4.2 Retention times obtained for the *Streptomyces* strains used.

Streptomyces spp	**Retention time (Min) (Before expression)**	**Retention time (Min) (After expression)**
S. clavuligerus	2.5, 4.6, 5.9, 11.4	5.1, 7.2
S. cattleya	4.6, 7.2, 14.0	2.2, 2.6, 3.2, 5.1
S. griseus	3.0, 4.6, 7.2, 14.0	2.6,3.2
S. hygroscopicus	2.6, 4.6, 7.1, 13.9	2.2, 2.6, 4.6
S. jumonjinensis	2.6, 4.6, 7.1, 14.0	7.1
S. lipmanii	2.6, 7.9, 14.0	2.6
S. tanashiensis	2.6, 4.6, 7.2, 14.0	-
Amycolaptosis spp	2.6, 3.0, 4.6, 7.2, 14.0	3.2, 5.1, 7.2
Thermophilic A	4.6, 7.2, 14.0	2.2, **2.6**, 3.2, 5.1
Thermophilic B	4.6, 7.9, 14.0	2.2, **2.6**, 7.2, 7.9

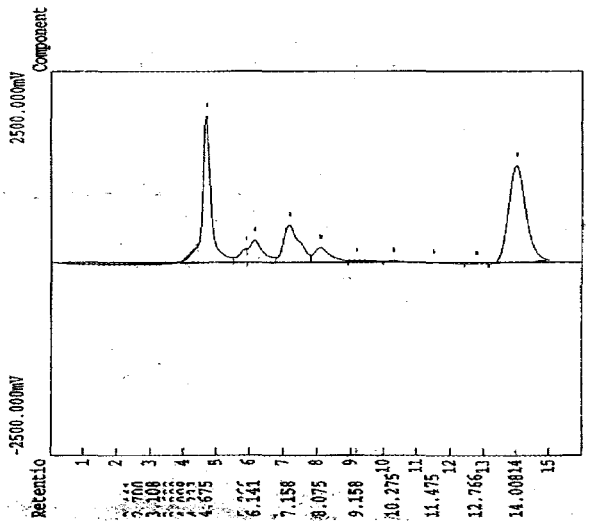


Figure 4.11. HPLC chromatogram of Thermophilic A showing a retention time similar to that of 6-aminopenicillanic acid at 4.6 and another peak at 14.0 similar to that of DTT.

The HPLC chromatograms of Thermophilic A (Figure 4.11 and 4.12) are used to show the retention times obtained from *Streptomyces* and their IPNS gene after being expressed in *E. coli* respectively. The results show that the IPNS gene was successfully expressed in *E. coli* since the IPNS activity was indicated by the retention times which is similar to that of penicillin, after expression.

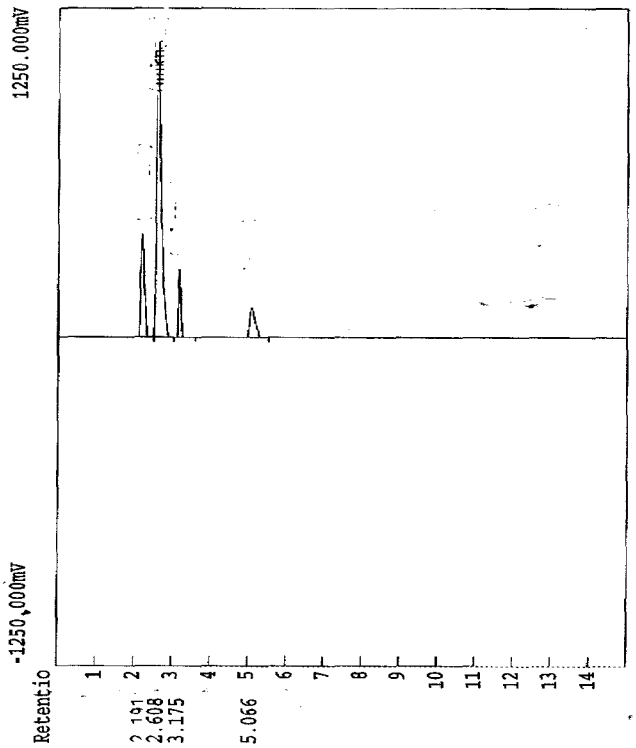


Figure 4.12. HPLC chromatogram of Thermophilic A IPNS gene after being expressed in *E. coli* showing a main peak with retention time of 2.6 similar to pen VK.

4.4 DISCUSSION

p-TrcHis2-TOPO kit was used for cloning of the IPNS gene into *E. coli*. The number of positive transformants was rather low for all the strains used. However this low number enabled screening of all the clones obtained. All the clones were found to have the right insert (IPNS gene).

The correct inserts were verified by use of *HindIII* and *EcoRI*. Three clones from each respective IPNS gene cloned into *E. coli* from the strains used were used to determine the success of the expression of the IPNS protein. IPTG was used to induce expression and analysis was carried out using SDS-PAGE gel electrophoresis with the commercial protein markers to analyse the protein profiles obtained. From these results the time at which there was maximum expression was determined and this corresponded to the lane at which that time point sample was loaded and also taking into account that the expected IPNS protein had a molecular weight of 39 kDa.

There was variation in the amounts of protein produced as determined from the SDS-PAGE gel protein profiles from the IPNS gene from the respective strains used and expressed in *E. coli*. Thermophilic B, *Amycolatopsis* and *S. cattleya* (Figures 4.3; 4.6 and 4.7) respectively showed the highest amount of expressed protein. This was also confirmed from the result obtained by Bradford's assay. This assay was carried out for all the time points collected from the clones of the respective strains used.

The stability of enzymes is essential for any enzymatic analysis to be meaningful therefore it was imperative that the IPNS assay be done on cell-free extracts which were prepared in a way that

IPNS activity and stability were ensured and this was achieved by use of DTT in the IPNS wash and assay buffers (Appendix C). IPNS enzyme is inactivated when oxidized therefore use of DTT ensured that this activity was maintained. The use of antibiotic standards enabled comparison of the profiles obtained against the test organisms. The similar retention times obtained for the standards and the test organisms confirmed the presence of the antibiotics even though we would have to go further with the analysis making use of Mass-spectroscopy and NMR to actually decipher the structure of the antibiotic found to be present.

IPNS assays using HPLC is a method recommended for use to determine the presence of antibiotics. HPLC analysis by virtue of being more sensitive though expensive would be the preferred method of determining the presence of the antibiotics being produced. The use of the HPLC for IPNS assay will allow the determination of changes in amounts of antibiotics being produced after cloning and expressing the IPNS gene in an *E. coli* system. HPLC confirmed the activity of IPNS since *E. coli* cannot make Pen VK. Thermophilic A and B produced Pen VK after expression (Table 4.2).

CHAPTER 5

GENE SHUFFLING

5.1 INTRODUCTION

Protein-engineering technology involves creating new proteins by modifying existing ones. This approach has been made possible by progress in two techniques: site-directed mutagenesis and computer-assisted modeling of the three dimensional structure of proteins. However, this approach is only applicable to protein families in which three-dimensional structure of at least one member protein has been resolved. Therefore, in order to obtain improved enzymes for potential commercial benefit, methods other than rational design are needed.

A practical strategy for altering enzyme properties is to introduce a random base substitution and then to select or screen for desired variants. The error-prone PCR which allows the introduction of random mutations is currently the preferred method for this molecular-breeding procedure. A critical point of this method is that mutation frequencies need to be carefully tuned. Generally, beneficial mutations are rare, while deleterious mutations are common because the combination of a beneficial mutation and a deleterious mutation may result in the formation of an inactive enzyme, the mutation frequency should not be too high if the maximum number of improved enzymes are to be obtained. The mutation frequency should also not be too low, otherwise the wild-type sequence would predominate in the population. The ideal number of base substitutions per targeted gene is usually between 1.5 and 5 (Harayama 1998).

5.2 MATERIALS AND METHODS

5.2.1 Gene Shuffling Methods

Two protocols were used successfully, which were carried out as follows: 10µl of the IPNS PCR product was digested with DNase I at a final concentration of 0.002 units/µl, and incubated at room temperature for 15 min. This was followed by performing PCR without primers with the components shown in the table below.

TABLE 5.1. Volumes of reagents for primer less PCR reactions in shuffling.

Component	Volume (µl)
DNase digested sample	10.0
Sterile dH_2O	30.0
10X buffer	5.0
dNTPs	2.0
$MgCl_2$	2.0
Taq	1.0

Taq DNA polymerase was added after initial denaturation when the temperature was decreased to 80^0C using the PCR temperatures previously mentioned, except that the program was set in such a way that 5µl samples were collected after 15 cycles, 20, 30, and 35 which corresponded with 90 min for 15 cycles and thereafter every 25 - 30 min. The program was set in such a way that 5µl samples were collected with the temperature held at 42^0C.

Phase 2 of the shuffling program was carried out by diluting the primer-less PCR product 1:10 as follows.

TABLE 5.2. Reassembly reagents.

Primer-less PCR product	5.0 (µl)
Sterile dH_2O	42.0
F` Primer	1.0
R` Primer	1.0
Taq	1.0

The PCR cycle was carried out on the above components for 15 cycles using the previously mentioned PCR program.

The alternative protocol was carried out using PCR products that had been purified using the Quaquick Gel extraction kit following the manufacturer's protocol using the microcentrifuge. DNase I digestion was carried out in the presence of Mn^{2+} in order to reduce the high rates of point mutations expected with gene shuffling. A digestion buffer (500mM Tris-HCl, pH 7.4, 100mM $MnCl_2$) was added to the PCR product to be shuffled together with DNase I to obtain a final concentration of 0.02 units. The mixture was incubated at room temperature for 5 min. This digestion was terminated by heating in the water-bath at 90°C for 5 min. Primer-less PCR was carried out as described previously, followed by a 1:10 dilution of the primer-less PCR product which was then subjected to the same PCR conditions as previously described.

5.2.2 Agarose Gel Electrophoresis

Agarose (LE, analytical Grade, Promega) gels were prepared as outlined in appendix D. For amplification of the IPNS gene (PCR product), 0.8% and 1.0% were prepared. Ethidium

Bromide, a fluorescent dye was incorporated into the gels to a final concentration of 0.5μg ml^{-1}(AppendixB).

All the agarose gels were run at 80V for 1-2 hrs at room temperature in either 1x TBE or 1x TAE using Hoefer Scientific Instruments (San Francisco) Ps 500x DC power supply. Gels were visualized by UV transilluminator using UV light at 300nm wavelength on a UVP incorporation Transilluminator and photographed using Kodak Imaging Systems (Kodak Digital Science 1D™). Ethidium bromide solutions were decontaminated before being sent for toxic waste disposal.

5.2.3 Restriction Endonuclease Digestion of the IPNS Gene

Restriction digests were carried out on all amplified IPNS (PCR product) of the ten strains used before and after shuffling using the enzymes *AluI*, *Sau3A*, *HaemIII*, and *RsaI* to facilitate comparison of the restriction profiles as a result of shuffling. The restriction enzymes were selected based on the result obtained from the Vector NTI programme that showed restriction sites on the available IPNS sequences from the database. Other enzymes were also used including *BamHI*, *BglII* and *MspI*.

Table 5.3. Restriction digest cocktail prepared as shown below for all the enzymes used.

Component	Volume (μl)	Final concentration
ddH_2O	5.5	-
Buffer (10X)	1.0	1X
Enzyme	0.5	1unit/μl
DNA	3.0	100ng

The volumes of each component were adjusted to obtain the above indicated final concentration based on the DNA concentration which was different for most of the strains used.

5.3. RESULTS

The IPNS gene were shuffled using modified protocol according to Stemmer (1994). These genes were fragmented with DNase I and they were found to reassemble together reconstituting the correct sized IPNS gene suggesting an internal rearrangement. The other shuffling procedure that was carried was between IPNS genes from two strains using *S. clavuligerus* with each of the following, *S. hygroscopicus*, *S. tanashiensis*, Thermophilic A and Thermophilic B. These resulted in a change in the size of amplified product by about 20 - 30 bp. The lanes 2 - 6 where no bands were seen is due to the DNA having been fragmented to small sizes with DNase such that we were unable to see them on the gel, however incorporating the primers allowed the fragment to reassemble together (Figure 5.1).

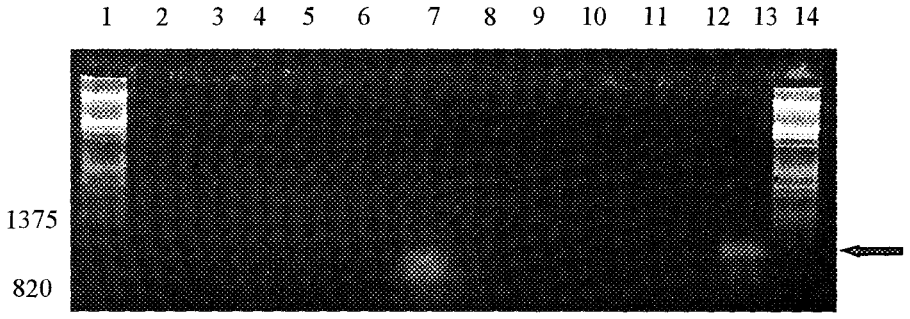


Figure 5.1. Agarose gel electrophoresis of shuffled PCR product, Lane 1 Molecular weight marker III; Lane 2- 6 DNase I digested primerless PCR product from *S. lipmanii* after 15, 20, 25, 30 & 35 cycles; Lane 7 *S. lipmanii* with primers following primerless PCR, Lane 8 - 12 DNase I digested primerless PCR product from *S. jumonjinensis* after 15, 20, 25, 30 & 35 cycles; Lane 13 *S. jumonjinensis* with primers following primerless PCR; Lane 14 Marker III.

The table below show changes in molecular weight for the *Streptomyces* as result of shuffling. The shuffling was done in two ways by shuffling the IPNS gene on its own for the respective strains used and also between the *Streptomyces spp* used in this study. There was no change observed in some strains, suggesting that either shuffling did not occur or there was an internal rearrangement which can only be confirmed by sequencing the shuffled product.

Table 5.4. Changes in size (bp) as a result of shuffling.

Streptomyces spp	**Size bp (shuffled on their own)**	**Shuffling between *S. clavuligerus* & other *Streptomyces* respectively**
S. clavuligerus	980	Not done
S. hygroscopicus	840	-
S. tanashiensis	840	920
Thermophilic A	**880**	940
Thermophilic B	**820**	860
S. cattleya	980	-
S. griseus	**880**	-
S. jumonjinensis	**920**	840
S. lipmannii	**940**	-
Amycolatopsis spp	960	-

Bold sizes represent changes in bp that occurred as result of shuffling IPNS genes on their own, whereas there were changes for others in column 2. *S. tanashiensis*, *S. jumonjinensis* and the two novel strains Thermophilic A and B were successfully shuffled together with *S. clavuligerus*.

(A) show the position of the IPNS band before shuffling and B is after shuffling. Figure 5.2 show the IPNS band at the same position for *S. hygroscopicus* before and after shuffling (lane 1 and 6) respectively and the same is observed for *S. tanashiensis*. The bottom wells in the gel show differences in the position of the IPNS band as a result of shuffling for the Thermophilic A and B.

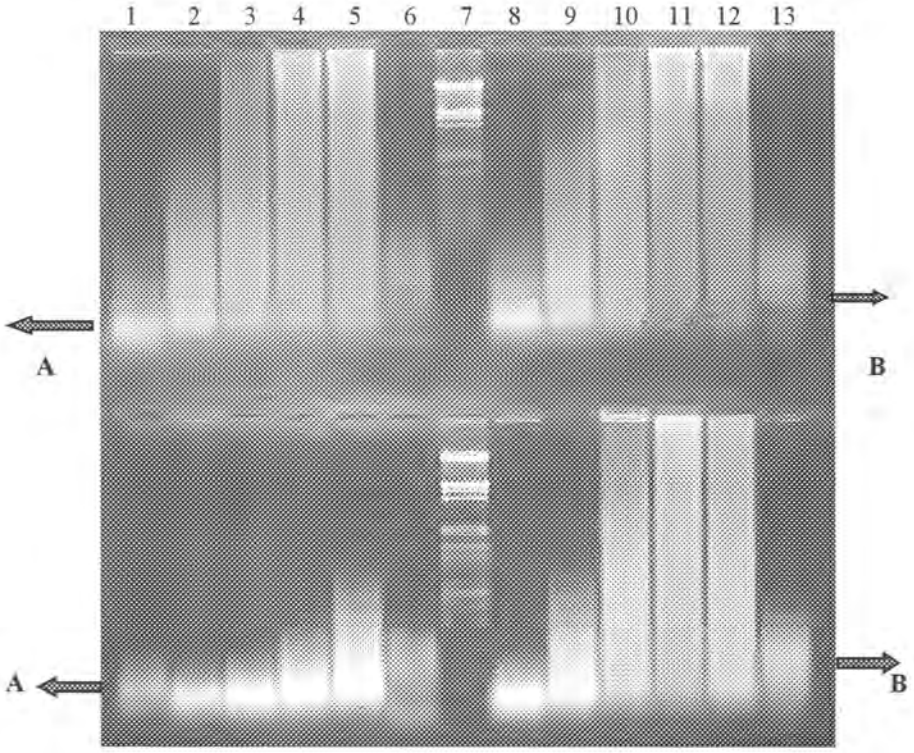


Figure 5.2. Agarose gel electrophoresis of the shuffled IPNS gene product obtained when shuffling two genes together. (Top gel) Lane 1 *S. hygroscopicus* IPNS PCR product before shuffling, Lanes 2- 5 Primerless PCR products after 15, 20, 25 and 30 cycles respectively; Lane 6 PCR product obtained with incorporation of primers to the shuffled primerless PCR product; Lane 7 Molecular weight marker III; Lane 8 - 13 *S. tanashiensis* as described for *S. hygroscopicus* (Bottom gel) Lane 1 - 6 Thermophilic A as described for the other two strains above; Lane 7

Molecular weight marker III; Lane 8 - 13 Thermophilic B as described for the other three strains. Shuffling was carried out on the IPNS gene for all the strains and there was no change in fragment size as detected from the migration profile of the PCR products for *S. clavuligerus*, *S. tanashiensis*, *S. cattleya*, *S. hygroscopicus* and *Amycolatopsis spp* (Figure 5.3). The unchanged sizes do not necessarily mean that shuffling did not take place but may suggest an internal rearrangement which would be confirmed by sequencing the shuffled DNA gene products.

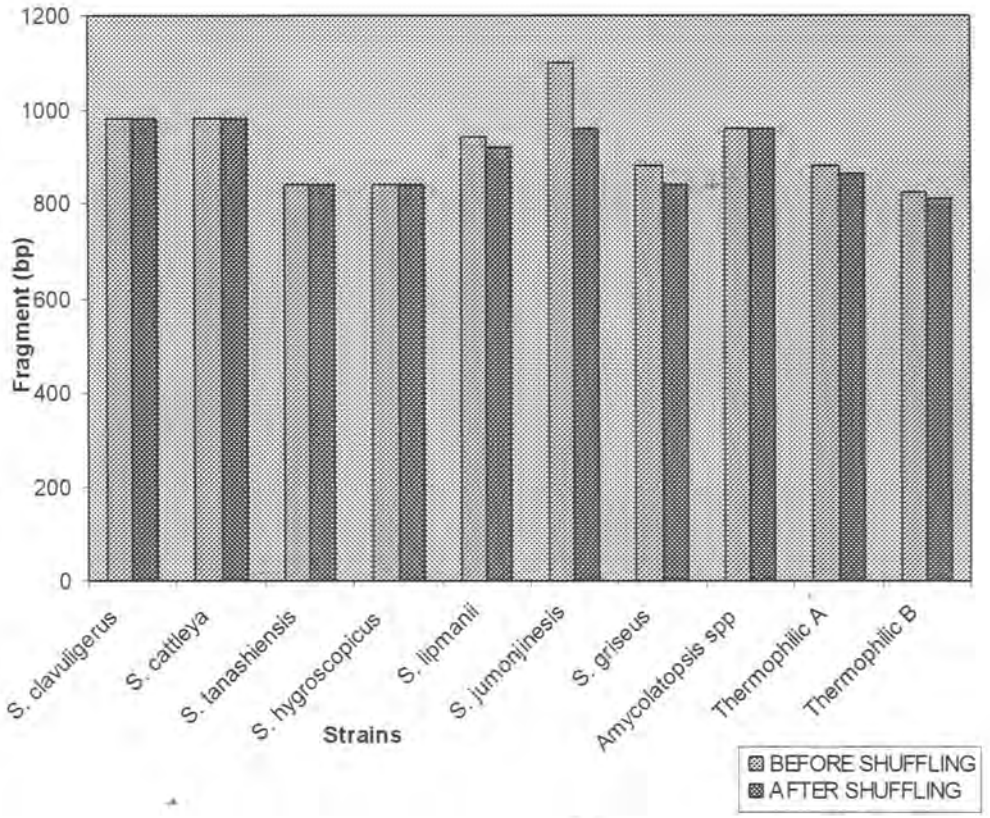


Figure 5.3. Graph showing change in base pairs of the IPNS gene as a result of shuffling.

The IPNS genes that were restrict digested with *Sau3A*, *BglII* and *BamHI* before shuffling showed differences in restriction profiles for some strains as shown (Figure 5.4). However, the shuffled IPNS genes did not yield any fragments after digesting with the three enzymes that were used before shuffling despite the noted changes in fragment size (Figure 5.3). This could be due to either that the enzymes used did not have recognition sites in the shuffled gene product or the concentration of the shuffled product was low due to gel purification in the shuffling protocol.

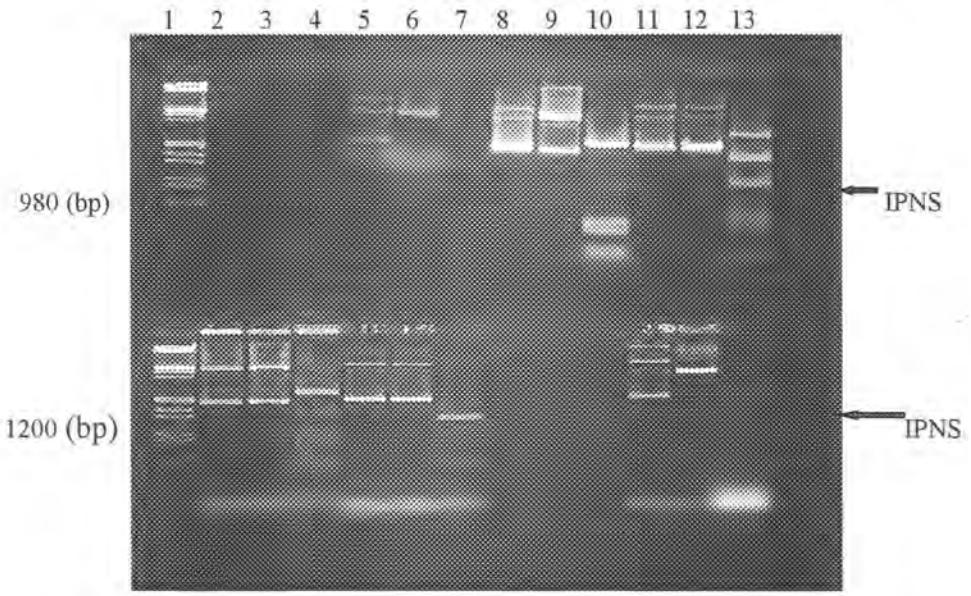


Figure 5.4. Agarose gel electrophoresis of Restriction digests using Plasmid DNA from a representative clone of all the strains used before shuffling, Lane 1 λPst I III; Lane 2 *S. hygroscopicus* (BS) *BamHI*; Lane 3 *S. hygroscopicus* (BS) *BglII*; Lane 4 *S. hygroscopicus Sau3A* (BS); Lane 5 - 7 *S. tanashiensis*; Lane 8 - 10 Thermophilic A; Lane 11- 13 Thermophilic B; Bottom gel Lane 1 λPst I ; Lane 2 - 4 *S. clavuligerus*; Lane 5 - 7 *S. griseus*; Lane 8 - 10 *S. jumonjinensis*; Lane 11 - 13 *S. lipmanii* and these were loaded as described for *S. hygroscopicus*.

