INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI®

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

The Biosynthesis of the Thiopeptide Antibiotic Thiostrepton

Paul R. Shipley

A dissertation submitted in partial fulfillment

of the requirements for the degree of

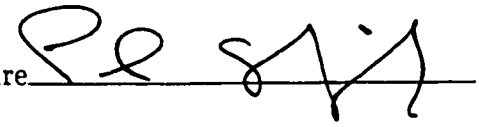
Doctor of Philosophy

University of Washington

1999

Program Authorized to Offer Degree: Chemistry

In presenting this dissertation in partial fulfillment of the requirements for the Doctorial degree at the University of Washington, I agree that the Library shall make its copies freely available for inspection. I further agree that extensive copying of this thesis is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for copying or reproduction of this dissertation may be referred to University Microfilms, 1490 Eisenhower Place, P.O. Box 975, Ann Arbor, MI 48106, to whom the author has granted "the right to reproduce and sell (a) copies of the manuscript in microform and/or (b) printed copies of the manuscript made from microform."

Signature 

Date November 3, 1999

University of Washington

Graduate School

This is to certify that I have examined this copy of a doctoral dissertation by

Paul R. Shipley

and have found that it is complete and satisfactory in all respects,

and that any and all revisions required by the final

examining committee have been made.

Chair of Supervisory Committee:

Heinz G. Floss

Reading Committee:

Michael H. Gelb

Tomikazu Sasaki

James T. Staley

William F. Trager

Date: November 1, 1999

University of Washington

Abstract

The Biosynthesis of the Thiopeptide Antibiotic Thiostrepton

by Paul R. Shipley

Chair of Supervisory Committee

Professor Heinz G. Floss
Chemistry

Thiostrepton is a sulfur rich, highly modified peptide antibiotic produced by the soil bacteria *Streptomyces laurentii*, *Streptomyces azureus*, and *Streptomyces hawaiiensis*. It is a member of the structurally similar thiopeptide antibiotic family. Thiostrepton has been found to show strong antibiotic activity towards Gram-positive bacteria by inhibition of ribosomal protein biosynthesis. Thiostrepton is used as a veterinary antibiotic, and as a selective marker in Streptomycete research. Previous studies on this system have elucidated the amino acid precursors of thiostrepton, demonstrated that the peptide likely is synthesized nonribosomally, and found that the biosynthetic genes are not clustered with the known gene conferring resistance to

the antibiotic. Three approaches to finding the thiostrepton biosynthetic gene cluster in *S. laurentii* are presented in this study. Attempted first was reverse genetics method in which a pathway enzyme was chosen for purification in order to obtain amino acid sequence information for use in designing oligonucleotides for Southern hybridization experiments. The enzyme was not successfully purified, due to its low stability and concentration in cell-free extract. Second, specific hybridization to restriction endonuclease digested genomic DNA from three thiopeptide producers was not observed using a Southern probe sythesized from a gene encoding an enzyme catalyzing nonribosomal peptide synthesis from a related species. A novel approach developed by Marahiel and co-workers was undertaken. In this approach, PCR primers are designed based on conserved amino acid sequences found in nonribosomal peptide synthetases. These primers were used to amplify DNA fragments from nonribosomal peptide synthetases in *S. laurentii*. A cluster of these enzymes was found by this method, partially characterized, and proven to not be involved in thiostrepton biosynthesis by single crossover integrational mutants in that region. The experiment was repeated, and two other clusters of nonribosomal peptide synthetases were isolated and demonstrated to not be located near the known cluster, or to each other. Whether these are involved in thiostrepton biosynthesis remains an open question.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