The restriction digests for *Amycolatopsis spp* and *S. cattleya* before shuffling (Figure 5.5) indicate the profiles obtained significantly with *Sau3A* compared to the other two enzymes *BamHI* and *Bgl II* whereas the shuffled samples did not show digestion with any of the enzymes used instead the gels showed the position of the sample prior to digestion.

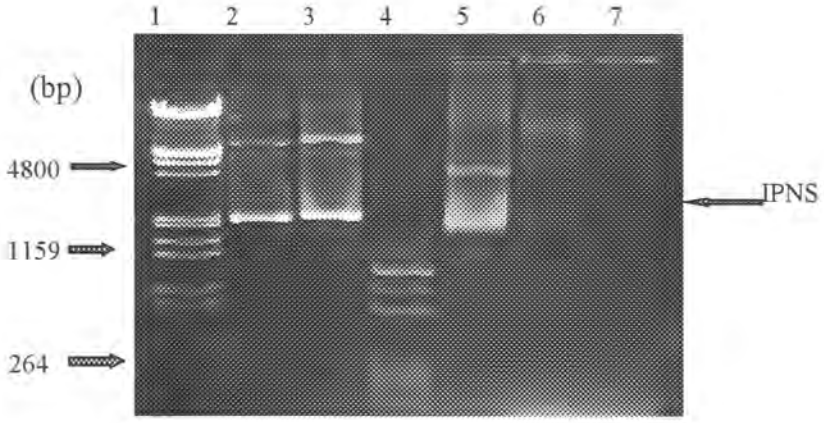


Figure 5.5. Agarose gel electrophorsis of Restriction digests from Plasmid DNA of cloned IPNS gene before being shuffled, Lane 1 λPstI; Lane 2 *Amycolatopsis spp BamHI* (BS); Lane 3 *Amycolatopsis spp BglII*; Lane 4 *Amycolatopsis spp Sau3A*; Lanes 5 - 7 *S. cattleya* loaded as per *Amycolatopsis spp.*

Figure 5.6 is an example of the IPNS activity increase noted as a result of shuffling. For all the strains used ten clones each were used for IPNS assay using the HPLC.

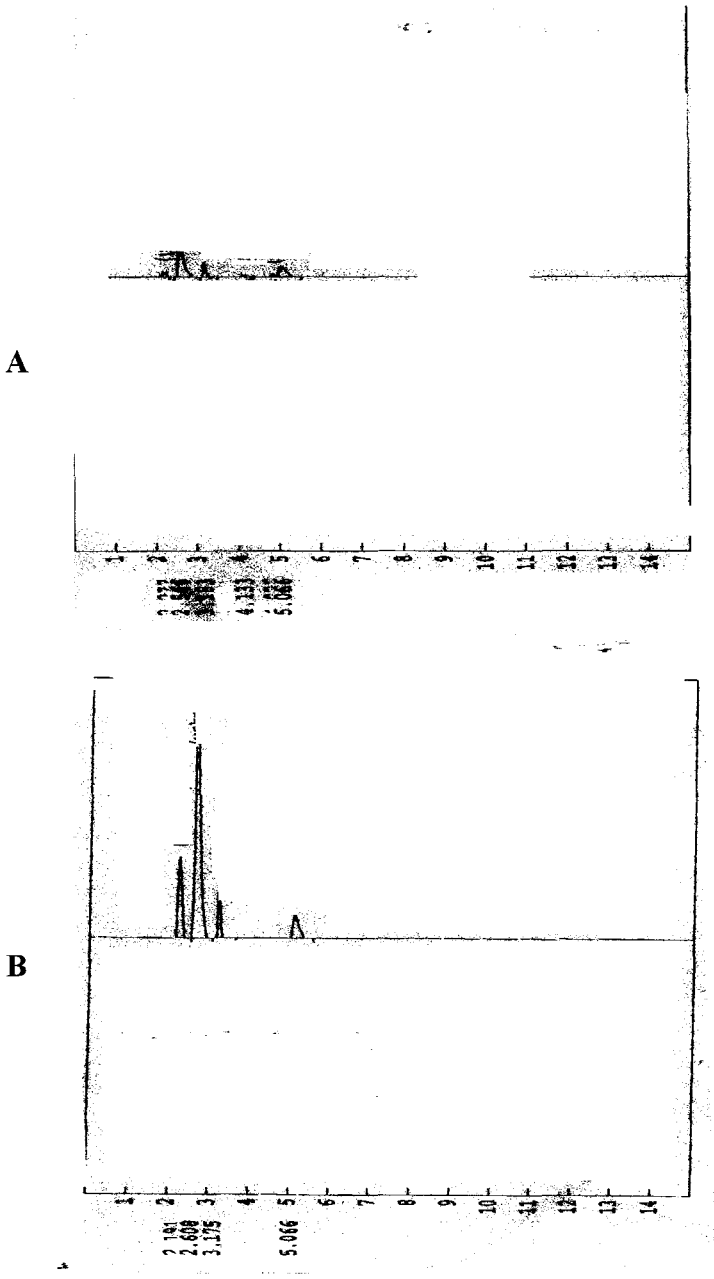


Figure 5.6. HPLC chromatogram of *S. cattleya* clones before (A) and after (B) shuffling.

The graph shows comparison of HPLC peak areas for the same retention time of 2.6 obtained from *E. coli* containing the IPNS gene cloned from each organism compared to the shuffled gene cloned in *E. coli*. IPNS expression was obtained in *E. coli* which showed (Figure 5.7) an improvement in IPNS activity. The retention time of 2.6 is used as an example to show the improved IPNS activity as a result of shuffling and successful expression in *E. coli*.

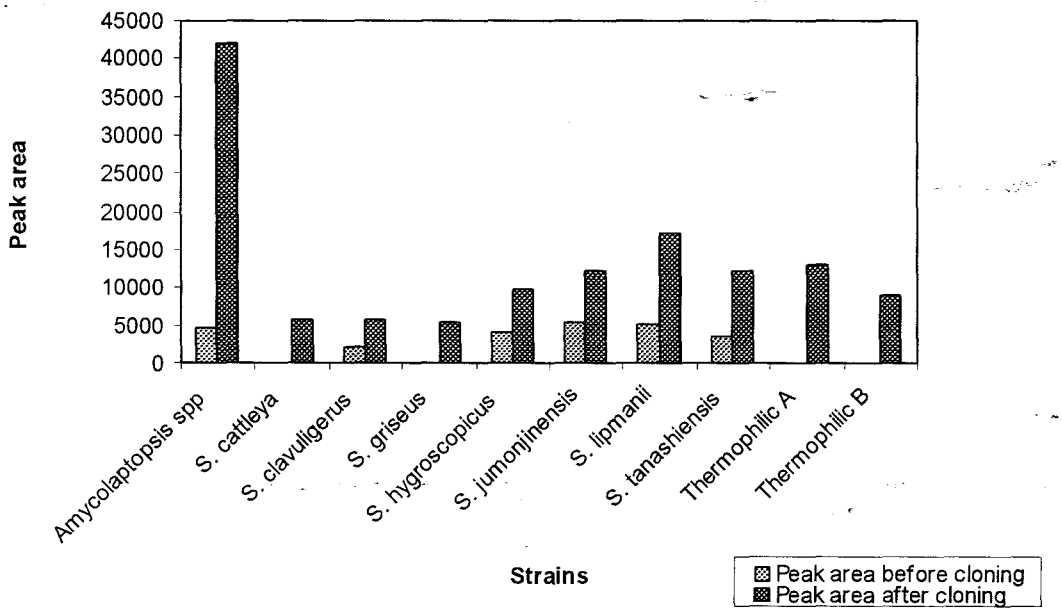


Figure 5.7. Graph of retention time of 2.6 obtained in samples before and after cloning using an *E. coli* expression system.

5.4 DISCUSSION

The use of PCR for molecular biology studies is immense seen from being able to develop conditions for gene shuffling. This technique allowed us to breakdown the IPNS gene using DNase I and reassembling the gene using PCR.

No major change in fragment size was found from the gel migration profiles (Figure 5.3) for five of the strains used. However, thermophilic A and B showed a decrease in size by a few base pairs which was seemingly essential to the functioning of the IPNS gene since the shuffling effect resulted in antibiotics being detected which were not present before shuffling when HPLC showed the presence of 6-aminopenicillanic acid and not any antibiotic. After shuffling, the HPLC results for IPNS assay showed the presence of penicillin antibiotic confirmed by the presence of the peak with similar retention time to the penicillin peak obtained from the commercial Pen VK and ampicillin antibiotic standards used. However there were problems encountered with analysing the shuffled products on the gel since there was heavy streaking of the DNA on the gel (Figure 5.2) noted which could be due to the DNA reassembling. However, the use of kodak digital science software enabled photographing of the gels as well as determining exactly where the bands were. The bands were cut out and purified using QiaQuick gel purification kit, in order to be able to clone and sequence the shuffled gene products.

The other analysis that was carried out was restriction digests on clones obtained before shuffling (Figure 5.4 and 5.5) and after shuffling using similar enzymes to see if there would be any differences in the restriction profiles. This was unsuccessful for the clones from shuffled products, probably because the concentration of the shuffled DNA was too dilute due to

purification carried out twice in the shuffling protocol, that is before the primer-less PCR and the post shuffling product, or it could be that the resulting shuffling product had been rearranged such that there were no sites for the enzymes used in the samples restricted before shuffling. The choice of enzymes used was also restricted to a few enzymes due to unavailability of enzymes that could have been used from a number of possible or suitable enzymes recommended from the information in the Vector NTI programme.

The other most reliable method by which one can confirm that shuffling is taking place is by sequencing even though this is a far more expensive way of ascertaining that shuffling is indeed taking place. Sequencing will allow detection of even one base pair loss, rearrangement or change in the position of the base pairs prior to shuffling. It would be desirable to sequence each an every potentially useful clone that is obtained after shuffling but time and cost would be the determining factor.

Genes can be switched on as a result of shuffling which was probably the case with the IPNS gene in *E. coli* from the novel thermophilic A and B since they did not yield any antibiotic before shuffling. A ten fold increase in the amount of the antibiotic produced (Figure 5.7) with *Amycolatopsis spp* confirms that shuffling improves the yield of enzymes. The other IPNS genes in *E. coli* from other strains also show a marked increase in the amount of the antibiotics produced as a result of shuffling therefore shuffled gene expression and or activity has increased as a result of shuffling.

CHAPTER 6

BIOINFORMATIC STUDIES

6.1 INTRODUCTION

Sequencing is no longer an art, it is a graduate student routine. Sequence analysis has become the cornerstone of bioinformatics because a sequence is the primary macromolecular structure data. It is the most easily obtainable structural data about a molecular function. DNA sequence allows the following conclusions to be drawn. Predicting the theoretical amino acid sequence which in turn enable secondary structure and in many cases match it with other protein sequences for which science has already obtained the elusive three-dimensional protein structure. When all the above are done, we get close to knowing the substance of the molecular functional unit (Schena 1999).

However, knowing the molecular sequence and structure of a gene/protein is not sufficient to know its real use to cellular processes. Every gene is present on its respective chromosome in every cell of an organism but is not necessarily being transcribed into an active gene product. The current assumption is that many of the genes are silent or expressed at very low levels in certain cell types. The other factor to consider is that, when a gene is expressed within a particular cell, its use is limited by whether or not other genes, that play a vital role in the same cellular process are present. If other genes, essential to the pathway are not being expressed, then the gene/protein in question will not execute its expected function (Schena 1999).

DNA sequences have received a lot of attention partly because they define the essence of what has been molecular biology. However, a few points need to be taken into account and this

include the fact that a particular gene sequence is only one instance of a gene, there is no ideal sequence existing in nature, every gene has many partial sequences associated with it, numerous intervals along a gene sequence can be defined and associated with other information and a gene sequence itself occupies an interval along a chromosomal sequence (Schena 1999).

6.1.1 BIOINFORMATIC DATABASE

The data contained in bioinformatics databases about genes and proteins is derivative. Some examples include, the protein sequence in SwissProt which is a result of translating the DNA sequence in GenBank or EMBL into an amino acid sequence using the appropriate codon translation table for that species. Sequence features are often a result of using multiple sequence alignment tools. Full-length sequences are derived from EST data through clustering algorithms and gene sequences are located within the larger genomic sequences through the use of gene finding or ORF finding algorithms (Peruski and Peruski 1997).

Computer aided analysis of nucleic acid and protein sequences enables us to scan the primary database for similar sequences in order to determine novelty, function and homology, to scan speciality databases for structural and regulatory motifs, domains and patterns to determine function, localization and family relationships, to model secondary structures to postulate two and three dimensional (2D and 3D) structure. Sequences by themselves do not yield much in the way of information, so they must be analysed by comparative methods against the existing database to deduce function and similarity (Peruski and Peruski 1997).

6.2 MATERIALS AND METHODS

6.2.1 Preparation of Samples for Sequencing

Amplified IPNS PCR products from *S. hygroscopicus*, *S. tanashiensis*, thermophilic A and thermophilic B were cloned into the vectors used in this study as previously described in section 4.2.2. The presence of the correct inserts were verified by restrict digesting the clones with Eco RI since it did not have a recognition site on the pGEM-T easy vector. Representative clones from each of the strains were inoculated into LB broth with ampicillin as described in section 4.2.4 (Plasmid DNA isolation). The DNA was quantified accurately using the (Pharmacia Biotech RNA/DNA calculator) GeneQuant to obtain the A260/280 ratio and also determined the concentration in $\mu g\ ml^{-1}$ before being sent to University of Cape Town for sequencing using M_{13} Universal pUC primers in both directions. Other representative clones from both vectors for the four strains were sequenced at Rhodes University, department of Biochemistry and Microbiology using ABI 3700 automatic sequencer. M_{13} Universal pUC primer and the sequencing primers that were supplied with the pTrcHis2-TOPO kit were used for sequencing in the forward direction only. The sequence data was analysed and aligned against documented *Streptomyces* IPNS genes using a software package Genetools (Biotools incorporated, Canada).

6.3 RESULTS

The sequence chromatogram below is an example from one of the samples sequenced at Rhodes University for *S. hygroscopicus* (Figure 6.1).

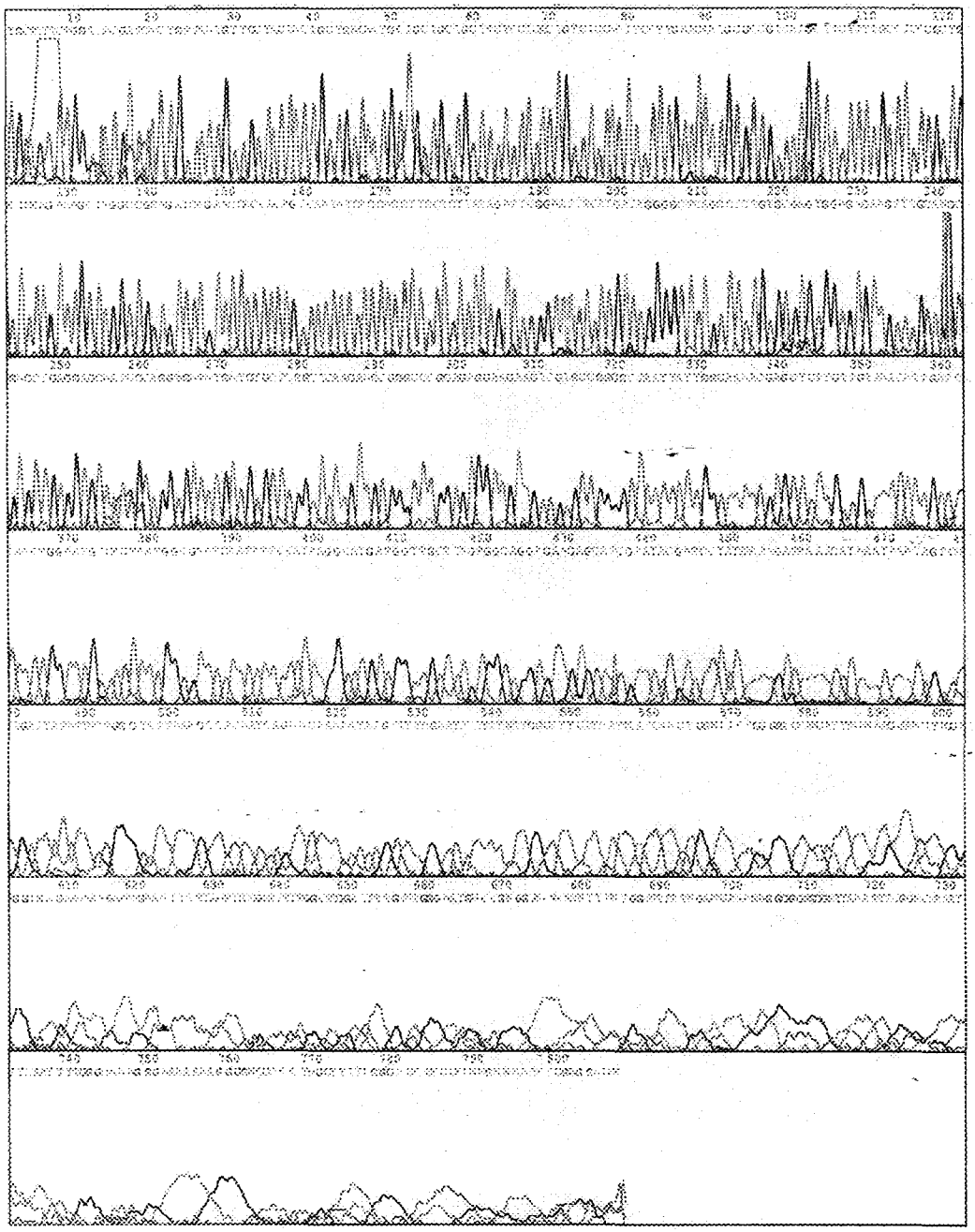


Figure 6.1. Electropherogram showing sequence obtained from *S. hygroscopicus* clone 1.

The sequence data for *S. hygroscopicus* (Figure 6.2) against other IPNS sequences obtained from the database show that the sequence was good in some parts with the T and A being the most conserved of the nucleotides. The sequence for *S. hygroscopicus* was found to be 64.6%, 62.0%, 62.2%, 62.0% identical to the IPNS sequences for *S. cattleya* (GenBank AB045856), *S. clavuligerus* (GenBank M19421), *S. hygroscopicus* subspecies *hygroscopicus* (GenBank AB045852), and *S. fimbriae* (GenBank AB045855) respectively (Figure 6.3). However, this was significantly lower than the expected 70-85% identity. The sequence alignments are shown in Figure 6.4.

TGCCGCGCGC CTGCCCAGTC CTGTTAATGT TGCTACTACT GCTGAGAGGT TGCATGCCGC
TTCTCCTCAC TGTCTCCACT TCCTTGAACA ATGCGCCGTC CGCCGCTGTC CGGACGACGC
TGCTCCAGAA AGCTAGGCCG CAGATCAGAA CCACCACAGT CAATATCACC ACCTTCCTCT
TATAGATTCG GAATCTCATG ATAGGGGCTC AGCCTCTGTG CGAGTGGAGA GAAGTTTGCA
ACGCGAGCTG AGGAGCAATG GCAGGTGATA TGATGTGCTC GGCTCAAGAA GCGGGCCCGG
AGAGGAAGAA GTCGTGCCGG GGCTAATTAT TGGCAAAACG AGCTCTTGTT GTAAACATTG
ATCCAACTGG AATGTCACTA ATGGCGAATC AATATTCCAT AAGGCATGAT GGTTGCTCAG
AGGCAGGAGA AGAGCATCGA ATACGAGAGA AGCACCCGCG GTTCCGGCCC TTCTGCGAGG
ACTACTATAG CAGCTGCTCC GGCTTTCCAC GGTGATCATG CGGGGGGCTA CGGCCTGGCG
CTGTTGGCCT TCTTCGACGA GGCACTGGGC GCTCGGGGCC AGCACGCTGT AAGAAGAACT
ATCTGGGTGA GTATAAGGAC ACGAAGTCAA TACTTTCGCG TACTTGGCTT GGACTTTGTA
CGGGAAGCTC AGAGGGGACC ATCTGGTGGG ATTGTGGAAC CGGACAGGGG GCGGGCTCGG
ACTTCGGTGC GGACTTCCTG GTTGGGAACT GGGACGTACA TGGGCCACCG TCGCTGCCGG
CCGTTCTTCC ACGGTGAAGA ACCCGACGAC A

Figure 6.2. Nucleotide sequence of IPNS gene from *S. hygroscopicus*.

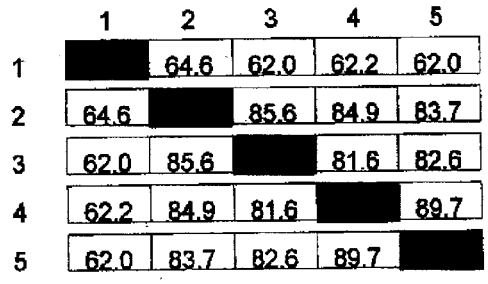


Figure 6.3. Percentage identity matrix of *S. hygroscopicus*; 1= *S. hygroscopicus*; 2= *S. cattleya*; 3= *S. clavuligerus*; 4= *S. hygroscopicus subspecie hygroscopicus*; 5= *S. fimbriae*.

```
 Consensus ATGCCAGTTCTGATGCCGTCGGCCCACGTTCCGACCATCGACATCT--CC-C----C 57
     hygro ----------------------------------------------T-GCCGC-G--C 8
  cattleya ----------------------------------------------TCGCCGCTGTCC 12
      clav ATGCCAGTTCTGATGCCGTCGGCCCACGTTCCGACCATCGACATCTCGCCGCTGTTC 57
hygro sub  ----------------------------------------------TCGCCGCTCTTC 12
  fimbriae ----------------------------------------------TCACCCCTGTTC 12

 Consensus G----C-G---C-G-C--G--A--G---C---T-C-GC---G-AG-T-------GC- 114
     hygro G---CCTG-CCCAGTCCTGTTAATGTTGCTACTACTGCTGAG-AGGTT------GCA 54
  cattleya GGGTCC-GACGCCGAC--GCCAA-GTCACGGGT-C-GCGCGGGAGATCGACGACGCG 63
      clav GGAACC-GACGCCGCC--GCGAA-GAAGCGCGT-C-GCCGAGGAGATACACGGGGCC 108
hygro sub  GGCAGC-GATCCGGAC--GCGAG-GACGCGTGT-C-GCGCAGGAGATCAACAAGGCG 63
  fimbriae GGCGAC-GACCCGGAC--GCCAA-GACACACGT-C-GCCCAGCAGATCAACAAGGCG 63

 Consensus TGCCG----TC-----T-TCT-C-AC--CC----AC-A--CG-C-T--G-CG---TC 171
     hygro TGCCGCTTCTCCTCACTGTCTCC-ACTTCCTTGAACAATGCGCCGTCCGCCGCTGTC 110
  cattleya TGCCGCGGATCCGG-CT-TCTTCTACG-CCTCGCACCA--CGGCGTC-GACG---TC 111
      clav TGCCGCGGCTCGGG-CT-TCTTCTACG-CCACGAACCA--CGGCGTG-GACG---TC 156
hygro sub  TGCCGGGGGTCCGG-TT-TCTTCTACG-CCTCCCACCA--CGGCATC-GACG---TC 111
  fimbriae TGCCGCGGCTCGGG-CT-TCTTCTACG-CCTCCCACCA--CGGCATC-GACG---TC 111

 Consensus C--------C--CTCCAG-A-G-T-GG-------T-A--AC-A-----GT---T--- 228
     hygro CGGACGACGCTGCTCCAGAAAGCTAGGCCGCAGATCAGAACCACCACAGTCAATATC 167
  cattleya C--A----GCTGCTCCAGGACG-T-GG-------TCA--ACGA-----GT---T--- 140
      clav C--A----GCAGCTCCAGGACG-T-GG-------TGA--ACGA-----GT---T--- 185
hygro sub  C--G----ACTGCTCCAGGACG-T-GG-------TCA--ACGA-----GT---T--- 140
  fimbriae C--A----GCAACTCCAGGACG-T-GG-------TCA--ACGA-----GT---T--- 140

 Consensus -CCAC---------AT-GA--C-G---C--A-GA-A------C-G-C-CTG-GCGA- 285
     hygro ACCACCTTCCTCTTATAGATTCGGAATCTCATGATAGGGGCTCAGCCTCTGTGCGAG 224
  cattleya -CCACCGGAC-C--AT-GA--CCGA--C-CAGGAGAAG--TAC-GAC-CTG-GCGA- 181
      clav -CCACGGCGC-C--AT-GA--CCGA--C-CAGGAGAAG--CAC-GAC-CTG-GCGA- 226
hygro sub  -CCACCGGAA-A--AT-GA--CCGC--C-GAGGAGAAA--CAC-GAC-CTG-GCGA- 181
  fimbriae -CCACGGGAC-C--AT-GA--CCGA--C-GAGGAGAAG--TAC-GAC-CTG-GCGA- 181

 Consensus T--A--G--GT---CA--C--------G--CAA----C---T---A--A-G-GCT- 342
     hygro TGGAGAGAAGTTTGCANGCGAGCTGAGGAGCAATN-GCAGGTGAT-ATGATGTGCTC 279
  cattleya TCCAC-GC-GTA--CAA-C-A----AGGC-CAACCCGCACATCCGCA--ACG-GCTA 224
      clav TCCAC-GC-GTA--CAA-C-C----CGGA-CAACCCGCATGTGCGCA--ACG-GCTA 269
hygro sub  TCAAC-GC-GTA--CAA-C-A----AGGA-CAACCCCCATGTGCGCA--ACG-GCTA 224
  fimbriae TCAAC-GC-GTA--CAA-C-A----GCGC-CAATCCGCGGGTGCGCA--ACG-GCTA 224

 Consensus ---T-CA-G--GC-G----GG---G-A-GAAG----GCCG----T------------ 399
     hygro GGCT-CAAGAAGCGGGCCCGGAGAGGAAGAAGTCGTGCCGGGGCTAATTATTGGCAA 335
  cattleya --CTACATG--GCGGTGAAGG---GGAAGAAG----GCCG----T------------ 254
      clav --CTACAAG--GCGGTCCCGG---GCAGGAAG----GCCG----T------------ 299
hygro sub  --TTACATG--GCCGTCAAGG---GCAAGAAG----GCCG----T------------ 254
  fimbriae --CTACATG--GCCGTCGAGG---GCAAGAAG----GCCG----T------------ 254

 Consensus ---GAG-TC-T--TG--A-C--T-A-CC--C-G---T-TC----A-G-C-A--C-AT 456
     hygro AACGAGCTCTTGTTGTAAACATTGATCCAACTGGAATGTCACTAATGGCGAATCAAT 392
  cattleya --CGAG-TCCTTCTGCTACC--TCAACC--C-GGCCT-TCACCGACGACCACCCGAT 302
      clav --CGAG-TCCTTCTGTTACC--TCAACC--C-GGACT-TCGGCGAGGACCACCCGAT 347
hygro sub  --GGAG-TCCTGGTGCTATC--TGAACC--C-GTCGT-TCACCGAGGACCACCCGAT 302
```

```
  fimbriae --CGAG-TCCTGGTGCTATC--TGAACC--C-GTCGT-TCGGCGAGGACCACCCGAT 302
 Consensus -AT-C-------C--G--G----C-C----GCA-GAG---A-C----G-----ACGA 513
     hygro -ATTCCATAAGGCATGATGGTTGCTCAGAGGCAGGAGAAGAGCATC-GAAT--ACGA 445
  cattleya GAT-C---AAGGCC-G--GTACGC-C-CCTGCACGAGGTGAACCGGTGGCCGGACGA 350
       clav GAT-C---GCCGCG-G--GGACGC-C-GATGCACGAGGTGAACCTCTGGCCCGACGA 395
 hygro sub  GAT-C---AGGTCC-G--GGGCGC-C-GCTGCACGAGGTGAACATCTGGCCGGACGA 350
  fimbriae GAT-C---CGGTCA-G--GGACAC-C-GATGCACGAGGTCAACATCTGGCCGGACGA 350

 Consensus G-A----CA---GCG-TTCCGG-C-TTCTG-GAG----TACTA--G--A--TG-T-- 570
     hygro G-AGAAGCACCCGCGG**TTCCGG**CCC**TTCTG**CGAGGA-C**TACTA**TAG-CAGCTGCTCC 499
  cattleya GGAACGCCATCCGCGGTTCCGGTCGTTCTGCGAG-AGCTACTACCGGCAGATGCTGC 406
       clav GGAGCGGCACCCGCGCTTCCGGCCGTTCTGCGAGG-GCTACTACCGGCAGATGCTGA 451
 hygro sub  GAAGCGGCACGAGCGGTTCCGGCCGTTCTGTGAGCAG-TACTACCGGGACATGTTCC 406
  fimbriae GAAGCGGCACGAGCGGTTCCGGCCGTTCTGCGAGCAG-TACTACCGGGACATGTTCC 406

 Consensus -GCT-TCCA----G-T-ATGCG-GG---T---GC-CT-GC-CT------GCC----- 627
     hygro GGCTTTCCACGGTGATC**ATGCG**GGGGGCTACGGC-CTGGCGCTGTTG--GCC----- 548
  cattleya GGCTCTCCACGGTGATCATGCGGGGG-CT--GGCGCTGGCGCTGGGGAAGCCGGAGC 460
       clav AGCTCTCCACCGTGCTCATGCGGGGG-CT--GGCGCTGGCGCTCGGGAGGCCGGAGC 505
 hygro sub  GGCTTTCCAAGGCGCTCATGCGAGGT-TT--CGCGCTTGCCCTGGGCAAGCCCGAGG 460
  fimbriae AGCTCTCCAAGACGCTGATGCGGGGC-TT--CGCGCTGGCGCTGGGCAAGCCCGAGG 460

 Consensus --TTCTTCGACG-----CT----G-G-------C-CT-----C-G---C--C--A-- 684
     hygro --**TTCTTCGACG**AGGCACT----G-GG------CGCT-----CGGGGCCAGC--A-- 583
  cattleya ACTTCTTCGACGCCGCGCTCGCCGAGGCTGACACGCTCTCCTCGGTCTC-CCTGATC 516
       clav ACTTCTTCGACGCGGCGCTCGCCGAGCAGGACTCCCTGTCGTCCGTCTC-GCTGATC 561
 hygro sub  ACTTCTTCGACGCCCCGCTCTCGGAGGCCGACACGCTGTCCGCCGTGTC-GCTGATC 516
  fimbriae ACTTCTTCGACGCGAACCTGCCGGAGGACGACACGCTGTCCGCCGTGTC-GCTCATC 516

 Consensus CGCT------A---G-AG---TA-C----GGT-A--A-A--GG-C--GA-G-C---- 741
     hygro CGCT----GTAA--GAAGAACTATCTG--GGTGAGTATA-AGGACACGAAGTCAATA 631
  cattleya CGCTACCCGTACCTGGAGGACTACCCCCCGGTCA--AGACCGGCCCGGACGGCACGA 571
       clav CGCTACCCGTATCTGGAGGAGTACCCGCCGGTGA--AGACGGGTCCCGACGGCCAGC 616
 hygro sub  CGCTACCCCTACCTGGAGGACTATCCGCCGGTCA--AGACAGGGCCCGACGGCACGA 571
  fimbriae CGCTACCCCCACCTGAAGGCCTACCCGCCGGTGA--AGACGGGGCCGGACGGCACGA 571

 Consensus --CT----TTCG-G-AC-----TGGAC-T--T-CG--GA--A--G--CT----CAG 798
     hygro --CT----**TTCG**CGTACTTGGCT**TGGAC**TTTGTACGG-GA--A--G--CT----CAG 671
  cattleya AGCTGAGCTTCGAGGACCA--CCTGGACGT-GT-CGATGATCACCGTGCTGTTCCAG 624
       clav TCCTGAGCTTCGAGGACCA--TCTGGACGT-CT-CGATGATCACCGTGCTCTTCCAG 669
 hygro sub  AGCTCAGCTTCGAGGACCA--CCTGGACGT-GT-CGATGATCACCGTCCTCTTCCAG 624
  fimbriae AGCTGAGCTTCGAGGACCA--CCTGGACGT-GT-CGGTGATCACCGTCCTGTTCCAG 624

 Consensus A----GG-G--GA-CC-TC--GGT-G-GA--GT--A--CGG-C-GG---G-C---C- 855
     hygro A----GG-G--GA-CCATCT-GGTGG-GATTGTGGAACCGGACAGG-GGG-CGGGCT 716
  cattleya ACGGAGGTGCAGAACC-TCCAGGTGGAGACGGTGGA--CGG-CTGGCGTGACCTGC- 676
       clav ACCCAGGTGCAGAACC-TCCAGGTGGAGACGGTCGA--CGG-CTGGCGGGACATCC- 721
 hygro sub  ACCGAGGTGCAGAACC-TCCAGGTCGAGACTGTGGA--CGG-CTGGCAGGACCTGC- 676
  fimbriae ACCGAGGTGCAGAACC-TCCAGGTCGAGACGGTGAA--CGG-CTGGCAGGACCTGC- 676

 Consensus C--AC-TC-G-----**GACTTCCTGGT**----**AACTG**-GG-AC-TAC-TGG---A---- 912
     hygro CGGACTTC-GGTGCGGACTTCCTGGTTGGGAACTG-GG-ACGTACATGGGCCAC--- 767
  cattleya CG-ACGTCCGGGGAGGACTTCCTGGT---GAACTGCGGCACGTACCTGGGCCACGTC 729
       clav CG-ACGTCGGAGAACGACTTCCTGGT---CAACTGCGGTACCTACATGGCGCATGTC 774
 hygro sub  CC-ACGTCCGGGGAGGACTTCCTGGT---GAACTGCGGTACCTACATGGGGTACCTC 729
  fimbriae CG-ACGTCGGGTGACGACTTCCTGGT---GAACTGCGGCACCTACATGGGGTACCTG 729
```

```
Consensus -C---CG-CT-C----CGGC-------CCAC--GGTGAAG------AAC-C-GA-CG 969
    hygro -CGT-CG-CTGC----CGGCCGTTCTTCCAC--GGTGAAG------AAC-CCGA-CG 807
  cattleya ACCAACGACTACTTCCCGGC-CCCCAACCACCGGGTGAAGTTCGTCAACGCCGAACG 785
     clav ACGAACGACTACTTCCCGGC-GCCCAACCACCGGGTGAAGTTCGTGAACGCGGAGCG 830
hygro sub  ACGAACGACTACTTCCCGGC-GCCGAACCACCGGGTGAAGTTCATCAACGCGGAGCG 785
  fimbriae ACGAACGACTACTTCCCGGC-GCCGAACGACCGGGTGAAGTTCATCAACGCGGAGCG 785
```

Figure 6.4. Nucleotide sequence alignment of *S. hygroscopicus* against *S. cattleya*; *S. clavuligerus*; *S. hygroscopicus* sub *hygroscopicus* and *S. fimbriae*.

Highly conserved regions are shown in bold in the *S. hygroscopicus* sequence data.

Open reading frame predictions for the sequence data obtained for *S. hygroscopicus* revealed two open reading frames (Figure 6.5): the first being 5' to 3' in the 3 frame on the sequenced DNA strand, and the second being 3' to 5' in the -2 frame on the complimentary DNA strand.

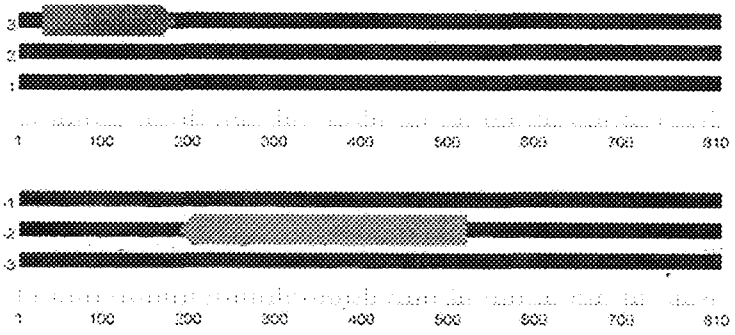


Figure 6.5. Predicted open reading frames for the sequence data obtained for *S. hygroscopicus*.

The sequence data obtained for the IPNS gene for Thermophilic A (Figure 6.6) showed 72% identity to the IPNS gene for *S. cattleya* (Figure 6.7). This finding confirms that the amplified gene product was indeed an IPNS gene present in Thermophilic A. The identity data as before (Including *S. griseus* (GenBank X54609)) obtained is shown below (Figure 6.6).

TCGCCGCTGT CCGGACCGAC GCCGACGCCA TGGAATGGGT AGCCCAGAAT
AGAACAACGA GGGGTCCGAG ATGATCCGGA TTGTTCAGAC AGCGCACCAG GCGGCGACGT
GCAGTGGCTG AGAGGACGTG GTCGGCGAGT TCCAACAAG AGTACACAAA CCAGTAGTAT
CACGAACCAG GAGAGTACGA CGGCAGAAGG GGAAGAAGGT AGTTGAAGAT
TCTGACTCAG GGAGCCGGTC GACCAGCCAG GCATGATCAA GACATACGCC GTGACGAGTG
AACCAGTTAG GTCTAGTCAG TTAACGCCGG AGCGCCCCAG CAAGAGAACT GATAGGCTCT
GTGTTCGACG ATGGCCGTGT GCAGGTATGA GGAGCGTAAG TGAACGAGAC TTCCGGCAGA
TGCTGCGGCT CTCCACGGTG ATCATGCGGG GCTGGCGCTC GCGCTGGGGA AGATAGAGCT
ACTTCTTTCG
ATTCCGCATG AGCAAGATCG AGGCTGACAC TCTGTCTCAA CGTGAGCCCT GTGGCCACTG
ATAGTCGTCC GAAGCGATGA CCTTCTTACG ACACGGGCTA CACGCCGACC CGGTACCGGA
GGACTACCCC CGGTCAAGAG GCGACCAGGA CTGTACTCAG CACCCGGTAG TGCAGTCGCC
GTGAGGATCT CGGCCTAGAG GTGATCCGTC TGGATATAGT AGTCCGGTCA GGTAACAAAG
CGTGTGCACA GTAAAGCATC CGGTGGAGAT CACAGGTGTA CCGAGCAAAT GTACAACCAG
AAGCCGGATG AACTTACATA AGCTATCTAG GCGCGCTCGG GATGGTATAG AACGGATCGG
GTCAGAAGAG GAGCGACAGA GTAAAGGGCC GCGCATCCCC GCAGTGCCAG
AAGCGCGAGG GATCGCGAGT ACGAAGAGAC GAACGAAGTA CGACTAGAAT
CGTTCTCACC GCCCGTTCAA CCAGCGACCC GCGAATGAAC AAAGGCGACG TGCCGTACCC
AGTCCGTCTA GCGCGACCGC CAAGAGCCGA GGCCCATTAC GCCACACTAG CCGCGTGCCT
CGCTGGGCCG ACCTGTAGCA CGATCTCGCG GCGGACACGT CCACTGGGCN CAGNATTGTG
ACGAGCCGTA GCGCACGAGC CAAGCGAGAG GTGACGATAC TACTAGCACG
ACGGAGGCTC CAGCTGAGAG AGCACAACAG TCTACACGCG TCAGATACTG GCGTTCATAC
AATAGTGCAT ACTCGCAGAC AGGACAAGCA ACCAGACGCA GCGAGCCTCC

TCAGATGGAG ACGAGCCGGG ACTGGGAGCC AGGACAGGCA CCAGCCAGTC

GGCAGCTCAG ACCCCGGCTG CGGCCACGTT CCGATCACCA TGCGAGCCCG ACCATCAGAG

AGCGACAGAG CTCGCCTAGT CCCTCAGAAG GAAGGCGTAA TCACTCGAGA GACGTAGGAG

CGCGACGGCG CCGGCCGCGA ACCGG

Figure 6.6. Nucleotide sequence of IPNS gene from Thermophilic A.