ACVS: δ-(*L*-α-aminoadipyl)-*L*-cysteinyl-bis-*D*-valine synthetase

AMP: Adenosine diphosphate

ATCC: American Type Culture Collection, Rockville, MD

ATP: Adenosine triphosphate

BP: Base pairs

CPM: Counts per minute

DA: Daltons

DEAE: Diethylaminoethyl

DNA: Deoxyribonucleic acid

DTT: Dithiothreitol

FPLC: Fast protein liquid chromatography

HEQ: 4-(1-Hydroxyethyl)quinoline-2-carboxylate

HPLC: High performance liquid chromatography

KB: Kilobases

KDA: Kilodaltons

MNNG: N-methyl-N'-nitro-N-nitrosoguanidine

NT: Nucleotides

PEG: Polyethylene glycol

PP_I: Pyrophosphate

PVDF: Polyvinyl difluoride

RNA: Ribonucleic acid

SDS: Sodium dodecyl sulfate

SDS-PAGE: Sodium dodecyl sulfate - polyacrylamide gel electrophoresis

TCA: Trichloracetic acid

TLC: Thin layer chromatography

TRIS: Tris-(hydroxymethyl)aminomethane

ACKNOWLEDGMENTS

I would like to first thank Professor Heinz Floss for allowing me to work on a project that let me explore such a diversity of different fields in one thesis project, for his advice and support, for his considerable help in preparing this thesis, and for always finding the time to talk when I knocked on his door. I would like to also thank several people who have helped over the years, including Dr. Todd Smith and Professor Nigel Priestley, who worked on this project previous to my joining the group, and who took the time to get me started on it. Ya-Fen Jiang worked beside me on this project in its early stages, and did much of the work presented in the mapping and partial sequencing of the first cluster of nonribosomal peptide synthetases. Dr. Paul August and Dr. Tin-Wein Yu gave valuable advice on the molecular biology experiments presented in this study, and Professor Yu-Sam Kim devoted considerable time and energy into teaching me the subtleties of protein purification. A heartfelt thanks is due to Kay Kampsen for all of her help, conversation, and friendship. Many thanks also to all of the members of the Floss research group for their friendship and support. My deepest gratitude to my parents and grandparents for their love, companionship, faith in my success, and brownies throughout my many years of college education. There are not words to express my thanks to my wonderful wife, Dr. Rebecca Tyson, who has provided love, understanding, respite, caring, and a

prolific red pen to my thesis writing efforts. This thesis is dedicated to the memory of Donna Beth Richmond.

Chapter 1

INTRODUCTION

Thiostrepton (figure 1.1) is a member of the family of sulfur-rich peptide natural products known as the thiopeptide antibiotics. It is produced by the soil bacteria *Streptomyces laurentii* [71], *Streptomyces azureus* [76], and *Streptomyces hawaiiensis* [10]. Other members of this family of antibiotics include nosiheptide [4] (figure 1.2), the siomycins [70], the thiopeptins [21] (figure 1.3), the micrococcins [77], the thiocillins [61] (figure 1.4), glycothiohexide α [44], cyclothiazomycin [3], GE2270 A [27] (figure 1.5), A10255B, -G, and -J [14], and Sch 40832 [54]. Thiostrepton has been found to show strong antibiotic activity against Gram-positive bacteria [50][67]. Thiostrepton is used in veterinary medicine as a topical antibiotic (Panalog®, SOLVAY Animal Health, Inc. Mendota Heights, MN 55120-1139) [42], but is limited in its usefulness as an effective antibiotic by its low solubility in water which greatly reduces its bioavailability. Recently, studies have been reported which show that thiostrepton inhibits protein synthesis in the blood stage of the malaria parasite at low concentrations [35]. The thiostrepton resistance gene from *S. azureus* was one of the first Streptomycete genes to be cloned [69], and is widely used as a selectable marker in cloning vectors designed to be used in streptomycetes [22].

A defining characteristic of the thiopeptide antibiotics are the thiazole and dihydrothiazole ring structures formed presumably by the condensation of a cysteine residue with the carboxyl carbon of the following amino acid. The major macrocycle found in all of the members of this family, which differentiates it from other families of thiazole containing antibiotics, is formed by the creation of a pyridine, or tetrahydropyridine, ring structure. This macrocycle contains a variable number of thiazole, thiazoline, or oxazole rings, as well as several modified amino acids such as dehydrobutyrine and dehydroalanine. In some of the members of this family, including thiostrepton, the N-terminus is joined to a threonine residue by a highly modified tryptophan derivative, forming a second macrocycle. There is a wealth of diversity in the modification of the amino acids in these antibiotics involving dehydration, dehydrogenation, epoxidation, methylation, and more.