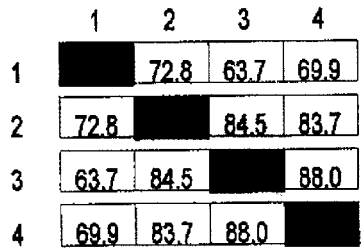


Figure 6.7. Percentage identity matrix for Thermophilic A, 1= Thermophilic A; 2= *S. cattleya*; 3= *S. Griseus*; 4= *S. fimbriae*.

```
Consensus    ATGCCCATTCCCATGCTGCCGGCACACGTGCCGACGATCGACATCTC-CC-CTGT-C 57
     Therm A ----------------------------------------------TCGCCGCTGTCC 12
    cattleya ----------------------------------------------TCGCCGCTGTCC 12
     griseus ATGCCCATTCCCATGCTGCCGGCACACGTGCCGACGATCGACATCTCACCGCTGTCC 57
    fimbriae ----------------------------------------------TCACCCCTGTTC 12

   Consensus GG---CGAC-C-GACG-CA-G------GT-GC-C-G-A--GA-C-AC-A-G-GT-CC 114
     Therm A GGA-CCGACGCCGACGCCATGGAATGGGTAGCCCAGAATAGAACAACGAGGGGT-CC 67
    cattleya GGGTCCGACGCCGACGCCAAGTCACGGGTCGCGCGGGA--GATCCACGACGCGTGCC 67
     griseus GGCGGCGACGCGGACGACAAGAAGCGCGTCGCACAGGA--GATCAACAAGGCGTGCC 112
    fimbriae GGCGACGACCCGGACGCCAAGACACACGTCGCCCAGCA--GATCAACAAGGCGTGCC 67

   Consensus G-G----TC-GG-TT-TTC-A--C--C-CACCA-GGC--CGACGT-CAG---CT--- 171
     Therm A GAGATGATCCGGATTGTTC-AGACAGCGCACCA-GGCGGCGACGTGCAGTGGCTGAG 122
    cattleya GCG--GATCCGGCTTCTTCTACGCCTCGCACCACGGCGTCGACGTCCAGCTGCTCC- 121
     griseus GGG--AGTCGGGGTTCTTCTACGCCTCCCACCACGGCATCGACGTCCAGCTGCTGA- 166
    fimbriae GCG--GCTCGGGCTTCTTCTACGCCTCCCACCACGGCATCGACGTCCAGCAACTCC- 121

   Consensus AGGACGT-GTC--CGAGTTCCA-C--GA--A--AC--AC-AG-AG-A--ACGA-CC- 228
     Therm A AGGACGTGGTCGGCGAGTTCCAACAAGAGTAC-ACAAACCAGTAGTATCACGAACCA 178
    cattleya AGGACGTGGTCAACGAGTTCCA-CCGGACCATGACCGACCAGGAGAAGTACGA-CCT 176
     griseus AGGACGTCGTCAACGAGTTCCA-CCGGACCATGACGGACGAGGAGAAGTACGA-CCT 221
    fimbriae AGGACGTGGTCAACGAGTTCCA-CGGGACCATGACCGACGAGGAGAAGTACGA-CCT 176

   Consensus  GG-GA-T-C-ACG---A-AA-------AA---G-------G-AA-G---A-T-C-TG 285
     Therm A GGAGAGTACGACG-GCAGAAGGGGAAGAAG--GTAGTT--G-AA-G---ATT-C-TG 223
    cattleya  GGCGA-T-CCACGCGTACAACAAGGCCAACCCGCACATCCGCAACGGCTACTACATG 231
     griseus  GGCGA-T-CAACGCTTACAACAAGAACAACCCGCGCACCCGCAACGGCTACTACATG 276
    fimbriae  GGCGA-T-CAACGCGTACAACAGCGCCAATCCGCGGGTGCGCAACGGCTACTACATG 231

   Consensus -C--T--AGGG--AG--GG-CG-C-AG-CC-----TG-T-----A-A--CG-C-T-- 342
     Therm A AC--TC-AGGG--AGCCGGTCGACCAG-CCAGGCATGAT-CAAGACATACGCCGT-- 271
    cattleya GCGGTGAAGGGGAAGAAGGCCGTCGAGTCCTTC--TGCTACCTCA-ACCCGGCCTTC 285
     griseus GCCGTCAAGGGCAAGAAGGCCGTCGAGTCCTGG--TGCTACCTGA-ACCCGTCGTTC 330
    fimbriae GCCGTCGAGGGCAAGAAGGCCGTCGAGTCCTGG--TGCTATCTGA-ACCCGTCGTTC 285

   Consensus --CGA--GA--ACC-G---G-TC--G-C-G--A-C-CC---GC-C----G-A-G--A 399
     Therm A GACGAGTGA--ACCAGTTAGGTCTAGTCAGTTAACGCCGGAGCGCCCCAGCAAGAGA 326
    cattleya ACCGAC-GACCACCCGAT-GATCAAGGCCGGTA-CGCCCCTGCAC----G-AGGTGA 334
     griseus AGCGAG-GACCACCCGCA-GATCAGGTCCGGGA-CGCCGATGCAC----G-AAGGTA 379
    fimbriae GGCGAG-GACCACCCGAT-GATCCGGTCAGGGA-CACCGATGCAC----G-AGGTCA 334

   Consensus AC---T-GGC-C-G-G--CGA--A--G-C----GC-G-T-T--GG--C-T---G-GA 456
     Therm A ACTGATAGGCTCTGTGTTCGACGATGGCCGTGTGCAGGTATGAGGAGCGTAA-GTGA 382
    cattleya ACCGGT-GGC-C-G-GA-CGAGGAACGCCATCCGC-GGT-TCCGGT-CGTTCTGCGA 383
     griseus ACATCT-GGC-C-G-GA-CGAGAAGCGCCACCAGC-GCT-TCCGGC-CCTTCTGCGA 428
    fimbriae ACATCT-GGC-C-G-GA-CGAGAAGCGGCACGAGC-GGT-TCCGGC-CGTTCTGCGA 383

   Consensus ---AG--ACT-CCG--A--T--T---GCTCTCCA-G----T-ATGCG-GG--T-GCG 513
     Therm A ACGAG--ACTTCCGGCAGATGCTGCGGCTCTCCACGGTGATCATGCGGGGC-TGGCG 436
    cattleya G--AGCTACTACCGGCAGATGCTGCGGCTCTCCACGGTGATCATGCGGGGGCTGGCG 438
     griseus GC-AG-TACTACCGCGACGTCTTCTCGCTCTCCAAGGTACTGATGCGAGGTTTCGCG 483
    fimbriae GC-AG-TACTACCGGGACATGTTCCAGCTCTCCAAGACGCTGATGCGGGGCTTCGCG 438
```

```
Consensus CT-GC-CTGGG-AAG---GAG--ACTTCTT-CGA---CGC-----C--------GG- 570
  Therm A CTCGCGCTGGGGAAGATAGAGCTACTTCTTTCGATTCCGCATGAGCAAGATCGAGGC 493
 cattleya CTGGCGCTGGGGAAGCCGGAGC-ACTTCTT-CGA---CGC-CGCGCTCGC-CGAGGC 488
  griseus CTGGCCCTGGGCAAGCCCGAGG-ACTTCTT-CGA---CGC-TTCACTCTC-ACTGGC 533
 fimbriae CTGGCGCTGGGCAAGCCCGAGG-ACTTCTT-CGA---CGC-GAACCTGCC-GGAGGA 488

Consensus -GACAC-CT-TC-C--------C---GT--C--CT-AT------CC---------- 627
  Therm A TGACACTCTGTCTCAACGTGAGCCCTGTGGCCACTGATAGTCGTCCGAAGCGATGAC 550
 cattleya TGACACGCTCTC-C-------TCG--GTCTCC-CTGAT------CCG---------- 518
  griseus CGACACACTGTC-C-------GCC--GTGACG-CTGAT------CCA---------- 563
 fimbriae CGACACGCTGTC-C-------GCC--GTGTCG-CTCAT------CCG---------- 518

Consensus CT----------------------ACCC---ACC-G-AGG-CTACCC-CCGGT-A 684
  Therm A CTTCTTACGACACGGGCTACACGCCGACCCGGTACC-GGAGGACTACCC-CCGGTCA 605
 cattleya CT----------------------ACCCG-TACCTGGAGGACTACCCCCCGGTCA 550
  griseus CT----------------------ACCCG-TACCTGGAGGACTACCCGCCGGTCA 595
 fimbriae CT----------------------ACCCC-CACCTGAAGGCCTACCCGCCGGTGA 550

Consensus AGA---G-CC-GGAC-G------GCAC--G--A----AG-C----TGAG---CT--- 741
  Therm A AGAGGCGACCAGGACTGTACTCAGCACCCGGTAGTGCAGTCGCCGTGAGGATCTCGG 662
 cattleya AGACCGGCCC-GGAC-G------GCAC--G--A----AG-C----TGAG---CT--- 580
  griseus AGACAGGCCC-GGAC-G------GCAC--G--A----AG-C----TGAG---CT--- 625
 fimbriae AGACGGGGCC-GGAC-G------GCAC--G--A----AG-C----TGAG---CT--- 580

Consensus --T-GAGG--A-CC--CTGGA----GT-GTC-G-T--G--A--------T---CAC- 798
  Therm A CCTAGAGGTGATCCGTCTGGATATAGTAGTCCGGTCAGGTAACAAAGCGTGTGCACA 719
 cattleya --TCGAGG--A-CCACCTGGAC---GT-GTC-GAT--G--A--------T---CACC 612
  griseus --TCGAGG--A-CCACCTGGAC---GT-GTC-GAT--G--A--------T---CACC 657
 fimbriae --TCGAGG--A-CCACCTGGAC---GT-GTC-GGT--G--A--------T---CACC 612

Consensus GT--------C-GT-----TC-CAG----AC-GAG-----GT-CA----GAA-CC-- 855
  Therm A GTAAAGCATCCGGTGGAGATCACAGGTGTACCGAGCAAATGTACAACCAGAAGCCGG 776
 cattleya GT-------GCTGT-----TC-CAG----ACGGAG-----GTGCA----GAA-CC-- 640
  griseus GT-------CCTGT-----TC-CAG----ACCGAG-----GTACA----GAA-CC-- 685
 fimbriae GT-------CCTGT-----TC-CAG----ACCGAG-----GTGCA----GAA-CC-- 640

Consensus ------------------TC-AGG-----T-G-GA-GG------ACGG-T-GG--C- 912
  Therm A ATGAACTTACATAAGCTATCTAGGCGCGCTCGGGATGGTATAGAACGGATCGGGTCA 833
 cattleya ------------------TCCAGG-----TGGAGACGGT---GGACGGCT-GG--CG 668
  griseus ------------------TCCAGG-----TCGAGACGGC---CGACGGAT-GG--CA 713
 fimbriae ------------------TCCAGG-----TCGAGACGGT---GAACGGCT-GG--CA 668

Consensus ---GA---GC-GAC---GT---G-G-----C-TCC--G--GTG----AA--GCG-G- 969
  Therm A GAAGAGGAGC-GACAGAGTAAAGGGCCGCGCATCCCCGCAGTGCCAGAAGCGCGAGG 889
 cattleya T--GACCTGCCGAC---GTCCGGGGAGGA-CTTCCT-G--GTG----AACTGCG-G- 710
  griseus G--GACCTGCCGAC---GTCCGGGGAGAA-CTTCCT-G--GTG----AACTGCG-G- 755
fimbriatus G--GACCTGCCGAC---GTCGGGTGACGA-CTTCCT-G--GTG----AACTGCG-G- 710

Consensus ---C-C---TAC---G-G--AC---AC-AA---CGACTA----C-TTC-C-C-GC-C 1026
  Therm A GATCGCGAGTACGAAGAG--ACGA-ACGAAGTACGACTAGAATCGTTCTCACCGCCC 943
 cattleya ---CACG--TAC-CTGGGCCACGTCACCAA---CGACTA----C-TTC-C-CGGCCC 751
  griseus ---CACC--TAC-ATGGGGTACCTCACCAA---CGACTA----C-TTC-C-CGGCCC 796
fimbriatus ---CACC--TAC-ATGGGGTACCTGACGAA---CGACTA----C-TTC-C-CGGCGC 751
```

```
Consensus ----AACCA-CG----G-GAA--T---CAA-G--GA-CG--C-GT-CCC-G-C--TC 1083
  Therm A GTTCAACCAGCGACCCGCGAA--TGAACAAAGGCGA-CGTGCCGTACCCAGTCCGTC 997
 cattleya C--CAACCACCG---GGTGAAGTTCGTCAACGCCGAACG-G--------------- 786
  griseus C--CAACCACCG---CGTGAAGTTCATCAACGCGGAGCG-GCTGT-CCCTGCCGTTC 846
 fimbriae C--GAACCACCG---GGTGAAGTTCATCAACGCGGAGCG-C--------------- 786

Consensus T--C-C-AC--C----A--C-A-----AT---GCC---C-AGCC-CG-G----C-CG 1140
  Therm A TAGCGCGACCGCCAAGAGCCGAGGCCCATTACGCCACACTAGCCGCGTGC---CTCG 1051
 cattleya ---------------------------------------------------------- 786
  griseus TTCCTCCACGCCGGGCACACCACCGTGATGGAGCCGTTC-AGCCCCGAGGACACCCG 902
 fimbriae ---------------------------------------------------------- 786

Consensus C-GG---GA-CTG-A-C---CG-TC-CGC--CGG--AC--CC-C--G---CAG--TT 1197
  Therm A CTGGGCCGACCTGTAGCA--CGATCTCGCGGCGGACACGTCCACTGGGCNCAGNATT 1106
 cattleya ---------------------------------------------------------- 786
  griseus CGGGAAGGAGCTGAACCCGCCGGTC-CGCTACGGCGACTACCTCCAG---CAGGCTT 955
 fimbriae ---------------------------------------------------------- 786

Consensus ---ACG--C-G-A-CGC---AG----GC-AGA--TGACGATACTACTAGCACGACGG 1254
  Therm A GTGACGAGCCGTAGCGCACGAGCCAAGCGAGAGGTGACGATACTACTAGCACGACGG 1163
 cattleya ---------------------------------------------------------- 786
  griseus CCAACGCCCTG-ATCGCGA-AGAACGGCCAGACCTGA-------------------- 990
 fimbriae ---------------------------------------------------------- 786

Consensus AGGCTCCAGCTGAGAGAGCACAACAGTCTACACGCGTCAGATACTGGCGTTCATACA 1311
  Therm A AGGCTCCAGCTGAGAGAGCACAACAGTCTACACGCGTCAGATACTGGCGTTCATACA 1220
 cattleya ---------------------------------------------------------- 786
  griseus ---------------------------------------------------------- 990
 fimbriae ---------------------------------------------------------- 786
```

Figure 6.8. Nucleotide alignment of Thermophilic A showing a consensus against *S. cattleya*; *S. griseus* and *S. fimbriae*.

Highly conserved regions are shown in bold in the Thermophilic A sequence data.

Open reading frame predictions for the sequence data obtained for Thermophilic A revealed three open reading frames (Figure 6.9): the first being 5' to 3' in the 3 frame on the sequenced DNA strand, and the second and third being 3' to 5' in the -1 frame on the complimentary DNA strand.

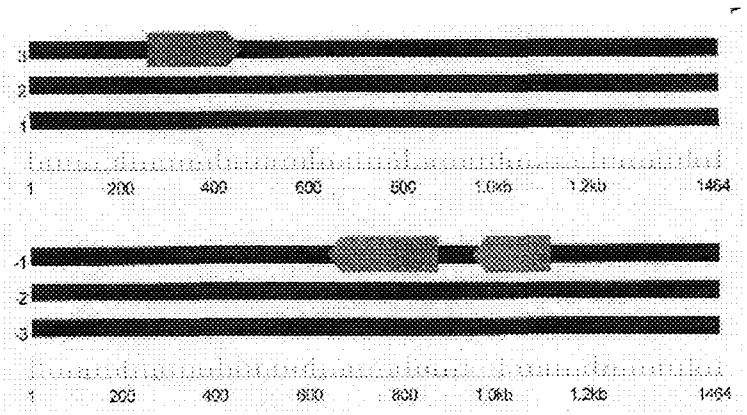


Figure 6.9. Predicted open reading frames for the sequence data obtained for Thermophilic A.

Numerous attempts were made to obtain a sequence for the amplified IPNS gene from Thermophilic B which would give a typical *Streptomyces* IPNS gene sequence, however many attempts resulted in chromatograms with poor detection peaks. No predicted open reading frames could be found for Thermophilic B sequence data (Figure 6.10). The following identity matrix (Figure 6.11) was however obtained from sequence comparisons (Figure 6.12) with the best data obtained for Thermophilic B.

ATGCCAGTTC TGATGCCGTC GGGCCGTCTG CCGACATCGG CTATCATCTA CTGTTAAGGT

TACTACTACT GTGTGACAAT GCTGCTGCTG CCTCTACCTG CACTGCCTCC ACTTGCTTGA

ACAATGCGCC GACCACGCTG CTTTGGCCTG CCGCTGCATA CAGTAGCGCT GGGCCGCAGA

TCAGAACCAC CACAGTTCTG TTACCTCAAC CCCGGACTTC GATGCGGAAC CTGATGATAC

GGCCGCGGGG ACTGGGCGAG TGGAGAGATG GCCGCAGGCG AGGCCGTCGA GTCCCTTCTG

TGATATTGAT GGGCTCTAGC TCAAGAAGCG GGCCCTGAGA GGAAGAATTC GAGCCGCGGC

TAATTATTGG CCGAACGAGC TCAGTGTCGA GAACAATGAC CCGACTGACA TGGCTCTCGT

GGAGAATCGA TATTGCATCA AGCGTGAGGC TTGCTCAGAG GGCATGAGAC GAGCGCGCGG

TACTATTCCT ATCGGAGCAT TCTTCGATAA GAGAAGGCCG CTCGGAGTTA TACCCTGTCG

TCCGTAAAGA CGGCCAGCTC CTGAGCTTCG AGGACCATCT GGAGGGAAGC TGATTACTCT

GGTCCTTCGG TTGCGTTTAG CCGAACGTTG CTTCAGCGGG CTCTTCCCAG ATTGAGGACT

GAACCTGCGG CGGGGCTACG GGGAGCTTGA TCGACGGCTG GCGGGACATC CCCGAGTTCT

TGGCAGACGA CTTCCTGGTC AACAACCGGC GCCTCGAAAA CTGTTCCTAC GAGACTTCCC

GGCGCCCAAC CACCGGGTGA ATTCGTTACG CGAGTGTCAA TTAATGCCGG AGGGCGCGAG

CGAGGAGGTG AGGAACGAAC AAGAATGGAG GCCAGACCGG TTGGAAAGGG ACCTGA

Figure 6.10. Nucleotide sequence of IPNS gene from Thermophilic B.

	1	2	3	4	5	6
1		64.6	62.1	63.1	62.8	62.3
2	64.6		85.4	82.7	79.2	81.6
3	62.1	85.4		83.8	84.9	85.0
4	63.1	82.7	83.8		88.3	89.7
5	62.8	79.2	84.9	88.3		89.1
6	62.3	81.6	85.0	89.7	89.1	

Figure 6.11. Percentage identity matrix of Thermophilic B against 2= *S. clavuligerus*; 3= *S. cattleya*; 4= *S. fimbriae*; 5= *S. griseus* and 6= *S. hygroscopicus* sub *hygroscopicus*.

```
 Consensus ATGCC--TTC--ATGC-G-CGG--C---T-CCGAC-ATCG-C-T-TC--C--CT-T- 57
   Therm B ATGCCAGTTCTGATGCCGTCGGGCCGTCTGCCGAC-ATCGGC-TATCATCTACTGTT 55
      clav ATGCCAGTTCTGATGCCGTCGGCCCACGTTCCGACCATCGACATCTCGCCG-CTGTT 56
  cattleya ----------------------------------------------TCGCCG-CTGTC 11
  fimbriae ----------------------------------------------TCACCC-CTGTT 11
   griseus ATGCCCATTCCCATGCTGCCGGCACACGTGCCGACGATCGACATCTCACCG-CTGTC 56
Hygroscopi ----------------------------------------------TCGCCG-CTCTT 11

 Consensus --GG---C-A---C-G---G--A--G---C----T-GC--------T--AC---GC- 114
   Therm B AAGGTTACTACTACTGTGTGACAATGCTGCTGC-T-GC-C------TCTACCT-GCA 102
      clav C-GGAACCGACG-CCGCC-GCGAA-GAAGC-GCGTCGCCGAGGAGATACACGGGGC- 107
  cattleya C-GGGTCCGACG-CCGAC-GCCAA-GTCAC-GGGTCGCGCGGGAGATCCACGACGC- 62
  fimbriae C-GGCGACGACC-CGGAC-GCCAA-GACAC-ACGTCGCCCAGCAGATCAACAAGGC- 62
   griseus C-GGCGGCGACG-CGGAC-GACAA-GAAGC-GCGTCGCACAGGAGATCAACAAGGC- 107
Hygroscopi C-GGCAGCGATC-CGGAC-GCGAG-GACGC-GTGTCGCGCAGGAGATCAACAAGGC- 62

 Consensus -TGCC-----TC----TT-CTT--AC---GC--C--ACCACG--GC-T-G-C-T-CC 171
   Therm B C**TGCC**-----TCCA-CTTGCTTGAACAATGCG-CCG**ACCACG**CTGCTTTGGCCTGCC 152
      clav CTGCCGCGGCTCGGGCTT-CTTCTAC---GCCACGAACCACG--GCGTGGACGT-CC 157
  cattleya GTGCCGCGGATCCGGCTT-CTTCTAC---GCCTCGCACCACG--GCGTCGACGT-CC 112
  fimbriae GTGCCGCGGCTCGGGCTT-CTTCTAC---GCCTCCCACCACG--GCATCGACGT-CC 112
   griseus GTGCCGGGAGTCGGGGTT-CTTCTAC---GCCTCCCACCACG--GCATCGACGT-CC 157
Hygroscopi GTGCCGGGGGTCCGGTTT-CTTCTAC---GCCTCCCACCACG--GCATCGACGT-CC 112

 Consensus --C--C-T--AG-A-CG-T-G----CG-AG-TC-A--G---A---C---C-AG--- 228
   Therm B -GCTGCATACAGTAGCGCTGGGC--CGCAGATC-A--GAACCA---CC-AC-AGTTC 198
      clav AGCAGC-TCCAGGA-CG-TGGTGAACG-AGTTCCACGGCGCCATGACCGACCAGGA- 209
  cattleya AGCTGC-TCCAGGA-CG-TGGTCAACG-AGTTCCACCGGACCATGACCGACCAGGA- 164
  fimbriae AGCAAC-TCCAGGA-CG-TGGTCAACG-AGTTCCACGGGACCATGACCGACGAGGA- 164
   griseus AGCTGC-TGAAGGA-CG-TCGTCAACG-AGTTCCACCGGACCATGACGGACGAGGA- 209
Hygroscopi GACTGC-TCCAGGA-CG-TGGTCAACG-AGTTCCACCGGAAAATGACCGCCGAGGA- 164

 Consensus -G--A---C-ACC--G-GA--TC-A-GC--A--AC------A-A---CC-C------ 285
   Therm B TGTTACCTCAACCCCG-GACTTCGATGCGGA--ACCTGATGATACGGCCGCGGG--- 249
      clav -GA-AGCACGACCTGGCGA--TCCACGCGTACAACCCGGACA-AC--CCGCATGTGC 259
  cattleya -GA-AGTACGACCTGGCGA--TCCACGCGTACAACAAGGCCA-AC--CCGCACATCC 214
  fimbriae -GA-AGTACGACCTGGCGA--TCAACGCGTACAACAGCGCCA-AT--CCGCGGGTGC 214
   griseus -GA-AGTACGACCTGGCGA--TCAACGCTTACAACAAGAACA-AC--CCGCGCACCC 259
Hygroscopi -GA-AACACGACCTGGCGA--TCAACGCGTACAACAAGGACA-AC--CCCCATGTGC 214

 Consensus G--AC--GGC-A-T----A-A-GGC-G-----GG---GA-GGCCGT-GAGTCC-T-- 342
   Therm B G--ACTGGGCGAGTGGAGAGATGGCCG-CA--GGC--GA-**GGCCGTCGAGTCC**CTTC 298
      clav GCAAC--GGCTACT----ACAAGGCGGTCCCGGGCAGGAAGGCCGTCGAGTCC-TTC 309
  cattleya GCAAC--GGCTACT----ACATGGCGGTGAAGGGGAAGAAGGCCGTCGAGTCC-TTC 264
  fimbriae GCAAC--GGCTACT----ACATGGCCGTCGAGGGCAAGAAGGCCGTCGAGTCC-TGG 264
   griseus GCAAC--GGCTACT----ACATGGCCGTCAAGGGCAAGAAGGCCGTCGAGTCC-TGG 309
Hygroscopi GCAAC--GGCTATT----ACATGGCCGTCAAGGGCAAGAAGGCCGTGGAGTCC-TGG 264

 Consensus TG---TA--T-A----C-C-----TC---GA-G-C---CCC---GA-----G---G- 399
   Therm B TGTGATAT-TGATGGGCTC-TAGCTCAA-GAAG-CGGGCCC--TGA--GAGGA-AG- 345
      clav TGT--TACCTCAA---CCCGGACTTCGGCGAGGACCA-CCCGATGATCGCCG-CGGG 359
  cattleya TGC--TACCTCAA---CCCGGCCTTCACCGACGACCA-CCCGATGATCAAGG-CCGG 314
  fimbriae TGC--TATCTGAA---CCCGTCGTTCGGCGAGGACCA-CCCGATGATC-CGGTCAGG 314
   griseus TGC--TACCTGAA---CCCGTCGTTCAGCGAGGACCA-CCCGCAGATC-AGGTCCGG 359
Hygroscopi TGC--TATCTGAA---CCCGTCGTTCACCGAGGACCA-CCCGATGATC-AGGTCCGG 314
```

```
Consensus -----C---GC-CG--GG-T-AA-----TGGCC--ACGAG---A----CG---A--- 456
  Therm B -AATTCGA-GC-CGC-GGCT-AATTAT-TGGCCGAACGAGCTCAGTGTCGAGAACAA 396
     clav GACGCCGATGCACGA-GG-TGAAC-CTCTGGCCCGACGAGG--AG---CG-GCACCC 407
 cattleya TACGCCCCTGCACGA-GG-TGAAC-CGGTGGCCGGACGAGG--AA---CG-CCATCC 362
 fimbriae GACACCGATGCACGA-GG-TCAAC-ATCTGGCCGGACGAGA--AG---CG-GCACGA 362
  griseus GACGCCGATGCACGAAGG-T-AAC-ATCTGGCCGGACGAGA--AG---CG-CCACCA 407
Hygroscopi GGCGCCGCTGCACGA-GG-TGAAC-ATCTGGCCGGACGAGA--AG---CG-GCACGA 362

Consensus -G---CG--T--C--GG---T-TC-TG-GAG-----A-TA--G--A-------T--- 513
  Therm B TGACCCGACTGACATGGC--TCTCGTG-GAGAATCGA-TATTG-CATCAAGCGTGAG 448
     clav -G---CG-CTTCC--GGCCGT-TC-TGCGAGGGC-TACTACCGGCA-GATGC-TGAA 452
 cattleya -G---CG-GTTCC--GGTCGT-TC-TGCGAG-AGCTACTACCGGCA-GATGC-TGCG 407
 fimbriae -G---CG-GTTCC--GGCCGT-TC-TGCGAGCAG-TACTACCGGGA-CATGT-TCCA 407
  griseus -G---CG-CTTCC--GGCCCT-TC-TGCGAGCAG-TACTACCGCGA-CGTCT-TCTC 452
Hygroscopi -G---CG-GTTCC--GGCCGT-TC-TGTGAGCAG-TACTACCGGGA-CATGT-TCCG 407

Consensus GCT---TC-A--------T-A-G-CG-G-----GCGC--G--CT--------C-GAG 570
  Therm B GCTTGCTC-AGAGG-GCATGA-GACGAG----CGCGC-GGTACTATTCCTATCGGAG 497
     clav GCT--CTCCA-CCGTGC-TCATG-CGGGGGCTGGCGCTGGCGCTCGGGAGGCCGGAG 504
 cattleya GCT--CTCCA-CGGTGA-TCATG-CGGGGGCTGGCGCTGGCGCTGGGGAAGCCGGAG 459
 fimbriae GCT--CTCCA-AGACGC-TGATG-CGGGGCTTCGCGCTGGCGCTGGGCAAGCCCGAG 459
  griseus GCT--CTCCA-AGGTAC-TGATG-CGAGGTTTCGCGCTGGCCCTGGGCAAGCCCGAG 504
Hygroscopi GCT--TTCCA-AGGCGC-TCATG-CGAGGTTTCGCGCTTGCCCTGGGCAAGCCCGAG 459

Consensus -A-TTCTTCGA---G-------C---C---G----A--C-CT-TC--C-GT-----C 627
  Therm B CA-TTCTTCGATAAGAGAAGGCCGCTCGGAG--TTATACCCTGTCGTCCGTAAAGAC 551
     clav CACTTCTTCGAC--GCGGCG--CTCGCCGAGCAGGACTCCCTGTCGTCCGTCTCG-C 556
 cattleya CACTTCTTCGAC--GCCGCG--CTCGCCGAGGCTGACACGCTCTCCTCGGTCTCC-C 511
 fimbriae GACTTCTTCGAC--GCGAAC--CTGCCGGAGGACGACACGCTGTCCGCCGTGTCG-C 511
  griseus GACTTCTTCGAC--GCTTCA--CTCTCACTGGCCGACACACTGTCCGCCGTGACG-C 556
Hygroscopi GACTTCTTCGAC--GCCCCG--CTCTCGGAGGCCGACACGCTGTCCGCCGTGTCG-C 511

Consensus ----CC--CT-CC---A-CT---AGG---A-C----GG--AAG-C-G----C-C--G 684
  Therm B GG--CCAGCT-CCTG-AGCTTCGAGGACCATCTGGAGG-GAAG-CTGATTACTCTGG 602
     clav TGATCC-GCTACCCGTATCTG-GAGGAGTACCCGCCGGTGAAGACGGGT--C-C-CG 607
 cattleya TGATCC-GCTACCCGTACCTG-GAGGACTACCCCCCGGTCAAGACCGGC--C-C-GG 562
 fimbriae TCATCC-GCTACCCCCACCTG-AAGGCCTACCCGCCGGTGAAGACGGGG--C-C-GG 562
  griseus TGATCC-ACTACCCGTACCTG-GAGGACTACCCGCCGGTCAAGACAGGC--C-C-GG 607
Hygroscopi TGATCC-GCTACCCCTACCTG-GAGGACTATCCGCCGGTCAAGACAGGG--C-C-CG 562

Consensus -C--C---G----T--G-TT--AG--CC----G-ACGT-T-G-T--TCA-CG--CT- 741
  Therm B TC--CTTCG--GTTGCGTTT--AG--CC----GAACGT-T-GCT--TCAGCGGGCTC 643
     clav ACGGCC-AGCTCCTGAGCTTCGAGGACCATCTGGACGTCTCGATGATCACCGTGCTC 663
 cattleya ACGGCA-CGAAGCTGAGCTTCGAGGACCACCTGGACGTGTCGATGATCACCGTGCTG 618
 fimbriae ACGGCA-CGAAGCTGAGCTTCGAGGACCACCTGGACGTGTCGGTGATCACCGTCCTG 618
  griseus ACGGCA-CGAAGCTGAGCTTCGAGGACCACCTGGACGTGTCGATGATCACCGTCCTG 663
Hygroscopi ACGGCA-CGAAGCTCAGCTTCGAGGACCACCTGGACGTGTCGATGATCACCGTCCTC 618

Consensus TTCC-AGA---AGG--C-GAACCT-C--C-GG--T---G-GA-C--G----ACGG-T 798
  Therm B TTCCCAGATTGAGG-ACTGAACCTGCGGCGGGGCTACGGGGAGCTTGATCGACGGCT 699
     clav TTCC-AGACCCAGGTGCAGAACCT-C--CAGG--T--GGAGA-CG-G-TCGACGGCT 709
 cattleya TTCC-AGACGGAGGTGCAGAACCT-C--CAGG--T--GGAGA-CG-G-TGGACGGCT 664
 fimbriae TTCC-AGACCGAGGTGCAGAACCT-C--CAGG--T--CGAGA-CG-G-TGAACGGCT 664
  griseus TTCC-AGACCGAGGTACAGAACCT-C--CAGG--T--CGAGA-CG-G-CCGACGGAT 709
Hygroscopi TTCC-AGACCGAGGTGCAGAACCT-C--CAGG--T--CGAGA-CT-G-TGGACGGCT 664
```

```
 Consensus GGC--GAC-T-CC--A-GT-C-------A--ACTTCCTGGT-AAC--C-GG--C-T- 855
   Therm B GGCGGGACATCCCCGA-GTTCTTGGCAGACGACTTCCTGGTCAACAACCGGCGCCT- 754
      clav GGCGGGACATCCC-GACGT-CGGAG--AACGACTTCCTGGTCAACTGC-GGTACCTA 761
  cattleya GGCGTGACCTGCC-GACGT-CCGGG--GAGGACTTCCTGGTGAACTGC-GGCACGTA 716
  fimbriae GGCAGGACCTGCC-GACGT-CGGGT--GACGACTTCCTGGTGAACTGC-GGCACCTA 716
   griseus GGCAGGACCTGCC-GACGT-CCGGG--GAGAACTTCCTGGTGAACTGC-GGCACCTA 761
Hygroscopi GGCAGGACCTGCC-CACGT-CCGGG--GAGGACTTCCTGGTGAACTGC-GGTACCTA 716

 Consensus C---G---A---T--C--ACGA--ACTTCCCGGC-CC-AACCACCG-GTGAA-TTC- 912
   Therm B C---GAAAACTGTT-CCTACGAG-ACTTCCCGGCGCCCAACCACCGGGTGAA-TTCG 805
      clav CATGGCGCAT-GTCACGAACGACTACTTCCCGGCGCCCAACCACCGGGTGAAGTTCG 817
  cattleya CCTGGGCCAC-GTCACCAACGACTACTTCCCGGCCCCCAACCACCGGGTGAAGTTCG 772
  fimbriae CATGGGGTAC-CTGACGAACGACTACTTCCCGGCGCCGAACCACCGGGTGAAGTTCA 772
   griseus CATGGGGTAC-CTCACCAACGACTACTTCCCGGCCCCCAACCACCGCGTGAAGTTCA 817
Hygroscopi CATGGGGTAC-CTCACGAACGACTACTTCCCGGCGCCGAACCACCGGGTGAAGTTCA 772

 Consensus T-A-CGC-GA--G-C--T---TGCCG----------A-G-CG-G--C------GT-A 969
   Therm B TTA-CGC-GAGTGTCAATTAATGCCGG---------AGGGCGCGAGCGAGGAGGTGA 851
      clav TGAACGCGGAGCGGCTGTCCCTGCCGTTCTTCCTCAACGGCGGGCACGAGGCGGTCA 874
  cattleya TCAACGCCGAACGG---------------------------------------------- 786
  fimbriae TCAACGCGGAGCGC---------------------------------------------- 786
   griseus TCAACGCGGAGCGGCTGTCCCTGCCGTTCTTCCTCCACGCCGGGCACACCACCGTGA 874
Hygroscopi TCAACGCGGAGCGA---------------------------------------------- 786

 Consensus --GA-C----------GAGG-C-----CG-G--GG---G-AC--G-C--T---CT 1026
   Therm B -GGAACGAACAAGAATGGAGGCCA-GACCG-GTTGGAAAGGGACCTGA--------- 896
      clav TCGAGCCGTTCGTGCCGGAGGGCGCGAGCGAGGAGGTGAGGAACGAGGCCCTGTCCT 931
  cattleya ---------------------------------------------------------- 786
  fimbriae ---------------------------------------------------------- 786
   griseus TGGAGCCGTTCAGCCCCGAGGACACCCGCGGGAAGGAGCTGAACCCGCCGGTCCGCT 931
Hygroscopi ---------------------------------------------------------- 786

 Consensus ACGG-GACTACCTCCAGCA-G--------GC-CTGATCG--AAGAACGGCCAGACCT 1083
   Therm B ---------------------------------------------------------- 896
      clav ACGGGGACTACCTCCAGCACGGGCTGCGGGCGCTGATCGTCAAGAACGGCCAGACCT 988
  cattleya ---------------------------------------------------------- 786
  fimbriae ---------------------------------------------------------- 786
   griseus ACGGCGACTACCTCCAGCAGGCTTCCAACGCCCTGATCGCGAAGAACGGCCAGACCT 988
Hygroscopi ---------------------------------------------------------- 786
```

Figure 6.12. Nucleotide sequence alignment of Thermophilic B against known IPNS sequences.

Highly conserved regions are shown in bold in the Thermophilic B sequence data.

Sequencing data was obtained for *S. tanashiensis* from the samples that were sent to U.C.T. (Figure 6.13). The sequence for *S. tanashiensis* (Figure 6.14) was found to be 66.6%, 68.1%, 68.8%, 65.8% identical to the IPNS sequences for *S. clavuligerus*, *S. griseus*, *S. jumonjinensis* (GenBank M36687) and *S. lipmanii* (GenBank AF133827) respectively (Figure 6.15). However, this was lower than the expected 70-85% identity. The sequence alignments are shown in Figure 6.16.

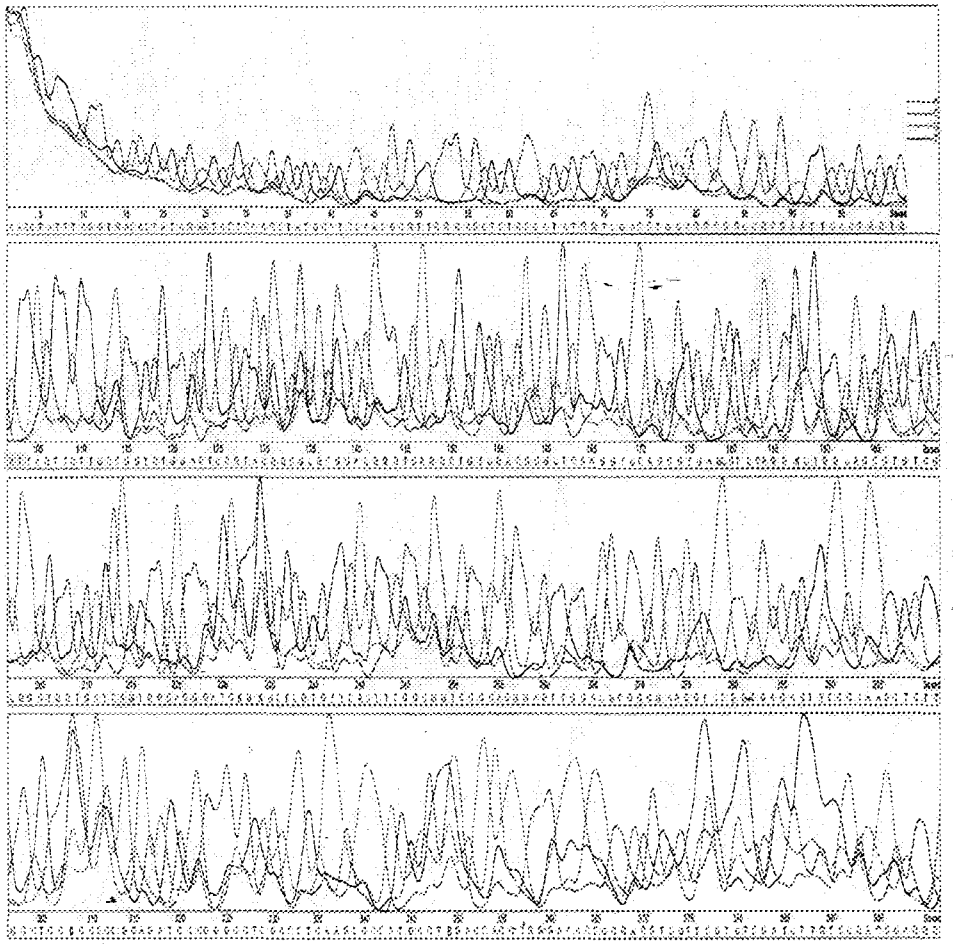


Figure 6.13. Electropherogram of sequence obtained for *S. tanashiensis* from clone 1.

TGCCCGCGGT CCCAGTGCGG AGTAGCGGGA TCATTTGCAC TGCCCCATTC TGGGCAGCAC

GAGCGGCGGA TTCTGCTAGC TACGTATTAG TACATCTCGC CGCTGTCCGG AGACGACGTG

TGATCGCGAA AGGACGCGTC GAAGACGGTG TTTTGGACGT AGATAGGGGT CCGGCTTCTT

CTACGCGTGA ACCCGGTGTG GACGTACAGC TGCTCCAGGA CAGTAATCGA AGTTCACCGG

AACAGGAGGC GGATAGACGG AGGATGCACG ACCTGGCATC GGAGTTCCAC CGGAACATGG

GAGAGACCGG TAGGAGCGCC GGCGGGATGC CGTAGGTGGT ACAGAAGGCC GTCGATGAGT

CCATCAGCTT CGGAGAACCC CGTCGTGTTC GACCGAACAC GACTACTGAA AGTGCGAGCG

AGTGCAGCAA ACGAGGATGG GGCGCGTCGG TTTGCGCCAC GGAGAGGACC ATCTAGAGCC

AGGTCGAGCG CTACGGTAGG ACCAACCAGA GGTTCCCCCG CTCGCGGATA CGATTGTAAT

TGAGAAACAG GATTGTCCTA GCACACAGCA GCTCATGGGT CGTCAAGCGA GCATTCGCTG

CGCGCTATAC TAGCGTCGGA GTGGGTGTAG GTCTAATATC AGATCCGAGC CAGCGTATCG

TAGCAATCCG ACTCGCCATG CTTGTCCACC AGGCACCTGG TAACAGCTCG TGGACGACAA

AGCGAGTGCG GCCAGCGATA AAAGACGCTC GGTAGCACCA TCCAGCCTGG AAGCAGACAG

TCCAGCGGCA TGCGAGTAGA CGCTAGCCGT CGCGCGATTA TGGAACACAC ATCGTGGACG

GTAACGGCAG TGCGACGCAG TGCCAGCCGG ATAGCATTGC GCGCCGCGAC TGCGCGCGTG

GGCCTCTGCA CTAGCAAGGT TAGCGCGCCG CGAGCGCGTG CGTGGAAGGC TCGGTAGCGG

AGCAGCGACC GGAGCGAGCT AAGGTGAGCT AGTTACATGT GTCGCCCGGG ATCGGCGCTC

GCGCCAGCTA GACGAGAGCA CGAGGCCGAT CCTACAAGCG AGCGACGAGT ACTAGGCGTG

CTCTTGCTGA TCAAGCGGAG CGCAGCTCGG CGGTGCGAGC CAACGGACAG CGCGAGGATT

CAGACCTGGA ACGCGTAGGA GCAAGAACTG A

Figure 6.14. Nucleotide sequence from IPNS gene of *S. tanashiensis*.

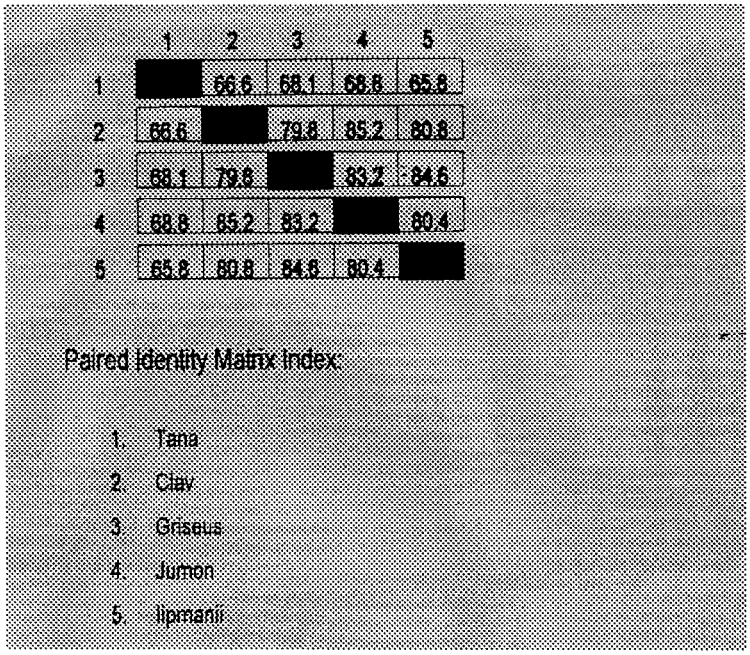


Figure 6.15 Percentage identity matrix of *S. tanashiensis* against 2= *S. clavuligerus*; 3= *S. griseus*; 4= *S. jumonjinensis* and 5= *S. lipmanii.*

```
Consensus TGCCGCGCGCCTGCCC-G----CA---CCAGGTGCGGAGTACA-GG----TTTGCA- 57
     tana ----------CTGCCC-GAGGTCA---CCAGGTGCGGAGTACAGGGTTCATTTGCAC 43
    jumon TGCCGCGCGCCTGCCCAGTCACCAGCGCCAGGTGCGGAGTACACGGAGG-TTTGCA- 55
     catt ---------------------------------------------------------- 0
  griseus ----------------------------------------------------------A- 1
     clav ----------------------------------------------------------A- 1

Consensus TGCC---TTC-----A-GC--G--CGGC--A---T-C---C-AC--AT--G-ACATC 114
     tana TGCCCCATTCTGGGCA-GCACGAGCGGCGGATTCTGCTAGCTACGTATTAGTACATC 99
    jumon TGCCC-ATTCTG---ATGC-CGT-CGGCGGAAG-TTC---CGACC-ATC-G-ACATC 99
     catt ---------------------------------------------------------- 0
  griseus TGCCC-ATTCCC---ATGC-TGC-CGGCACACG-TGC---CGACG-ATC-G-ACATC 45
     clav TGCCA-GTTCTG---ATGC-CGT-CGGCCCACG-TTC---CGACC-ATC-G-ACATC 45

Consensus TC-CCGCTGT-CGG---CGACG-------G--AA-G---CG-GTCG-----GG-G-T 171
     tana TCGCCGCTGTCCGGAGACGACGTGTGATCGCGAAAGGA-CGCGTCGAAGACGGTGTT 155
    jumon TCGCCGCTGTCCGGAGACGACGCG-AA-GGCGAA-GCAGCGGGTCGCGCA-GGAGAT 152
     catt TCGCCGCTGTCCGGGTCCGACGCC-GA-CGCCAA-GTCACGGGTCGCGCG-GGAGAT 53
  griseus TCACCGCTGTCCGGCGGCGACGCG-GA-CGACAA-GAAGCGCGTCGCACA-GGAGAT 98
     clav TCGCCGCTGTTCGGAACCGACGCC-GC-CGCGAA-GAAGCGCGTCGCCGA-GGAGAT 98

Consensus -------G-C--------G--TC-GG-TTCTTCTACGC-----ACC-CGG--T-GAC 228
     tana TT----GGACGTAGATAGGGGTCCGGCTTCTTCTACGCGT-GAACC-CGGTGTGGAC 206
    jumon CAACAAGG-CTG-CCCGCGGGTCCGGCTTCTTCTACGCGTCGAACCACGGTGTGGAC 207
     catt CCACGACG-CGT-GCCGCGGATCCGGCTTCTTCTACGCCTCGCACCACGGCGTCGAC 108
  griseus CAACAAGG-CGT-GCCGGGAGTCGGGGTTCTTCTACGCCTCCCACCACGGCATCGAC 153
     clav ACACGGGG-CCT-GCCGCGGCTCGGGCTTCTTCTACGCCACGAACCACGGCGTGGAC 153

Consensus GT-CAGC-GCT--AGGAC---GT-AA-CGA-GTTC-AC-G---CA-GA----GA--- 285
     tana GTACAGCTGCTCCAGGACA--GT-AATCGAAGTTC-ACCGGAACAGGAGGCGGATAG 259
    jumon GTACAGCTGCTCCAGGACGTGGTGAA-CGA-GTTCCACCGGAACATGA--GCGAC-C 259
     catt GTCCAGCTGCTCCAGGACGTGGTCAA-CGA-GTTCCACCGGACCATGA--CCGAC-C 160
  griseus GTCCAGCTGCTGAAGGACGTCGTCAA-CGA-GTTCCACCGGACCATGA--CGGAC-G 205
     clav GTCCAGCAGCTCCAGGACGTGGTGAA-CGA-GTTCCACGGCGCCATGA--CCGAC-C 205

Consensus A-GGAG-A-G-ACGACCTGGC----GA--TC-AC-C--A--AC--G----A-ACC-G 342
     tana ACGGAGGATGCACGACCTGGCATCGGAGTTCCAC-CGGA--ACATGGGAGAGACCGG 313
    jumon A-GGAGAA-GCACGACCTGGC----GA--TCAACGCGTACAACAAGG-ACA-ACCCG 306
     catt A-GGAGAA-GTACGACCTGGC----GA--TCCACGCGTACAACAAGG-CCA-ACCCG 207
  griseus A-GGAGAA-GTACGACCTGGC----GA--TCAACGCTTACAACAAGA-ACA-ACCCG 252
     clav A-GGAGAA-GCACGACCTGGC----GA--TCCACGCGTACAACCCGG-ACA-ACCCG 252

Consensus ------CGC--CGGC------A-G-C--T---G-GG-A--GAAGGCCGTCGA---GT 399
     tana TAGGAGCGC--CGGCGGG---ATG-CCGT-AGGTGGTACAGAAGGCCGTCGATGAGT 363
    jumon CATGTGCGCAACGGCTACTACAAGGCGATCAAG-GGGA-AGAAGGCCGTCGA---GT 358
     catt CACATCCGCAACGGCTACTACATGGCGGTGAAG-GGGA-AGAAGGCCGTCGA---GT 259
  griseus CGCACCCGCAACGGCTACTACATGGCCGTCAAG-GGCA-AGAAGGCCGTCGA---GT 304
     clav CATGTGCGCAACGGCTACTACAAGGCGGTCCCG-GGCA-GGAAGGCCGTCGA---GT 304

Consensus CC----G-T-C----AACCC-G---T--TC---CGA----GAC-AC--------C-- 456
     tana CCATCAGCTTCGGAGAACCCCGTCGTGTTCGACCGAACACGACTACTGAAAGTGCGA 420
    jumon CCTTCTGCTACCT-GAACCC-GTCGT--TC-TCCGA---CGACCAC--------C-- 397
     catt CCTTCTGCTACCT-CAACCC-GGCCT--TC-ACCGA---CGACCAC--------C-- 298
  griseus CCTGGTGCTACCT-GAACCC-GTCGT--TC-AGCGA---GGACCAC--------C-- 343
     clav CCTTCTGTTACCT-CAACCC-GGACT--TC-GGCGA---GGACCAC--------C-- 343
```

```
Consensus -CG---G-A-C----------G---C-C--C--T--GC-C-A-GG--A--AC----T 513
     tana GCGAGTGCAGCAAACGAGGATGGGGCGCGTCGGTTTGCGCCACGGAGAGGACCATCT 477
    jumon -CGA-TG-ATCAA----ATCCGAGACCC--CGAT--GCACGA-GGTGA--ACC-TCT 439
     catt -CGA-TG-ATCAA----GGCCGGTACGC--CCCT--GCACGA-GGTGA--ACC-GGT 340
  griseus -CGC-AG-ATCAG----GTCCGGGACGC--CGAT--GCACGAAGGT-A--ACA-TCT 385
     clav -CGA-TG-ATCGC----CGCGGGGACGC--CGAT--GCACGA-GGTGA--ACC-TCT 385

Consensus -G-GCC--G---A-CG--A-G--A----CA--C-G-G-TTCC---C--TC-GCG-G- 570
     tana AGAGCCAGGTCGAGCGCTACGGTAGGACCAACCAGAGGTTCCC-CCGCTC-GCG-GA 531
    jumon -G-GCC-GG---A-CG--A-GG-AGAAGCA-CCCGCGGTTCCGGCCCTTCTGCGAGG 484
     catt -G-GCC-GG---A-CG--A-GG-AACGCCA-TCCGCGGTTCCGGTCGTTCTGCGAGA 385
  griseus -G-GCC-GG---A-CG--A-GA-AGCGCCA-CCAGCGCTTCCGGCCCTTCTGCGAGC 430
     clav -G-GCC-CG---A-CG--A-GG-AGCGGCA-CCCGCGCTTCCGGCCGTTCTGCGAGG 430

Consensus ----A-T---A--G-----C-G----G-C-T--------GC---TC-A--G-T--T- 627
     tana TACGATTGTAATTGAGAAACAGGATTGTCCTAGCACACAGCAGCTC-ATGGGTCGTC 587
    jumon -ACTACT---ACCG-G---CAG-CT-G-C-T-CCG----GCT-TTCCACGG-TGATC 522
     catt -GCTACT---ACCG-G---CAG-AT-G-C-T-GCG----GCT-CTCCACGG-TGATC 423
  griseus -AGTACT---ACCG-----C-G-AC-GTC-T-TCTC---GCT-CTCCAAGG-TACTG 468
     clav -GCTACT---ACCG-G---CAG-AT-G-C-T-GAA----GCT-CTCCACCG-TGCTC 468

Consensus A-GCG-G------CGCTG-GC-CT------GC----GAG----T-----TCT----T 684
     tana AAGCGAGCATT--CGCTGCGCGCTATACTAGCGTCGGAGTGGGTGTAGGTCTAATAT 642
    jumon ATGCGGGGGTACGCGCTG-GCGCTGGGCCGGC--GCGAG-GACT-----TCT----T 566
     catt ATGCGGGGGCTGGCGCTG-GCGCTGGGGAAGC--CGGAG-CACT-----TCT----T 467
  griseus ATGCGAGGTTTCGCGCTG-GCCCTGGGCAAGC--CCGAG-GACT-----TCT----T 512
     clav ATGCGGGGGCTGGCGCTG-GCGCTCGGGAGGC--CGGAG-CACT-----TCT----T 512

Consensus C-GA--CG---C--C---TC----------GAC-C-C------T-TC--C--G---C 741
     tana CAGATCCGAGCCAGCGTATCGTAGCAATCCGACTCGCCATGCTTGTCCACCAGGCAC 699
    jumon C-GA--CGAGGC-GC---TCGCCG-AGGCCGACACGC------TGTCCTCC-GTGTC 608
     catt C-GA--CGCCGC-GC---TCGCCG-AGGCTGACACGC------TCTCCTCG-GTCTC 509
  griseus C-GA--CGCTTC-AC---TCTCAC-TGGCCGACACAC------TGTCCGCC-GTGAC 554
     clav C-GA--CGCGGC-GC---TCGCCG-AGCAGGACTCCC------TGTCGTCC-GTCTC 554

Consensus -CTG-T--C--CT---C-T--A---C---AG-GA-T----CC--CG-T-AA-GAC-- 798
     tana -CTGGTAACAGCT---CGTGGACGACAA-AGCGAGTGCGGCCAGCGATAAAAGAC-- 749
    jumon GCTGAT--CCGCTACCCCT--AT--CTGGAG-GAGTAT--CCGCCGGTCAA-GACGG 655
     catt CCTGAT--CCGCTACCCGT--AC--CTGGAG-GACTAC--CCCCCGGTCAA-GACCG 556
  griseus GCTGAT--CCACTACCCGT--AC--CTGGAG-GACTAC--CCGCCGGTCAA-GACAG 601
     clav GCTGAT--CCGCTACCCGT--AT--CTGGAG-GAGTAC--CCGCCGGTGAA-GACGG 601

Consensus G--C-G-A--G-------C--AGC-T-GA-G--GAC----CA-C-GG-A--CG--T- 855
     tana GCTCGGTA--GCACCATCC--AGCCTGGAAGCAGACAGTCCAGC-GGCATGCGAGTA 801
    jumon GCGCGG-ACGGTACCAAGCTGAGCTTCGA-G--GAC----CATCTGG-A--CGTCTC 701
     catt GCCCGG-ACGGCACGAAGCTGAGCTTCGA-G--GAC----CACCTGG-A--CGTGTC 602
  griseus GCCCGG-ACGGCACGAAGCTGAGCTTCGA-G--GAC----CACCTGG-A--CGTGTC 647
     clav GTCCCG-ACGGCCAGCTCCTGAGCTTCGA-G--GAC----CATCTGG-A--CGTCTC 647

Consensus GA-G-T-A-CCGT--C----C--GA---A-GG--CA-A-C-TC---GT-G-GACGG- 912
     tana GACGCT-AGCCGTCGC-G--C--GATT-ATGGAACACA-CATC---GT-G-GACGGT 845
    jumon GATGATCA-CCGT-GCTGTACCAGACCGA-GGTGCAGAACCTCCAGGTCGAGACGGT 755
     catt GATGATCA-CCGT-GCTGTTCCAGACGGA-GGTGCAGAACCTCCAGGTGGAGACGGT 656
  griseus GATGATCA-CCGT-CCTGTTCCAGACCGA-GGTACAGAACCTCCAGGTCGAGACGGC 701
     clav GATGATCA-CCGT-GCTCTTCCAGACCCA-GGTGCAGAACCTCCAGGTGGAGACGGT 701
```

```
Consensus --ACGG--G-GC--GAC----T-CC-A-G-C-G---A--A-TT-C-C-G--G--A-C 969
     tana A-ACGGCAGTGC--GACGCAGTGCC-A-G-CCGGATAGCA-TTGCGC-GCCGCGA-C 893
    jumon CGACGGCTG-GCAGGAC-A--TACCGAGGTCCGACGAGGACTT-C-CTG--GTGAAC 804
     catt GGACGGCTG-GCGTGAC-C--TGCCGACGTCCGGGGAGGACTT-C-CTG--GTGAAC 705
  griseus CGACGGATG-GCAGGAC-C--TGCCGACGTCCGGGGAGAACTT-C-CTG--GTGAAC 750
     clav CGACGGCTG-GCGGGAC-A--TCCCGACGTCGGAGAACGACTT-C-CTG--GTCAAC 750

Consensus TGCG---C---C-TGG------T-CAC---C-A--A--TT--CG-C-CC---A-C-C 1026
     tana TGCG---CG--CGTGGGCCTC-TGCACTAGC-A--AGGTTAGCG-CGCCGCGAGCGC 940
    jumon TGCGGTACGTACATGGGCCACAT-CACCCACGACTAC-TTCCCGGCCCC-GAACCAC 858
     catt TGCGGCACGTACCTGGGCCACGT-CACCAACGACTAC-TTCCCGGCCCC-CAACCAC 759
  griseus TGCGGCACCTACATGGGGTACCT-CACCAACGACTAC-TTCCCGGCCCC-CAACCAC 804
     clav TGCGGTACCTACATGGCGCATGT-CACGAACGACTAC-TTCCCGGCGCC-CAACCAC 804

Consensus --G-GTG-AAG--TC--T-----A--A-CG--C-GA-CG-GCT---GT---CT-G-- 1083
     tana GTGCGTGGAAGGCTCGGTAGCGGAGCAGCGACCGGAGCGAGCTAAGGTGAGCTAGTT 997
    jumon C-GGGTG-AAGT-TCA-T--C--A--A-CG--CGGAGCG-GCT---GTC-GCT-GC- 895
     catt C-GGGTG-AAGT-TCG-T--C--A--A-CG--CCGAACG-G---------------- 786
  griseus C-GCGTG-AAGT-TCA-T--C--A--A-CG--CGGAGCG-GCT---GTC-CCT-GC- 841
     clav C-GGGTG-AAGT-TCG-T--G--A--A-CG--CGGAGCG-GCT---GTC-CCT-GC- 841

Consensus -C--GT-TC--CC----TC--CG--CG-GC-A-C-------G---A--GAG-CCG-T 1140
     tana ACATGTGTCGCCCGGGATCGGCGCTCGCGCCAGCTAGACGAGAGCA-CGAGGCCGAT 1053
    jumon -C--GT-TCTTCC----TCAACGC-CGGGC-A-CAACAGC-GT-CATCGAG-CCGTT 938
     catt ---------------------------------------------------------- 786
  griseus -C--GT-TCTTCC----TCCACGC-CGGGC-A-CACCACC-GT-GATGGAG-CCGTT 884
     clav -C--GT-TCTTCC----TCAACGG-CGGGC-A-CGAGGCG-GT-CATCGAG-CCGTT 884

Consensus C---C--G----C--C-----C--G--G-G-----G--GA-C----G---C---G-T 1197
     tana CCTACAAGCGAGCGACGAGTACTAGGCGTGCTCTTGCTGATCAAGCGGAGCGCAGCT 1110
    jumon CGTGCCGGAAGGCG-CGGC--CGGGACG-G----TGAAGAAC--CCGA--CGACG-T 982
     catt ---------------------------------------------------------- 786
  griseus CAGCCCCGAGGACA-CCCG--CGGGAAG-G----AGCTGAAC--CCGC--CG--G-T 926
     clav CGTGCCGGAGGGCG-CGAG--CGAGGAG-G----TGAGGAAC--GAGG--CCCTG-T 928

Consensus C-----T-CG--G---AC---CAGC--G-G--T-C---G--CTG-A-CG---AG-A- 1254
     tana CGGCGGTGCGA-GCCAACGGACAGCGCGAGGATTCA--GACCTGGAACGCGTAGGA- 1163
    jumon CG----TACGGGGAGTACCTTCAGCACG-GCCTCCGT-GCGCTG-ATCGTCAAGAAC 1032
     catt ---------------------------------------------------------- 786
  griseus CCGC--TACGGCGACTACCTCCAGCA-G-GCTTCCAACGCCCTG-ATCGCGAAGAAC 978
     clav CC----TACGGGGACTACCTCCAGCACG-GGCTGCGG-GCGCTG-ATCGTCAAGAAC 978

Consensus -GC-AGA-CTG                                                1265
     tana -GCAAGAACTG                                                1173
    jumon GGCCAGACCTG                                                1043
     catt -----------                                                786
  griseus GGCCAGACCTG                                                989
     clav GGCCAGACCTG                                                989
```