1.1 Mechanism of action

The thiopeptide antibiotics all inhibit ribosomal protein synthesis in Gram-positive bacteria. Gram-negative bacteria are not sensitive to these antibiotics *in vivo* because the antibiotics are not able to permeate the outer membrane. *E. coli* ribosomes, however, have been shown to be sensitive *in vitro*. All of the known mechanisms of action, excepting only that of GE2770, involve binding to the same site on the ribosome, but the antibiotics appear to act in very different ways.

The thiopeptides bind to the 50S ribosomal subunit, complexing with ribosomal protein L11 and 23S ribosomal RNA [12] [11]. Thiostrepton binds tightly ($K_{diss} < 10^{-9}$) to a ribosomal complex which has been prepared *in vitro* from *E. coli* cells.

It has been determined that thiostrepton binds the 23S ribosomal RNA complexed with ribosomal protein L11 with the same dissociation constant, but binds to the 23S rRNA alone with somewhat lower affinity, with $K_{diss} \simeq 5 \times 10^{-7}$. Ribosomal protein L11 showed no measurable affinity for thiostrepton alone. This suggests that thiostrepton binds primarily to the 23S rRNA, and the binding is enhanced by the presence of ribosomal protein L11. This interpretation is further supported by the known resistance mechanism in the producing organisms. A recent study has demonstrated that thiostrepton inhibits a conformational change in ribosomal protein L11 which is necessary for its activity in the ribosome [52].

An interesting associated characteristic of the thiopeptides' mode of action is their effect on the stringent response. When bacteria are subject to amino acid starvation the ribosome produces a pyrophosphate adduct of guanidine triphosphate (pppGpp) or guanidine diphosphate (ppGpp). In some Streptomycetes there appears to be a correlation between high levels of (p)ppGpp, and differentiation and production of secondary metabolites. To date there is no evidence, however, that high levels of (p)ppGpp are necessary for antibiotic production. Mutants of *Streptomyces coelicolor* that cannot produce (p)ppGpp are found to be lacking in ribosomal protein L11 [47] [46]. Consistent with this is the observation that cells subjected to sub-lethal concentrations of thiopeptin cease the production of (p)ppGpp. When thiopeptin binding to the ribosomes is inhibited by the action of the thiostrepton resistance gene (p)ppGpp synthesis is unaffected [45].

1.2 Resistance mechanisms

The known resistance genes found in the producing organisms block the binding of thiostrepton to the ribosome by O-methylating a single base, adenosine 1067, in the 23S ribosomal RNA [60]. This resistance gene, isolated from *S. azureus*, has been used to confer resistance to thiostrepton in a number of different species, and is used as a selectable marker in Stretomycete molecular biology [22]. The resistance genes which have been found in the thiostrepton producers *S. laurentii* and *S. actuosus* are highly homologous to the one from *S. azureus* and encode a very similar RNA O-methyl transferase.

Random mutagenesis of bacteria which have been shown to be sensitive to thiostrepton either *in vitro* or *in vivo* has yielded strains that are resistant to thiostrepton. It has been found that these resistant strains either have mutations in their 23S ribosomal RNA, specifically adenosine to guanine mutations at positions 1067 or 1095 [57][56], or that they exhibit a modification in ribosomal protein L11 [52] in which a proline to threonine or a proline to serine mutation has occured (P18T or P18S).

A frequent mode of antibiotic resistance found in non-producing strains is the transport of the antibiotic out of the cells. A DNA clone has been isolated from *Bacillus subtilis* which confers resistance to thiostrepton. A gene which is highly homologous to the *E. coli secE* protein transport gene has been found on this clone and postulated to be the gene responsible for resistance [24].