Figure 6.16. Nucleotide alignment sequence of *S. tanashiensis* against known IPNS sequences. Highly conserved regions are shown in bold in the *S. tanashiensis* sequence data.

Open reading frame predictions for the sequence data obtained for *S. tanashiensis* revealed six open reading frames (Figure 6.17). The first three open reading frame were 5' to 3' in the 1 frame on the sequenced DNA strand, and fourth, fifth, and sixth were 3' to 5' in the -1, -2, and -3 frames respectively on the complimentary DNA strand.

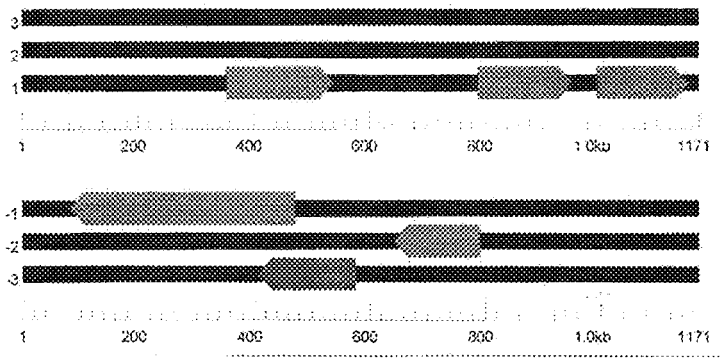


Figure 6.17. Predicted open reading frames for the sequence data obtained for *S. tanashiensis*.

In summary, the sequence data obtained for IPNS genes of *S. hygroscopicus*, Thermophilic A, Thermophilic B and *S. tanashiensis* suggest that the sequenced genes are *Streptomyces* IPNS genes, in view of the degree of similarity (generally greater than 70%) when compared with other known IPNS gene sequence data. The sequence data obtained also indicate a high GC content which is typical of the *Streptomyces* IPNS genes. The consensus amongst the strains used for comparison show regions which are conserved which also seems to be expected of *Streptomyces* IPNS genes.

The predicted open reading frames for *S. hygroscopicus*, *S. tanashiensis*, Thermophilic A and Thermophilic B were aligned against documented sequences for *S.clavuligerus*, *S. griseus*, *S. jumonjinensis* and *S. lipmanii* to obtain identity information (Table 6.1). The lowest percentage identity 58% is obtained for *S. hygroscopicus* against *S. clavuligerus* and the highest was obtained for *S. tanashiensis* against *S. lipmanii*.

Table 6.1. Percentage identity of the open reading frames against documented strains.

ORF	*S. clavuligerus* (%)	*S. griseus* (%)	*S. jumonjinensis* (%)	*S. lipmanii* (%)
Hygro 27-182	60.3	69.2	63.5	58.3
Hygro 183-518	58.0	64.0	68.2	59.5
Tana 355-540	67.2	70.4	65.6	72.0
Tana 406-576	67.3	66.7	69.6	61.4
Tana 642-791	68.0	67.3	62.0	68.0
Tana 790-954	64.2	63.6	63.0	64.9
Tana 83-472	65.6	65.6	68.7	65.4
Tana 997-1155	68.6	67.3	66.7	67.9
ThermA 252-452	65.7	60.2	67.2	61.2
ThermA 629-865	64.1	62.9	65.4	65.8
ThermA 938-1105	69.6	61.9	62.5	70.2

The sequencing of one of the shuffled gene products was carried out for *S. jumonjinensis* and the sequence data (Figure 6.18) obtained, when compared to the unshuffled gene (Figure 6.19) show a percentage identity of 77%.

TGCCGCGCGC GTTGGGAGTT GATGTACCTA AATGCCTAGC TAGGAGCCAT TTGCATGGGG
ATTCTGATGC GGTTGGTGCA GTTCGGTTGT CCATCGAGCG GAGAGCGCGC CACCACCGCG
ACCGTTGGTT AGTAGCGTAG CAGCCTCGTC GACGGAATGG AGATCAGCCC CCCGCGGCGT
CAGTAGCCTT CTCTTCTAGC GTCGAACCAC GGTGAGACGA GATCACTTAA GGCTGCGCGC
TTCGTTGTCT TCCACCAACA TGAGCGATTA GGCTAGCCGA CCTGTGATCA ACGCGTACAA
CAAGGACGGA AGGCATGTGC GCAACGGCTA CTACAAGGCG ATCAAGGGGT AAGTTGCTTG
CGATTCTTCT GCTACCTGAG AGGAGATGTG ACGACGGCTT ATCCGACGAC CACTTGATGA
TCAAATCCGA TACGGGCATG CACGAGGTGA ACCTCTGGTG GCCCAGAGAC CACCGCCGAA
GGTCGGCCCG TCTCGGAGGA CTACTACCGG CAGCTGCAGC GTTGTTCGAG TGCTCATGCC
CCCGTACGCG TGCGCTGTTG GGGCGCGAGG ACTTCGTCTA CGGGTGTCAT GCCCGAGGCC
AGTACTGTCC AGGCGCGGTC GTCTGTAGAT TCAGAGCCGC TTGGCGTGTG TGAGCCATGG
GTGTGATTTG AGTAAAGTGA ATTTTATCAG AGGGAGGTGG GAGTTCCTGC AGCTACATCT
TGATTTGAAG ACTGCGAGGA AGTCAGTTCT AGCCTATGGA GCGGTCGATC AGCTTTTTTT
GGGAGAGTGT CCAGGTATCC GTGTCGCTGA TCCGCTAGGG GTATCTGGAG GAGTATCCGC
CGGTCAAGAC GGGCGCGGAC GGTACCAAGC TGAGCTTCGA GGACCATCTG GACGTCTCGA
TGATCACCGT GCTGTACCAG ACCGAGGTGC AGAACCTCCA GGTCGAGACG GTCGACGGCT
GGCAGGACAT ACCGAGGGTT GAAGGACAGT GTGAGTAGGT CATGGGGGCC ATCACCCACG
ACTACAGGAG ACTTGGGGCC GGGGCAACGA CCGGGTGAAG TTCATCAACG CGGAGCGGCT
GTCGCTGCCG TTCGCGTGCT GTCTGGGCTT TCACCGGATG CTTCTGGCGT TCTGTATGTC
GCTATTGATG TGTGATGGAG GGCAGGCGCG GGGAATCGTC ACATATTTGG TAGGGGCCAC
TTCTTAACAG GGATCTCTAA CTGCGCAGAC CCAGATGTTG TGCTGGTAGT TAGCAGGGTA
AGCACATTCG CTCGTGTCGG TCTTGCAGCG GCACCTCGTG GTGTGGTTAG AGAGAGATAG
AC

Figure 6.18. Nucleotide sequence of shuffled *S. jumonjinesis* IPNS gene.

```
Consensus TGCCGCGCGC-T----AGT--------C-A--TGC--AG--C--GGAG---TTTGCA  57
jumonjinen TGCCGCGCGCCTGCCCAGTCACCAGCGCCAGGTGCGGAGTACACGGAGG--TTTGCA 55
Jumonshuff TGCCGCGCGCGTTGGGAGTTGATGTACCTAAATGCCTAG--CTAGGAGCCATTTGCA 55

Consensus TG---ATTCTGATGC-GT-GG-G-A-GTTC-G----CCATCGA-C-----GC-CGC-  114
jumonjinen TGCCCATTCTGATGCCGTCGGCGGAAGTTCCGA---CCATCGA-CATCTCGC-CGCT 107
Jumonshuff TGGGGATTCTGATGCGGTTGGTGCA-GTTCGGTTGTCCATCGAGCGGAGAGCGCGCC 111

 Consensus --C-C-G-GAC-G--G---AG--GCG-AGCAGC---GTCG-CG-A--GGAGATCA-- 171
jumonjinen GTC-CGGAGAC-GACGCGAAG--GCGAAGCAGCGG-GTCG-CGCA--GGAGATCAAC 156
Jumonshuff ACCACCGCGACCGTTGGTTAGTAGCGTAGCAGCCTCGTCGACGGAATGGAGATCA-- 166

 Consensus ---GC--CCCGCGG-GTC-G--GC-TTCT--TCTA-GCGTCGAACCACGGTG-G-AC 228
jumonjinen AAGGCTGCCCGCGG-GTCCG--GC-TTCT--TCTACGCGTCGAACCACGGTGTGGAC 207
Jumonshuff ---GCCCCCCGCGGCGTCAGTAGCCTTCTCTTCTA-GCGTCGAACCACGGTGAG-AC 218

 Consensus G-A-A-C--CT--AGG--G-G-----CG--GT--TCCACC--AACATGAGCGA--AG 285
jumonjinen GTACAGCTGCTCCAGGACGTGGTGAACGA-GT--TCCACCGGAACATGAGCGACCAG 261
Jumonshuff G-AGATCA-CTTAAGGCTGCGCGCTTCGTTGTCTTCCACC--AACATGAGCGATTAG 271

 Consensus G--A-GC-CGACCTG--GATCAACGCGTACAACAAGGAC-----GCATGTGCGCAAC 342
jumonjinen GAGAAGCACGACCTGGCGATCAACGCGTACAACAAGGACAACCCGCATGTGCGCAAC 318
Jumonshuff GCTA-GC-CGACCTGT-GATCAACGCGTACAACAAGGACGGAAGGCATGTGCGCAAC 325

 Consensus GGCTACTACAAGGCGATCAAGGGG-A----GC---CGA-TC-TTCTGCTACCTGA-A 399
jumonjinen GGCTACTACAAGGCGATCAAGGGGAAGAAGGCCGTCGAGTCCTTCTGCTACCTGA-A 374
Jumonshuff GGCTACTACAAGGCGATCAAGGGGTAAGTTGCTTGCGATTC-TTCTGCTACCTGAGA 381

 Consensus ----------CG-CG--TT-TCCGACGACCAC--GATGATCAAATCCGA-AC----A 456
jumonjinen CC--------CGTCG--TTCTCCGACGACCACCCGATGATCAAATCCGAGACCCCGA 421
Jumonshuff GGAGATGTGACGACGGCTTATCCGACGACCACTTGATGATCAAATCCGATACGGGCA 438

 Consensus TGCACGAGGTGAACCTCTGG--GG-C-AG-AGA--CACC--CG--G-TC-GGCCC-T 513
jumonjinen TGCACGAGGTGAACCTCTGGCCGGACGAGGAGAAGCACCCGCG--GTTCCGGCCCTT 476
Jumonshuff TGCACGAGGTGAACCTCTGGT-GGCCCAG-AGAC-CACCGCCGAAGGTC-GGCCCGT 491

 Consensus CT--GAGGACTACTACCGGCAGCTGC-------T-CG---GCT--T-CC-----ACG 570
jumonjinen CTGCGAGGACTACTACCGGCAGCTGC-------TCCG---GCT--TTCC-----ACG 516
Jumonshuff CTCGGAGGACTACTACCGGCAGCTGCAGCGTTGTTCGAGTGCTCATGCCCCCGTACG 548

 Consensus -GTG---------------A-----------------TCATGC--G-GG---GTAC 627
jumonjinen -GTG---------------A-----------------TCATGC--GGGG---GTAC 534
Jumonshuff CGTGCGCTGTTGGGGCGCGAGGACTTCGTCTACGGGTGTCATGCCCGAGGCCAGTAC 605

 Consensus -------GCGC-G-CG-CTG--G-------GCCG----GCG--------C------- 684
jumonjinen -------GCGCTGGCG-CTG--G-------GCCG----GCG--------C------- 555
Jumonshuff TGTCCAGGCGCGGTCGTCTGTAGATTCAGAGCCGCTTGGCGTGTGTGAGCCATGGGT 662

 Consensus --GA---------------------------------GGA-TTC------------ 741
jumonjinen --GA---------------------------------GGACTTC------------ 564
Jumonshuff GTGATTTGAGTAAAGTGAATTTTATCAGAGGGAGGTGGGAGTTCCTGCAGCTACATC 719

 Consensus ----TT-GA------CGAGG----C-G-TC--GCC----GAG--G-CGA---GCT-- 798
jumonjinen ----TTCGA------CGAGG----C-GCTC--GCC----GAG--GCCGACACGCT-- 596
Jumonshuff TTGATTTGAAGACTGCGAGGAAGTCAGTTCTAGCCTATGGAGCGGTCGATCAGCTTT 776
```

```
 Consensus ------------GTCC-----TCCGTGTCGCTGATCCGCTA----TATCTGGAGGAG 855
jumonjinen ------------GTCC-----TCCGTGTCGCTGATCCGCTACCCCTATCTGGAGGAG 636
Jumonshuff TTTTGGGAGAGTGTCCAGGTATCCGTGTCGCTGATCCGCTAGGGGTATCTGGAGGAG 833

 Consensus TATCCGCCGGTCAAGACGGGCGCGGACGGTACCAAGCTGAGCTTCGAGGACCATCTG 912
jumonjinen TATCCGCCGGTCAAGACGGGCGCGGACGGTACCAAGCTGAGCTTCGAGGACCATCTG 693
Jumonshuff TATCCGCCGGTCAAGACGGGCGCGGACGGTACCAAGCTGAGCTTCGAGGACCATCTG 890

 Consensus GACGTCTCGATGATCACCGTGCTGTACCAGACCGAGGTGCAGAACCTCCAGGTCGAG 969
jumonjinen GACGTCTCGATGATCACCGTGCTGTACCAGACCGAGGTGCAGAACCTCCAGGTCGAG 750
Jumonshuff GACGTCTCGATGATCACCGTGCTGTACCAGACCGAGGTGCAGAACCTCCAGGTCGAG 947

 Consensus ACGGTCGACGGCTGGCAGGACATACCGAGG---GA--AGGAC----TG-TGA----- 1026
jumonjinen ACGGTCGACGGCTGGCAGGACATACCGAGGTCCGACGAGGACTTCCTGGTGAACTGC 807
Jumonshuff ACGGTCGACGGCTGGCAGGACATACCGAGGGTTGA--AGGACAG--TG-TGA----- 994

 Consensus -GTA-GT-CATGGG---CATCACCCACGACTAC------C--GG--CCG----AAC- 1083
jumonjinen GGTACGTACATGGGCCACATCACCCACGACTACTT----CCCGGCCCCG----AACC 856
Jumonshuff -GTAGGT-CATGGGGGCCATCACCCACGACTACAGGAGACTTGGGGCCGGGGCAACG 1049

 Consensus ACCGGGTGAAGTTCATCAACGCGGAGCGGCTGTCGCTGCCGTTC-----C----C-G 1140
jumonjinen ACCGGGTGAAGTTCATCAACGCGGAGCGGCTGTCGCTGCCGTTCTTCCTCAACGCCG 913
Jumonshuff ACCGGGTGAAGTTCATCAACGCGGAGCGGCTGTCGCTGCCGTTCGCGTGCTGGTCTG 1106

 Consensus GGC---CA-C-G---C-TC--GC-GTTC-----GTGC----GA-G-G-G--G--GG- 1197
jumonjinen GGCA--CAACAGCGTCATCGAGCCGTTC-----GTGCCG--GAAG-GCGCGGCCGGG 960
Jumonshuff GGCTTTCACCGGATGCTTCTGGC-GTTCTGTATGTGCTATTGATGTGTGATGGAGGC 1162

 Consensus A-G-G-G--GAA-C---AC--A--T-GTA-GGGG---ACCTTC--A-CA-GG--CTC 1254
jumonjinen ACG-GTGAAGAACCCG-ACG-ACGTCGTACGGGGAGTACCTTC--AGCACGGC-CTC 1011
Jumonshuff AGGCGCGGGGAATCGTCACATATTTGGTA-GGGGCC-ACCTTCTTAACAGGGATCTC 1217

 Consensus ----TGCGC-GA-C---A-G----GC--G---T-AGCAGGGTAAGCACATTCGCTCG 1311
jumonjinen CG--TGCGCTGATCGTCAAGAACGGCCAGACCTGA---------------------- 1044
Jumonshuff TAACTGCGCAGACCCAGATGTTGTGCTGGTAGTTAGCAGGGTAAGCACATTCGCTCG 1274

 Consensus TGTCGGTCTTGCAGCGGCACCTCGTGGTGTGGTTAGAGAGAGATAGA           1358
jumonjinen ------------------------------------------           1044
Jumonshuff TGTCGGTCTTGCAGCGGCACCTCGTGGTGTGGTTAGAGAGAGATAGA           1321
```

Figure 6.19. Nucleotide sequence alignment of shuffled and unshuffled *S. jumonjinesis*.

Highly conserved regions are shown in bold in the consensus sequence data.

6.4 DISCUSSION

The sequence data obtained for the four strains sequenced; *S. hygroscopicus*, *S. tanashiensis*, and two novel thermophiles Thermophilic A and B, suggests the amplified genes to be an IPNS genes. Comparison with IPNS genes from *Streptomyces spp,* obtained from Genbank, indicated a high identity for each of the strains sequenced. Thermophilic A and Thermophilic B showed 72% and 65% identity to *S. cattleya* respectively. Identities amongst the *Streptomyces* IPNS genes ranges from 70 - 85%. Therefore the Thermophilic A IPNS gene shows high identity to *Streptomyces*, whereas Thermophilic B shows a lower identity to the expected. The sequence obtained for *S. tanashiensis* showed a 70% identity to *S. jumonjinensis* and the similarity was noted to be highly conserved at the beginning of the gene, as well as in the mid-region at 400bp. *S. hygroscopicus* sequence data was efficient with sharp detection peaks (Figure 6.1), indicating strong primer complementarity and therefore specificity, however the identity was lower than expected at 60%.

The sequence data for *S. hygroscopicus* (Figure 6.2) against other IPNS sequences obtained from the database show that the sequence was good in some parts with the T and A being the most conserved of the nucleotides. The sequence for *S. hygroscopicus* was found to be 64.6%, 62.0%, 62.2%, 62.0% identical to the IPNS sequences for *S. cattleya*, *S. clavuligerus*, *S. hygroscopicus subspecies hygroscopicus*, and *S. fimbriae* respectively (Figure 6.3). However, this was significantly lower than the expected 70-85% identity. The failure to obtain reading frames for these genes may be explained by high levels of sequencing errors due to only single direction sequencing and a relatively poor output in a number of cases. The sequencing errors in *S. hygroscopicus* gene may be due to low primer specificity. This explanation is supported by the

observation that many sequencing attempts resulted in DNA sequence that was found to be identical to the pGEM-T vector, DNA to which the primers were not specific. Hence, pGEM-T vector contamination of the sequencing *S. hygroscopicus* sample may have decreased the accuracy of the sequence data.

The open reading frames were predicted for *S. hygroscopicus*, *S. tanashiensis*, Thermophilic A and Thermophilic B. The sequence data contained two, six, three and no predicted open reading frames respectively. Since no antibiotics were produced by Thermophilic A before the gene was expressed in an *E. coli* pTrcHis2-TOPO expression system, this gene in Thermophilic A is believed to be either non-functional or actively repressed. The other clones showed functionality and it is believed that definitive sequencing which will take place in the near future will show good open reading frames.

The sequence data obtained for the clone post shuffling from *S. jumonjinensis* showed a 77% identity to the sequence of *S. jumonjinensis*, obtained from Genbank with accession number M36687. Again, we are convinced that definitive sequencing will show a closer similarity and gain us some insight into the mutations that increased the IPNS activity in this clone. There are regions which seem to be highly conserved between the two sequences especially from about 850 to 1140bp (Figure 6.15). This shows that the IPNS gene is highly conserved, with shuffling resulting in accelerated beneficial mutations due to the observation of an increase in IPNS activity after shuffling. The alignment shows regions that are highly conserved, confirming that the *Streptomyces* IPNS gene is conserved. The GC content remains fairly constant even after shuffling. There are however sections of what seems to be inserted DNA in the shuffled *S.*

jumonjinensis gene, which may again be explained by low primer specificity causing repetitive generation of DNA during the shuffling process.

Due to time and equipment restraint only one DNA strand was sequenced and future work would include sequencing of the complimentary strand resulting in meaningful sequence data. It also needs to be emphasised that the sequencing was not of high quality in a number of cases and this needs to be improved.

CHAPTER 7

FINAL DISCUSSION

This study has provided an insight into the effect of gene shuffling of the IPNS gene with the resultant increase in IPNS activity and hence an increase in antibiotic activity. Error-prone PCR was used to create point mutations at each round of amplification. This method has the advantage of simplicity and ease of use. However, it is limited by difficulties in identification of critical regions for mutagenesis, as well as the practical need to reduce randomly chosen clones, from unlimited libraries to manageable numbers. Hence, the number of gene products exhibiting increased IPNS activity is limited.

Storage of the *Streptomyces* cultures did not affect IPNS gene function, facilitating further characterization of the IPNS gene. Conventional procedures for the isolation of *Streptomyces* DNA were successfully optimised, such that DNA yields were high and of sufficient purity for subsequent molecular manipulations.

Competitive growth bioassays using methods such as agar plug and the cup plate methods were efficient in demonstrating the antimicrobial activity phenotype of the *Streptomyces* cultures against *M. luteus*, *S. aureus* and *E. coli* ESS. However, this technique allowed for detection of antimicrobial activity only, without specific identification and characterization of the antimicrobial agents produced by the *Streptomyces* cultures. No antimicrobial activity was detected for Thermophilic A and Thermophilic B against *M. luteus*, *S. aureus* and *E. coli* ESS. However, after IPNS gene shuffling of Thermophilic A and Thermophilic B, similar competitive growth bioassays successfully demonstrated antimicrobial activity phenotype of both the IPNS

gene-shuffled.

In order to PCR-amplify the IPNS gene for eight *Streptomyces spp* and two novel thermophiles, primers specific enough for PCR-amplification, yet sufficiently degenerate for use on all ten strains, had to be designed. Furthermore IPNS sequence data was only available for *S. clavuligerus*, *S. cattleya*, *S. griseus*, *S. jumonjinensis*, *S. lipmanii*, and *Amycolatopsis spp* (six strains in total). However, primer-specificity was found to be sufficient for the PCR-amplification of the ten strains, since PCR-products ranging in size from 840 to 1200 bp, were successfully obtained for *S. clavuligerus*, *S. cattleya*, *S. griseus*, *S. jumonjinensis*, *S. lipmanii*, and *Amycolatopsis spp*, *S. hygroscopicus*, *S. tanashiensis*, Thermophilic A and Thermophilic B. Therefore, not only was sequencing of these PCR products possible, but also a degree of conservation of the IPNS gene of all these strains was indicated.

The PCR products for all ten strains were cloned into the pGEM-T and pTrcHis2-TOPO vectors for amplification and expression of the DNA in *E. coli* (α DH5) and TOP 10F respectively. IPNS-gene shuffling products of *S. hygroscopicus*, *S. tanashiensis*, Thermophilic A and Thermophilic B were similarly cloned into pGEM-T and pTrcHis2-TOPO.

E. coli TOP 10F transformed with p-TrcHis2-TOPO PCR product constructs were cultured and the PCR product DNA expressed by IPTG induction. Induced *E. coli* (TOP 10) cells were lysed and cell-free extracts analysed by HPLC. HPLC peaks were found after analysis of the induced cell-free extracts which corresponded to the retention time of the penicillin standard peak. No such peak was found in the HPLC analysis of non-transformed *E. coli*. Therefore, penicillin was successfully produced in *E. coli* (TOP 10F) transformed with p-TrcHis2-TOPO containing IPNS

cDNA, indicating that the PCR-amplified cDNA within the pTrcHis2-TOPO constructs was transcribed and translated in the *E. coli* transformants to enable them to produce penicillin. Furthermore, the shuffled gene expression and the IPNS activity were shown to be greater than the original IPNS genes by SDS-PAGE gel electrophoresis analysis, total protein determination and HPLC.

Sequencing of the IPNS gene of *S. tanashiensis* and *S. hygroscopicus* and alignment of this data to documented *Streptomyces* IPNS gene data (*S. cattleya, S. clavuligerus, S. fimriatus* and *S. hygroscopicus* subspecies *hygroscopicus*), indicated greater than 70% identity. This suggest a high degree of conservation in the IPNS gene within the *Streptomyces* genus. The IPNS genes of Thermophilic A and Thermophilic B were similarly sequenced and aligned with the documented IPNS gene sequence data as before. These two novel thermophiles, Thermophilic A and Thermophilic B demonstrated a 72% and 65% identity to *Streptomyces* IPNS genes respectively. However, due to time constraints and set-up difficulties of the newly installed departmental sequencer, only one strand of the DNA was sequenced. Therefore, in order to increase the reliability of this sequencing data, further work would include the sequencing of the complimentary strand.

In summary, an IPNS like gene has ben confirmed in the novel Thermophilic A and Thermophilic B. These IPNS genes were made fully functional after shuffling, indicating that originally these IPNS genes were silent or actively repressed. Furthermore, gene shuffling has been shown to improve both IPNS activity and production in the other strains used.

APPENDICES

APPENDIX A

CULTURE MEDIA

A.1 IPNS solution

A.2 ISP_2

A.3 LB agar

A.4 M_3 agar

A.5 M_3 broth

A.6 M_{65} agar

A.7 Nutrient agar

A.8 R_3

A.9 R_4

A.10 Sporulation agar

A.11 Terific broth

APPENDIX B

STANDARD SOLUTIONS AND BUFFERS

B.1 Chloroform: Isoamyl alcohol

B.2 CTAB buffer

B.3 CTAB/NaCl solution

B.4 EDTA 0.5M (pH 8.0)

B.5 Etbr

B.6 Gel loading buffer 6X

B.7 Lysozyme

B.8 NaCl 5M

B.9 Phenol/Chloroform/Isoamyl alcohol

B.10 Proteinase K

B.11 RNase A

B.12 SDS 10%

B.13 Sodium acetate 3M

B.14 TAE 50X

B.15 TBE 10X

B.16 TE buffers

APPENDIX C

BIOASSAYS OF BETA-LACTAM COMPOUNDS AND HPLC ANALYSIS

C.1 ACV

C.2 6-Aminopenicillanic acid

C.3 DTT

C.4 IPNS assay buffer

C.5 IPNS wash buffer

C.6 PMSF

C.7 Potassium phosphate buffer

C.8 Tris-HCl (pH 7.0)

APPENDIX D

PCR REAGENTS

D.1 Acetamide 4%

D.2 Agarose with Etbr (1.0%)

D.3 BSA

D.4 DNase I

D.5 dNTPs

D.6 Primers

D.7 Sybr green

D.8 Taq DNA Polymerase

APPENDIX E

TRANSFORMATION AND EXPRESSION STUDIES

E.1 Acrylamide Stock

E.2 Bath Buffer Stock

E.3 Coomasie (Staining solution)

E.4 CTAB 5%

E.5 Destaining Solution

E.6 Dissociation Buffer

E.7 Glucose 2M

E.8 IPTG 1M

E.9 Mg^{2+} stock

E.10 Resolving Gel Buffer

E.11 R_F 1

E.12 R_F 2

E.13 Stacking Gel Buffer

E.14 SOC medium

E.15 Stet

E.16 XGal

All media and solutions used in this study were autoclaved at 121^{0}C for 15 minutes

A.1 IPNS solution

Malt extract 1%	10g
Bacto peptone 1%	10g
Glycerol 2%	20g

Distilled H_2O to 1000ml

pH was adjusted to 7.0 with NaOH before autoclaving

A.2 ISP_2 media

Bacto Yeast extract	4g
Bacto Malt extract	10g
Bacto Dextrose	4g
Agar Agar	20g

Distilled H_2O to 1000ml

A.3 LB media

Bacto Tryptone	10g
Bacto Yeast extract	5g
Sodium chloride	10g
Agar	15g
Distilled H_20	950ml

pH was adjusted to 7.0 with 5M NaOH before addition of agar

Final volume was adjusted to 1000ml

A.4 M_3 agar

Malt extract	24g
Yeast extract	5g
Agar Agar	15g

Distilled H_20 to 1000ml

A.5 M_3 broth

Bacto Malt extract	24g
Bacto Yeast extract	5g
Sucrose	200g

Distilled H_2O to 1000ml

A.6 M_{65} (*Streptomyces* media)

Glucose	4g
Yeast extract	4g
Malt extract	10g

Calcium carbonate	2g
Agar Agar	12g

pH was adjusted to 7.2 using KOH before addition of the agar

Volume was adjusted to 1000ml with distilled H_2O

A.7 Nutrient agar

Peptone	5g
Meat extract	3g
Agar	15g

Distilled H_2O to 1000ml

A.8 R_3 solution

Glucose	10g
Yeast extract	5g
Casamino acids	100mg
Proline	3g
$MgCl_2.6H_20$	10g
$CaCl_2.2H_2O$	4g
K_2SO_4	200mg
KH_2PO_4	50mg
TES	5.6g

Trace elements were also added to R_3 and R_4 solutions

$ZnCl_2$	40mg
$CuCl_2.2H_2O$	10mg

$MnCl_2.4H_2O$	10mg
$FeCl_2.6H_2O$	200mg
$Na_2B_4O_7.10H_2O$	10mg
$(NH_4)6Mo_7O_{24}.4H_2O$	10mg

Distilled H_2O to 1000ml

A.9 R_4 solution

Glucose	10g
Yeast extract	1g
Casamino acids	100mg
Proline	3g
$MgCl_2.6H_2O$	10g
$CaCl_2.2H_2O$	4g
K_2SO_4	200mg
TES	5.6g

Trace elements as in R_3 solution

Distilled H_2O to 1000ml

A.10 Sporulation agar

Yeast extract	4g
Malt extract	10g
Glucose	4g
Agar	20g

pH was adjusted to 7.3 prior to the addition of agar

A.11 Terific broth

Bacto Tryptone	12g
Bacto Yeast extract	24g
Glycerol	4ml
Distilled H_2O to 900ml	
Buffer medium	100ml
KH_2PO_4	0.17M
K_2HPO_4	0.72M

Buffer medium to maintain pH at 7.0

APPENDIX B

All solutions and buffers were autoclaved at 121°C for 15 minutes unless otherwise stated.

B.1 Chloroform : Isoamyl alcohol

24 : 1

The above solution was not autoclaved

B.2 CTAB solution (100ml)

CTAB	1g
NaCl	4.38g
Tris-base	0.61g
EDTA	0.29g
Distilled H_2O	90ml

pH was adjusted to 8.0 and the volume made to 100ml

B.3 CTAB/NaCl solution

NaCl 4.1g

CTAB 10g

Distilled H_20 to 100ml

B.4 EDTA 0.5M (pH 8.0)

EDTA.$2H_2O$ 186.1g

Distilled H_2O 800ml

pH was adjusted to 8.0 with NaOH pellets and volume adjusted to 1000ml

B.5 Ethidium Bromide (10mg/ml)

EtBr 1.0g

Distilled H_2O to 100ml

The Solution was wrapped in aluminium foil and stored in the dark at 4°C

B.6 Gel loading buffer (6X)

Bromophenol blue 0.025g

Xylene cyanol 0.025g

Sucrose 4.0g

Distilled H_2O to 10ml

The mixture was stirred thoroughly to allow the dye to mix properly

B.7 Lysozyme solution

Lysozyme 50mg

Tris-HCl 0.1M (pH 8.0) 1000µl

B.8 Sodium chloride (5M)

NaCl		292.2g
Distilled H2O	to	1000ml

B.9 Phenol / Chloroform / Isoamyl alcohol

25 : 24 : 1

Phenol	500.0g
Chloroform	500ml
Isoamyl alcohol	25ml
TE buffer	525ml

Solution was stirred to emulsify, phases were allowed to separate. The top phase was removed and TE addition was repeated. When phases had separated the top phase was removed leaving 1 cm of the aqueous phase and was stored at 4°C in a dark bottle.

B.10 Proteinase K

Proteinase K	20mg
Distilled H_2O	1000µl

100µl aliquots were prepared and stored at -20^{0}C

B.11 RNase A

RNase A	20mg
Tris-HCl 10mM (pH7.5)	1000µl

The solution was heated to 100^{0}C for 10min, cooled and aliquots of 100µl were stored at -20^{0}C

B.12 SDS (10%)

SDS (electrophoresis grade) 100g

Distilled H_2O 900ml

The solution was heated to 68^{0}C to assist dissolution and the pH was adjusted to 7.2 using HCl, followed by adjustment of the volume to 1000ml

B.13 Sodium acetate (3M)

Sodium acetate.$3H_2O$ 40.81g

Distilled H_2O to 80ml

The pH was adjusted to 5.2 with glacial acetic acid before making up the volume to 100ml

B.14 TAE 50X (250ml)

Tris base 60.5g

Glacial acetic acid 14.3ml

EDTA 0.5M (pH 8.0) 25ml

The volume was adjusted to 250ml before autoclaving

B.15 TBE 10X

Tris base 108g

Boric acid 55g

Distilled H_2O 900ml

EDTA 0.5M pH 8.0 40ml

The volume was adjusted to 1000ml

B.16 TE buffer

Tris-HCl 1M pH 8.0	10ml
EDTA 0.5M pH 8.0	2ml
Distilled H_2O	850ml

pH was adjusted to 8.0 with NaOH

The volume was adjusted to 1000ml before autoclaving

C.1 ACV

α-amino cysteinyl D-valine	1mg
Distilled H_2O	1000µl

C.2

6-aminopenicillanic acid	50mg
Methanol	1000µl

A drop of concentrated HCl was added to acidify enabling the antibiotic to go into solution

C.3 Dithiothreitol 1M

DTT	3.09g

Sodium acetate 0.01M (pH 5.2) 20ml

Solution was sterilized by filtration dispensed into 1ml aliquots and stored at -20^{0}C

C.4 IPNS assay buffer (1000ml) final concentrations of ingredients used

Tris-HCl pH 7.5	50mM
DTT	1.0mM
EDTA	0.01mM
Glycerol	100ml

C.5 IPNS wash buffer (1000ml)

Tris-HCl (pH 7.5)	50mM
DTT	0.1mM
EDTA	0.01M
Glycerol	100ml

C.6 Phenylmethyl-sulphonyl fluoride (10mm)

PMSF	1.74mg
Isopropanol	1000µl

C.7 Potassium phosphate buffer (1000ml)

KH_2PO_4	12.52g
K_2HPO_4	1.39g
Distilled H_2O to	1000ml

pH adjusted to 6.0 using NaOH (1M)

C.8 Tris-HCl pH 7.0 (1000ml)

Tris-HCl (1M)	948.8ml
Tris-base (1M)	51.2ml

D.1 Acetamide (4%)

Acetamide	2.0g
Distilled H_2O	50ml

D.2 Agarose 1%

Agarose	1.0g
TAE (1X)	100ml
EtBr (100mg/ml)	3.0µl

D.3 DNase I (0.02 Units)

DNase I	1.0µl
Final concentration was	0.002Unit/µl

D.4 Deoxyribonucleotides (dNTPs)100mM Stock

dATP	1.0µl
dCTP	1.0µl
dGTP	1.0µl
dTTP	1.0µl
Sterile Distilled H_2O	16.0µl

The above is diluted for PCR reaction to obtain 100µM

D.5	Primers	5.6µM
	Diluted to obtain	0.8µM

D.6 Sybr green (1000X)

Sybr green	1.0µl
DMSO	1.0ml

D.7 Taq (500U/µl)

Diluted to obtain 5U/µl for all the PCR reactions that were carried out

E.1 Acrylamide Stock

Acrylamide	75g
N,N-methylene bis-acrylamide	2.0g

E.2 Bath buffer stock

Tris	30.0g
Glycine	144.1g
SDS	10.0g

E.3 Coomasie (Staining solution)

Methanol	45.0ml
Acetic acid (glacial)	10.0ml
Coomasie Brilliant blue	0.2g
Distilled H_2O	45.0ml

E.4 CTAB 5%

Cetyl trimethyl ammonium bromide	5.0g
H_2O	100ml

E.5 Destaining solution (1000ml)

Methanol	450ml
Acetic acid	70ml
H_2O	480ml
Glycerol	100ml

E.6 Dissociation buffer (50ml)

SDS	5.0g
Mercaptoethanol	5.0ml
Glycerol	7.5ml
Bromophenol Blue (0.2%)	2.5ml
1M Tris pH 6.8 (stock)	6.3ml
H_20	28.7ml

E.7 Glucose 2M (10ml)

Glucose		3.6g
Sterile Milli Q H_20	to	10ml

E.8 IPTG stock solution (0.1M; 50ml)

IPTG	1.2g

Make up to final volume with distilled water. Filter-sterilize and store at 4°C

E.9 Mg^{2+} stock (2M; 100ml)

$MgCl_2.6H_2O$	20.33g
$MgSO_4.7H_2O$	24.65g

Make up to final volume with 100ml. Filter sterilize

E.10 Resolving Gel Buffer (1M Tris-HCl, pH 8.8; 500ml)

Tris	60.6g
HCl (conc)	7.3ml

Make up to final volume with distilled water

E.11 RF_1 solution (1000ml)

KCl (100mM)	7.456g
$MnCl_2$ (50mM)	9.8955g
CH_3COOK (30mM)	2.974g
$CaCl_2$	1.47g
Glycerol (15% m/v)	150g

Make up to final volume with distilled water. Adjust pH to 5.8. Filter-sterilize

E.12 RF_2 solution (1000ml)

MOPS (10mM)	2.093g
KCl	0.7456g
$CaCl_2$	11.025g
Glycerol (15%m/v)	150g

Make up to final volume with distilled water. Adjust pH to 6.8.

E.13 Stacking gel buffer (1M Tris-HCl; pH6.8; 500ml)

Tris	60.6g
HCl (conc)	41.0ml

Make up to final volume with distilled water

E.14 SOC medium

Tryptone	2.0g
Yeast extract	0.5g
NaCl (1M)	1.0ml
KCl (1M)	0.25ml

Mg^{2+} (2M)	1.0ml
Glucose (2M)	1.0ml

The tryptone, yeast extract, NaCl and KCl were add to 97ml distilled water, stirred to dissolve and autoclaved. Addition of Mg and glucose followed and thereafter the complete medium was filtered through a 0.2μm filter unit.

E.15 STET buffer

Sucrose (8% w/v)

Triton X-100 (0.1% v/v)

EDTA (50mM)

Tris-HCl (50mM, pH 8.0)

E.16 X-Gal (2ml)

5-bromo-4-chloro-3-indolyl-β-D-galactoside	100mg
N,N` - dimethyl-formamide	2ml

The X-Gal was dissolved , covered in aluminum foil and stored at -20^{0}C

Aceti D.J. and Champness W.C. (1998) Transcriptional regulation of *Streptomyces coelicolor* pathway-specific antibiotic regulators by the abs A and abs B loci. J. Bacteriol. **180**: 3100-3106.

Aharonowitz Y., Cohen G. and Martin J.F. (1992) Penicillin and cephalosporin biosynthetic genes : Structure, organization, regulation and evolution. Ann. Rev. Microbiol. **46**: 461- 495.

Aharonowitz Y., Av-Gay Y., Schreiber R. and Cohen G. (1993) Characterization of a broad-range disulfide reductase from *Streptomyces clavuligerus* and its possible role in β-lactam antibiotic biosynthesis. J. Bacteriol. **175**: 623-629.

Aidoo K.A., Wong A., Alexander D.C., Rittammer R.A. and Jensen S.E. (1994) Cloning sequencing and disruption of a gene from *Streptomyces clavuligerus* involved in clavulanic acid biosynthesis. Gene **147**: 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**: 4068- 4079.

Ambedkar S.S., Deshpande B.S. and Shewale J.G. (1997) Separation of cephalosporin C acylase and penicillin V acylase from *Aeromonas* ACY 95 by chemical treatment. Process Biochem. **32**: 305-307.

Baldwin J.E., Adlington R.M., Bryans J.S., Lloyd M.D., Sewell T.J. and Schofield C.J. (1997) Chemo-enzymatic synthesis of bicyclic γ-lactams using clavaminic acid synthase. Tetrahedron **53**: 7011-7020.