Another possible thiostrepton resistance mechanism exists in *Streptomyces lividans*. When *S. lividans* cells are subjected to low concentrations of thiostrepton it has been found that at least 8 proteins are expressed which are not found in cells

grown in the absence of thiostrepton. Two of these proteins have been characterized, TipAL and TipAS. Both of these proteins are expressed from the same gene, TipA, sharing the carboxy-terminal domain. TipAS has its own set of promoters inside the TipA gene, and TipAL is the entire gene product. TipAL has been demonstrated to bind thiostrepton by reacting one of two cysteines on the carboxy terminus of the protein with one of the dehydroalanine residues in thiostrepton, leading the investigators to classify the protein as an electrophile scavenger [9].

It has been reported that *Streptomyces incarnatus* contains an 8.7kb plasmid (pGIF3) which carries a thiostrepton resistance gene that is unrelated to the known resistance genes. When the plasmid is isolated, and used to transform thiostrepton sensitive *S. lividans* cells, it confers resistance to thiostrepton. When a Southern hybridization was performed using the known thiostrepton resistance gene as a probe no homology was found. The plasmid has been cloned into a vector and mapped, but no sequence information has been published to date [31].

1.3 Biosynthesis of peptide antibiotics

The mode of biosynthesis of the thiopeptide antibiotics has not been elucidated for any of the members of the family. In looking at the precursor peptide chains for this family of antibiotics it seems likely that they share the same mechanism of formation since they have a high degree of sequence identity (figure 1.6) and structural similarity. It is likely that if the biosynthetic genes for one of the antibiotics were found, they would prove to be valuable tools in finding the genes responsible for the production of the other compounds. There have been three biochemical pathways found which produce

peptide-based secondary metabolites; ribosomal biosynthesis, a direct condensation process, and the nonribosomal thiotemplate mechanism.

Ribosomally produced peptides

The best known mode of peptide biosynthesis is the ribosomal process. In this system amino acids are activated by formation of anhydrides with nucleic acids. The resulting activated amino acid is then transferred to the acceptor hydroxyl group on tRNA. This complex is transferred to the elongation domain of the ribosome, where the peptide chain is synthesized [82]. There are a number of secondary metabolites formed by this pathway, the most extensively studied being the lantibiotics family.

The lantibiotics are synthesized as large preproteins that are proteolytically cleaved prior to modification. The formation of lanthionine bridges occurs by the dehydration of serine or threonine residues, which act as electrophiles, reacting with the sulfhydryl group on cysteine residues yielding a thioether. The known gene clusters for these antibiotics contain regulatory genes, the gene encoding the preprotein, the gene encoding the protease, and post-translational modification genes [17] (figure 1.7).

There are other known gene clusters of ribosomally produced peptide antibiotics, including microcin B17 (figure 1.8), which is of interest to this project in that it also contains the thiazole and dihydrothiazole structures that are found in thiostrepton and the other thiopeptides.

Direct Condensation

Another known mechanism for the biosynthetic assembly of a peptide chain can be described as a direct condensation pathway. In this pathway, amino acids are condensed in a reaction exchanging ATP and pyrophosphate to activate the individual amino acids. The peptide antibiotic mycobacillin is synthesized by this mechanism [19], as is glutathione [82]. Glutathione is formed by first condensing glutamate with cysteine, in a reaction requiring ATP, followed by the condensation of this dipeptide with glycine. These two synthetases have been cloned, and *E. coli* mutants lacking the ability to produce this tripeptide have been found to grow naturally, only differing from the wild type in their sensitivity to sulfhydryl reagents, such as mercurials.

Thiotemplate Mechanism

The third known mechanism for the biosynthesis of peptide antibiotics is the now well studied thiotemplate mechanism. Many peptide antibiotic biosynthetic gene clusters have been characterized and shown to contain nonribosomal peptide synthetases of this type [28]. The multifunctional enzymes that synthesize the peptide chain have been examined and found to constitute a distinct family. The proteins show a high degree of similarity in core regions, and the catalytic function of many of these core regions has been elucidated either biochemically or by comparison to known activities. This family has also been found to be a member of a super family of ATP binding and hydrolyzing enzymes, including insect luciferases and acyl CoA synthetases.

Peptide synthetases work to build a peptide chain by a series of reactions in which an amino acid is activated to its acyl adenylate, condensed with the thiol group of

a phosphopantetheine cofactor which is covalently attached to the enzyme through a serine residue, and transferred to a similarly activated and attached amino acid. This process continues, with one elongation per activation module, until the peptide chain is constructed. Each elongation step requires an ATP to AMP plus pyrophosphate turnover as the primary energy source (figure 1.9).