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

Barredo J.L. and Martin J.F. (1991) Genes directly involved in the biosynthesis of beta-lactam antibiotics. Microbiologia **7**: 1-12.

Bellgardt K.H. (1998) Process models for production of beta-lactam antibiotics. Adv. Biochem. Eng. Biotechnol. **60**: 153-194.

Bibb M., Schottel J.L. and Cohen S.N. (1980) A DNA cloning system for interspecies gene transfer in antibiotic-producing *Streptomyces*. Nature **284**: 526-530.

Bradford M.M. (1976) A rapid and sensitive method for quantification of microgram of protein utilizing the principle of protein dye binding. Anal. Biochem. **72**: 248-252.

Brakhage A.A., Browne P. and Turner G. (1992) Regulation of *Aspergillus nidulans* penicillin biosynthesis and penicillin biosynthesis genes *acvA* and *ipnA* by glucose. J. Bacteriol. **174**: 3789-3799.

Brakhage A.A. (1998) Molecular regulation of beta-lactam biosynthesis in filamentous fungi. Microbiol. Mol. Bio. Rev. **62**: 547-585.

Burzalaff N.J., Rutledge P.J., Clifton I.J., Hensgens C.M.H., Pickford M., Adlington R.M., Roach P.L. and Baldwin J.E. (1999) The reaction cycle of Isopenicillin N Synthase observed by X-ray diffraction. Nature **401**: 721-724.

Butler P.R., Brown M. and Oliver S.G. (1996) Improvement of antibiotic titres from *Streptomyces* bacteria by interactive continuous selection. Biotechnol. and Bioengineer. **49**:185-196.

Castro J.M., Liras P., Cortes J. and Martin J.F. (1985) Regulation of α-aminoadipyl-cysteinyl-valine, isopenicillin N synthetase, isopenicillin N isomerase and deacetoxycephalosporin C synthetase by nitrogen sources in *Streptomyces lactamdurans*. Appl. Microbiol. Biotechnol. **22**: 32- 40.

Castro J.M., Liras P., Laiz S., Cortes J. and Martin J.F. (1988) Purification and characterization of Isopenicillin N synthase of *Streptomyces lactamdurans*. J.Gen. Microbiol. **134**: 133-141.

Chater K.F. and Bruton C. J. (1983) Mutational cloning in *Streptomyces* and the isolation of antibiotic production gene. Gene **26**: 67-78.

Chhabra S.R., Stead P., Bainton N.J., Salmond G.P.C., Steward G.S.A.B., Williams P. and Bycroft B.W. (1993) Autoregulation of carbapenem biosynthesis in *Erwinia carotovora* by analogues of N-(3-oxohexanoyl)-L-homoserine lactone. J. Antibiot. **46**: 441- 454.

Christians A.C., Sapozza L., Crameri A., Folkers G. and Stemmer W.P.C. (1999) Directed evolution of thymidine kinase for AZT phosphorylation using DNA family shuffling. Nature Biotech. **17**: 259-264.

Cochet N. and Demain A.L (1996) Effect of water activity on production of beta-lactam antibiotics by *Streptomyces clavuligerus* in submerged culture. J. Appl. Bacteriol. **180**: 333-337.

Cohen G., Shiffman D., Mevarech M. and Aharonowitz Y. (1990) Microbial isopenicillin N synthase genes: Structure, function, diversity and evolution. Tibtech. **8**: 105-111.

Cohen G.,Yanko M., Mislovati M., Argaman A., Schreiber R., Av-GayY and Aharonowitz Y. (1993) Thioredoxin-thioredoxin reductase system of *Streptomyces clavuligerus* :Sequences, expression and organization of the genes. J. Bacteriol. **175**: 5159-5167.

Coque J.R., Liras P., Laiz L. and Martin J.F. (1991) A gene encoding lysine 6-aminotransferase which forms the β-lactam precursor α-aminoadipic acid is located in the cluster of cephamycin biosynthetic genes in *Nocardia lactamdurans*. J. Bacteriol. **173**: 6258-6264.

Cortes J., Liras P., Castro J.M. and Martin J.F. (1986) Glucose regulation of cephamycin biosynthesis in *Streptomyces lactamdurans* is exerted on the formation of α-Aminoadipyl-cysteinyl-valine and Deacetoxycephalosporin C Synthase. J. Gen. Microbiol. **132**: 1805-1814.

Courvalin P. and Davies P. (1999) Antimicrobial, Current opinion Microbiol. **2**: 465- 466.

Crameri A., Whitehorn E.A., Tate E. and Stemmer W.P.C. (1996) Improved green fluorescent protein by molecular evolution using DNA shuffling. Nat. Biotechnol. **14**: 315-319.

Crameri A., Dawes G., Rodriguez E. Jr., Silver S. and Stemmer W.P.C. (1997) Molecular evolution of an arsenate detoxification pathway by DNA shuffling. Nat. Biotechnol. **15**: 436- 438.

Crameri A., Raillard S., Bermudez E. and Stemmer Willem P.C. (1998) DNA shuffling of a family of genes from diverse species accelerates directed evolution. Nature **391**: 288-291.

Davies J. (1990) What are antibiotics? Archaic functions for modern activities. Mol. Microbiol. **4**: 1227-1232.

Demain A.L. (1991) Production of beta-lactam antibiotics and its regulation. Proc. Natl. Counc. Repub. China **15**: 251-265.

Diez B., Mellado E., Rodriguez M., Fouces R. and Barredo J. (1997) Recombinant microorganisms for industrial production of antibiotics. Biotech. and Bioengineering **55**: 216-226.

Doran J.L., Leskiw B.K., Aippersbach S. and Jensen S.E. (1990) Isolation and characterization of a ß-lactamase inhibitory protein from *Streptomyces clavuligerus* and cloning and analysis of the corresponding gene. J. Bacteriol. **172**: 4909- 4918.

Durairaj M. and Jensen S.E. (1996) Purification and characterization of recombinant *Streptomyces clavuligerus* isopenicillin N synthase produced in *Escherichia coli*. J. Industrial Microbiol. **16**: 197-203.

Durairaj M., Doran J.L. and Jensen S.E. (1992) High level expression of the *Streptomyces clavuligerus* Isopenicillin N synthase gene in *Escherichia coli*. Appl. Env. Microbiol. **58**: 4038-4041.

Durairaj M., Leskiw B.K. and Jensen S.E. (1996) Genetic and biochemical analysis of the cysteinyl residues of isopenicillin N synthase from *Streptomyces clavuligerus*. Can. J. Microbiol. **42**: 870- 875.

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

Fang A., Keables P. and Demain A. L.(1996) Unexpected enhancement of ß-lactam antibiotic formation in *Streptomyces clavuligerus* by very high concentrations of exogenous lysine. Appl. Microbiol. Biotechnol. **44**: 705-709.

Fang A. and Demain A.L. (1995) Dependence of Nitrogen and Phosphorus-regulation of β-lactam antibiotic production by *Streptomyces clavuligerus* on aeration level. J. Industr. Microbiol. **15**: 407- 410.

Floss H.G. (1987) Hybrid antibiotics- the contribution of the new gene combinations. Tibtech. **5**: 111-115.

Forsman M., Haggastronn B., Lindgren L. and Jaurin B. (1990) Molecular analysis of beta-lactamase from four species of *Streptomyces* : Comparison of amino acid sequences with those of other beta-lactamases. J. Gen. Microbiol. **136**: 589-598.

Garcia-Dominguez M., Liras P. and Martin J.F. (1991) Cloning and Characterization of the Isopenicillin N synthase gene of *Streptomyces griseus* NRRL 3851 and studies of expression and complementation of the cephamycin pathway in *Streptomyces clavuligerus*. Antimicrob. Agents Chemother. **35**: 44-52.

Gusek T.W. and Kinsella J.E. (1992)Review of the *Streptomyces lividans* / vector *pIJ702* system for gene cloning. Crit. Rev. Microbiol. **18**: 247-260.

Hansson L.O., Bolton-Grob R., Massoud T. and Mannervik B. (1999) Evolution of differential substrate specificities in Mu class glutathione transferases probed by DNA shuffling. J. Mol. Biol. **207**: 265-276.

Hao J. and Kendrick K.E. (1998) Visualization of penicillin-binding proteins during sporulation of *Streptomyces griseus*. J. Bacteriol. **180**: 2125-2132.

Harayama S. (1998) Artificial evolution by DNA shuffling. Tibtech **16**: 76- 82.

Hidaka T., Hidaka M., Uozumi T. and Seto H. (1992) Nucleotide sequence of a carboxyphosphoenolpyruvate phosphonomutase gene isolated from a bialaphos-producing organism, *Streptomyces hygroscopicus* and its expression in *Streptomyces lividans*. Mol. Gen. Genet. **233**: 476- 478.

Hopwood D.A. (1978) Extrachromosomally determined antibiotic production. Ann. Rev. Microbiol. **32**: 373-392.

Hopwood D.A (1988) Towards an understanding of gene switching in *Streptomyces*, the basis of sporulation and antibiotic production. Proc. R. Soc. Lon. **235**: 121-138.

Hopwood D.A., Chater K.F., Bibb M.J., Buttner M.J., Bruton C.J. and Kieser T. (1996) "Practical *Streptomyces* Genetics". John Innes Centre, UK.

Jensen S.E., Westlake D.W.S. and Wolfe S.E. (1982) High Performance Liquid Chromatographic assay of cyclization activity in cell-free systems from *Streptomyces clavuligerus*. J. Antibiot. **35**: 1026-1032.

Jensen S.E., Leskiw B.K., Vining L.C., Aharonowitz Y. and Westlake D.W.S. (1986) Purification of Isopenicillin N synthetase from *Streptomyces clavuligerus*. Can. J. Microbiol **32**: 953-958.

Keleman G.H. and Buttner M.J. (1998) Initiation of aerial mycelium formation in *Streptomyces*. Current Opinion in Microbiol. **1**: 656- 662.

Khetan A., Malmberg L.H., Sherman D.H. and Hu W.S. (1996) Metabolic engineering of cephalosporin biosynthesis in *Streptomyces clavuligerus*. Ann. N Y Acad. Sci. **782**: 17-24.

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

Kirby R. (1992) The isolation and characterization of antibiotic biosynthesis genes. Biotech **10**: 561-576.

Kovacevic S., Weigel B.J., Tobin M.B., Ingolia T.D. and Miller J.R. (1989) Cloning, characterisation and expression in *Escherichia coli* of the *Streptomyces clavuligerus* gene encoding deacetoxycephalosporin C synthetase. J. Bacteriol. **171**: 754-760.

Kovacevic S., Tobin M.B. and Miller J.R (1990) The β-lactam biosynthesis genes for isopenicillin N Epimerase and deacetoxycephalosporin C Synthatase are expressed from a single transcript in *Streptomyces clavuligerus*. J. Bacteriol. **172**: 3952-3958.

Kovacevic S. and Miller J.R. (1991) Cloning and sequencing of the β-lactam hydroxylase gene (cef F) from *Streptomyces clavuligerus*: Gene duplication may have led to separate hydroxylase and expandase activities in actinomycetes. J. Bacteriol. **173**: 398- 400.

Krallis M. (1997) Isolation and Identification of beta-lactam producing microorganisms using PCR based methodologies. MSc Thesis.

Kriauciunas A., Frolik C.A., Hassell T.C., Skatrud P.L., Johnson M.G., Holbrook N.L. and Chen V.J. (1991) The functional role of cysteine in Isopenicillin N Synthase. J.Biol.Chem. **266**: 11779-11788.

Kudo N., Kimura M., Beppu T. and Horinouchi S. (1995) Cloning and Characterization of a gene involved in aerial mycelium formation in *Streptomyces griseus*. J. Bacteriol. **177**: 6401-6410.

Laemmli U.K. (1970) Cleavage of structural proteins during the assembly of the head of the bacteriophage T4. Nat. **227**: 680-685.

Leitao L.A., Enguita F.J., Fuente J.L., Liras P. and Martin J.F (1999) Inducing effect of diamines on transcription of the cephamycin C genes from the lat and *pcbAB* promoters in *Norcardia lactamdurans*. J. Bacteriol. **181**: 2379-2384.

Leskiw B.K., Aharonowitz Y., Mevarech M., Wolfe S., Vining L.C., Westlake D.W. and Jensen S.E. (1988) Cloning and nucleotide sequence determination of the Isopenicillin N synthetase gene from *Streptomyces clavuligerus*. Gene **62**: 187-196.

Liu Y., Ji D., Chou C., Chao H., Liao C. and Han S. (1994) Improvement of penicillin G Acylase production by using *pac* gene clone. Proceedings Nat. Sci. Council ROC **18**: 95-100.

Loke P. and Sim T (1998) Analysis of glutamines in catalysis in *Cephalosporium acremonium* Isopenicillin N synthase by site-directed mutagenesis. Biochemical and Biophysical Research Communications **252**: 472- 475.

Loke P. and Sim T. (1998) Mutational evidence for the role of serine-283 in *Cephalosporium acremonium* isopenicillin N synthase. FEMS **165**: 353-356.

Loke P., Chee P. N. and Sim T. (2000) PCR cloning, heterologous expression and characterization of IPNS from *Streptomyces lipmanii*: NRRL 3584. Can. J. Microbiol. **46**: 166-170.

Ludovice M., Martin J.F., Carrachas P. and Liras P. (1992) Characterization of the *Streptomyces clavuligerus* argC gene encoding N-acetylglutamyl-phosphate reductase: Expression in *Streptomyces lividans* and effect on clavulanic acid production. J. Bacteriol. **174**: 4606- 4613.

Madduri K., Stuttard C. and Vining L.C. (1989) Lysine catabolism in Streptomyces spp is primarily through cadaverine : β-lactam producers also make α-aminoadipate. J. Bacteriol. **171**: 299-301.

Madduri K., Stuttard C. And Vining L.C. (1991) Cloning and location of a gene governing lysine ε-aminotransferase, an enzyme initiating β-lactam biosynthesis in *Streptomyces spp*. J. Bacteriol. **173**: 985-988.

Martin J.F. and Gutierrez S. (1995) Genes for beta-lactam antibiotic biosynthesis. Antonie Van Leeuwenhoek 67: 181-200.

Martin J.F. (1998) New aspects of genes and enzymes for beta-lactam antibiotic biosynthesis. Appl. Microbiol. Biotechnol. **50**: 1-15.

Martin J.F. and Liras P. (1985) Biosynthesis of ß-lactam antibiotics: design and construction of overproducing strains. Trends in Biotechnol. **3(2)**: 39- 44.

McGowan S.J., Sebaihia M., Porter L.E., Stewart G.S.A.B., Williams P., Bycroft B.W. and Salmond G.P.C (1996) Analysis of bacterial carbapenem antibiotic production genes reveals a novel β-lactam biosynthesis pathway. Mol. Microbiol. **22**: 415- 426.

McGowan S.J., Bycroft B.W. and Salmond G.P.C. (1998) Bacterial production of carbapenems and clavams: evolution of β-lactam antibiotic pathways. Trends in Microbiol. **6**: 203-208.

Moore J.C., Jin H., Kuchner O. and Arnold F.H. (1997) Strategies for the *in vitro* evolution of protein function: Enzyme evolution by random recombination of improved sequences. J. Mol. Biol. **272:** 336-347.

Niessen W.M.A. (1998) Analysis of antibiotics by liquid chromatography-mass spectrometry. J. Chromatog. A. **812**: 53-75.

Orford C.D., Perry D. and Adlard M.W (1991) The determination of naturally produced penicillins and their biosynthetic precursors using pre-column derivitisation with dansylaziridine. J. Liq. Chromatog. **14**: 2665-2684.

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**: 1307-1314.

Paradkar A.S., Aidoo K.A., Wong A. and Jensen S.E. (1996) Molecular analysis of a beta-lactam resistance gene encoded within the cephamycin gene cluster of *Streptomyces clavuligerus*. J. Bacteriol. **178**: 6266-6274.

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.

Patten P.A., Howard R.J. and Stemmer W.P.C. (1997) Applications of DNA shuffling to pharmaceuticals and vaccines. Current opinion in Biotech. **8**: 724-733.

Penalva M.A., Rowlands R.T.and Turner G. (1998) The optimization of penicillin biosynthesis in fungi. Tibtech. **16:** 483-489.

Perez-Llarena F.J., Liras P., Rodriguez-Garcia A. and Martin J.F. (1997) A regulatory (ccaR) required for cephamycin and clavulanic acid production in *Streptomyces clavuligerus* : amplification results in overproduction of both beta-lactam compounds. J. Bacteriol. **179**: 2053-2059.

Perez-Llarena F.J., Rodriguez-Garcia A., Enguita F.J., Martin J.F. and Liras P. (1998) The *pcd* gene encoding piperideine-6-carboxylate dehydrogenase involved in biosynthesis of α-aminoadipic acid is located in the cephamycin cluster of *Streptomyces clavuligerus*. J. Bacteriol. **180**: 4753-4756.

Pertursson S. and Baldwin J.E. (1998) Synthesis of δ-(L-α-Aminoadipoyl)-L-cysteinyl-D-(O-methyl)-D- allothreonine a substrate for Isopenicillin-N-synthase and its O-methyl-D-threonine Epimer. Tetrahedron **54**: 6001-6010.

Petrich A., Wu X., Roy K.L. and Jensen S.E. (1992) Transcriptional analysis of the Isopenicillin N synthase-encoding gene of *Streptomyces clavuligerus*. Gene **111**: 77- 84.

Peruski L.F. Jr. and Peruski A.H. (1997) The internet and the new biology. 206-208.

Ramos F.R., Lopez-Nieto M.J. and Martin J.F. (1985) Isopenicillin N Synthetase of *Penicillium chrysogenum*, an Enzyme that converts δ-(L-α-Aminoadipyl)-L-Cysteinyl-D-Valine to Isopenicillin N. Antimicrob. Agents & Chemother. **27**: 380-387.

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**: 700-704.

Romero J., Liras P. and Martin J.F (1984) Dissociation of cephamycin and clavulanic acid biosynthesis in *Streptomyces clavuligerus*. Appl. Microbiol. Biotechnol. **20**: 318-325.

Rodriguez-Garcia A., Martin J.F. and Liras P. (1995) The argG gene of *Streptomyces clavuligerus* has low homology to unstable argG from other actinomycetes: effect of amplification on clavulanic acid biosynthesis. Gene **167**: 9-15.

Ryu D.D. and Nam D.H. (2000) Recent progress in biomolecular engineering. Biotechnol. Prog. **16**: 2-16.

Salas J.A. and Mendez C. (1998) Genetic manipulation of antitumour-agent biosynthesis to produce novel drugs. Tibtech **16**: 475-482.

Sambrook J., Fritch E.F. and Maniatis T. (1989) Molecular cloning: A laboratory manual 2nd ed., **Vol 1-3** Cold Spring Harbor laboratory press, Cold Spring Harbor N.Y.

Samson S.M., Belagaje R., Blankenship D.T., Chapman J.L., Perry D., Skatrud P.L., Van Frank R.M., Abraham E.P., Baldwin J.E. and Queener S.W. (1985) Isolation, sequence determination and expression in *Escherichia coli* of the isopenicillin N synthetase gene from *Cephalosporium acremonium*. Nat. **318**: 191-194.

Sanchez L. and Brana A.F. (1996) Cell density influences antibiotic biosynthesis in *Streptomyces clavuligerus*. Microbiol. **142**: 1209-1220.

Schena M. (1999) DNA miroarrays: A practical approach, 139-142.

Schugerl K. and Seidal G. (1998) Monitoring of the concentration of β-lactam antibiotics and their precursors in complex activation media by High-performance liquid chromatography. J. Chromatogra. **812**: 179-189.

Shiffman D., Mevarech M., Jensen S.E., Cohen G. and Aharonowitz Y. (1988) Cloning and Comparative sequence analysis of the gene coding for Isopenicillin N. Synthase in *Streptomyces*. Mol. Gen. Genet. **214**: 562-569.

Shima J., Hesketh A., Susumo H., Kawamoto S. and Ochi K. (1996) Induction of Actinorhodin production by rps L (encoding Ribosomal protein S12) Mutations that confer streptomyces resistance in *Streptomyces lividans* and *Streptomyces coelicolor* A3(2). J. Bacteriol. **178**: 7276-7284.

Sim TS and Tan D.S.H. (1995) Biophysical factors affecting the expression of Isopenicillin N Synthase isozymes in *E. coli*. Biochem. Mol. Biol. Internat. **35**: 1069-1078.

Skatrud P.L. (1992) Genetic engineering of beta-lactam antibiotic biosynthetic pathway in filamentous fungi. Trends Biotechnol. **10**: 324-329.

Stemmer W.P.C. (1994) Rapid evolution of a protein *in vitro* by DNA shuffling. Nature **370**: 389-391.

Stemmer W.P.C. (1994) DNA shuffling by random fragmentation and reassembly: In vitro recombination for molecular evolution. Proc. Natl. Acad. Sci. USA **91**: 10747-10751.

Sun F.Z. (1999) Modeling DNA shuffling. J. Computational Biol. **6**: 77-90.

Tan D.S.H. and Sim T. (1996) Functional analysis of conserved histidine residues in *Cephalosporium acremonium* Isopenicillin N synthase by site-directed metagenesis. J. Biol. Chem. **271**: 889-894.

Townsend C.A. (1993) Oxidative amino acid processing in β-lactam antibiotic biosynthesis. Biochem. Soc. Transactions **21**: 208-213.

Valegard K., Anke C., Scheltinga T., Lloyd M.D., Hara T., Ramaswamy S., Perrakis A. Thomson A., Lee H., Baldwin J.E., Schofield C.J., Hajdu J. and Andersson I. (1998) Structure of a cephalosporin synthase. Nature **394**: 805- 809.

Van Liempt H., Dohren H. and Kleinkauf H. (1989) δ-(L-α-aminoadipoyl)-L- cysteinyl-D-valine Synthatase from *Aspergillus nidulans*. J. Biol. Chem. **264**: 3680-3684.

Veilleux C. (1999). PCR technology. http://www.accessexcellence.org

Walter S., Wellmann E. and Schrempf H. (1998) The cell wall-anchored *Streptomyces reticuli* Avicel-binding protein (AbpS) and its gene. J. Bacteriol. **180**: 1647-1654.

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 **110**: 239-242.

Williamson J.M., Inamine E., Wilson K.E., Douglas A.W., Liesch J.M. and Albers-Schonberg G. (1985) Biosynthesis of the β-lactam antibiotic, thienamycin, by *Streptomyces cattleya*. J. Biological Chem. **260**: 4637- 4647.

Wolfe S. and Jokinen M.G. (1979) Total synthesis of δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine (ACV), a biosynthesis precursor of penicillins and cephalosporins. Can. J. Chem. **57**: 1388-1396.

Wolfe S., Demain A.L., Jensen S.E. and Westlake D.W. (1984) Enzymatic approach to synthesis of unnatural beta-lactams. Science **226**: 1386-1392.

Yu H., Serpe E., Romero J., Coque J. J., Maeda K.,Oelgeschlager M., Hintermann G., Liras P., Martin J.F. and Demain A.L. (1994) Possible involvement of the lysine epsilon-aminotransferase gene (lat) in the expression of genes encoding ACV synthetase (pcbAB) and isopenicillin N synthase (pcbC) in *Streptomyces clavuligerus*. Microbiology **140**: 3367-3377.

Zhang J., Dawes G. and Stemmer W.P.C. (1997) Directed evolution of a fucosidase from galactosidase by DNA shuffling and screening. Proc. Natl. Acad. Sci. USA **94**: 4504-45.

Zhao H. and Arnold F.H. (1997) Optimization of DNA shuffling for high fidelity recombination. Nucleic Acid Research **25**: 1307-1308.

http://www.ncbi.nlm.nih.gov

http://www.promega.com

http://www.invitrogen.com

http://www.yahoo.com