These peptide synthetases, when compared with each other, form a strongly conserved family of enzymes which show the same general organization, and a high degree of amino acid sequence similarity in certain conserved regions. Each amino acid in the antibiotic is activated by a module on the peptide synthetase which is 1100 amino acids in length. The module contains all of the activities required to bind, activate, and condense the amino acid into the growing chain. Most of the conserved domains in the modules have had activities tentatively ascribed to them. The adenylation region (the A core) is about 500 amino acids in length and is capable of independently and specifically activating an amino acid to its acyl adenylate, releasing pyrophosphate. The thiolation region, about 100 amino acids long and immediately following the adenylation domain, is the attachment point of the phosphopantetheine cofactor, which is the carrier arm for the amino acid. This cofactor is covalently attached to the thiolation domain through a phosphoester bond to a serine hydroxyl group. The 600 amino acids preceding the adenylation domain have been tentatively assigned a condensation activity by sequence homology (figure 1.10).

The adenylation region of a peptide synthetase module has been shown to be able to independently recognize and activate a specific amino acid. This was demonstrated by amplifying by PCR regions containing all of the adenylation (A) domains from the

gene clusters which produce the peptide synthetases for gramicidin S biosynthesis and tyrocidine biosynthesis. When these gene fragments were expressed as if they were complete genes the resulting proteins were able to recognize the specific amino acids which the complete genes activated, and converted them to their acyl adenylates [33].

A peptide synthetase is constructed of one or more of these modules, and often contains domains responsible for the modification of the amino acids in the antibiotic. Eleven modules have been found on a single enzyme for the cyclosporin A synthetase in the fungus *Tolypocladium niveum*. This enzyme catalyzes a total of 40 reactions, and at 1,689,243 Da is the largest known enzyme [33].

Many of these gene clusters have now been cloned and characterized, and it has been found that the majority of the bacterial peptide secondary metabolites are produced by this mechanism. A particularly well studied peptide antibiotic that is produced by this mechanism is Gramicidin S. Gramicidin S is a cyclic peptide dimer that is produced by *Bacillus brevis* ATCC 9999. The gene cluster for this antibiotic has been cloned and sequenced, and has been shown to contain a 4'-phosphopantetheine transferase, a thioesterase, a peptide synthetase which activates and epimerizes phenylalanine, and a larger peptide synthetase which activates and condenses into the growing chain the remaining four amino acids, and contains a thioesterase domain to cleave the product from the enzyme (figure 1.11).

1.4 Goals

The rapid appearance of pathogenic bacteria which have developed resistance to known antibiotics has made the development of new classes of antibiotics with novel

modes of action vital to the medical world [43]. A great effort is being undertaken in both the academic world and in industry to screen compounds from natural and synthetic sources for biological activity, and to develop methods to generate new structures with possible pharmaceutical utility. The goal of this project is to find the genes responsible for producing thiostrepton, and use them to develop strains of bacteria that are capable of producing a wide variety of different, potentially active, compounds.

The approach of combining genes from different antibiotic producing bacteria to produce novel structures has been at the forefront of research in the polyketide natural products field for the past few years. Several new structures have been produced, and are being investigated for biological activity [36]. The limitation to this method is that the polyketides have a limited number of building blocks. The structures are mostly based on acetate and propionate. If the same sort of approach could be used with the peptide antibiotics one could achieve a much broader diversity of structures. Starting with twenty-six building blocks, with an almost infinite variety of ways in which they can be modified, the libraries that could be generated have the potential to rival those produced chemically in diversity, and be well ahead in complexity.

Thiostrepton, and the other thiopeptides, are unique among the peptide antibiotics in their extraordinarily high degree of modification. If the encoding genes could be identified and cloned they could be used in conjunction with the genes responsible for assembling the amino acid chain to make new and potentially useful peptide-based structures.

The goal of this project is to find the genes coding for the enzymes which are

involved in the biosynthesis of thiostrepton in *S. laurentii*. Three approaches have been pursued. In the first attempt to find the biosynthetic gene cluster the DNA surrounding the resistance gene was analyzed, as it has been found that in all known cases an antibiotic's biosynthetic genes are clustered around a resistance gene. The second approach taken used the method developed by Marahiel and co-workers [73] to find peptide synthetases, a possible biosynthetic enzyme, using PCR primers based on known conserved amino acid sequences. The final approach involved purifying a protein known to be on the pathway, determining part of its amino acid sequence, and using that information to develop Southern probes and PCR primers to isolate DNA containing the biosynthetic genes from a cosmid library.

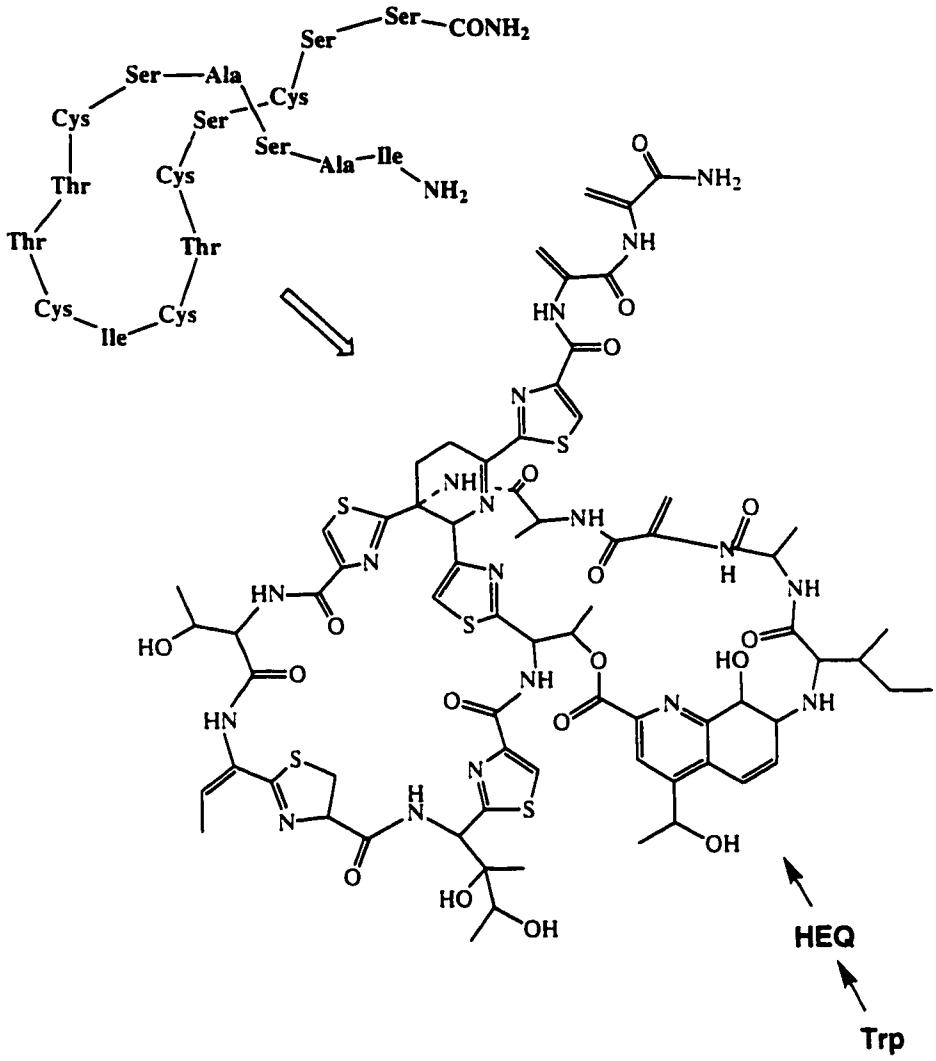


Figure 1.1: The thiopeptide antibiotic thiostrepton. The precursor amino acids are shown top left, and the origin of the hydroxyethyl quinaldic acid moiety (HEQ) is shown bottom right.

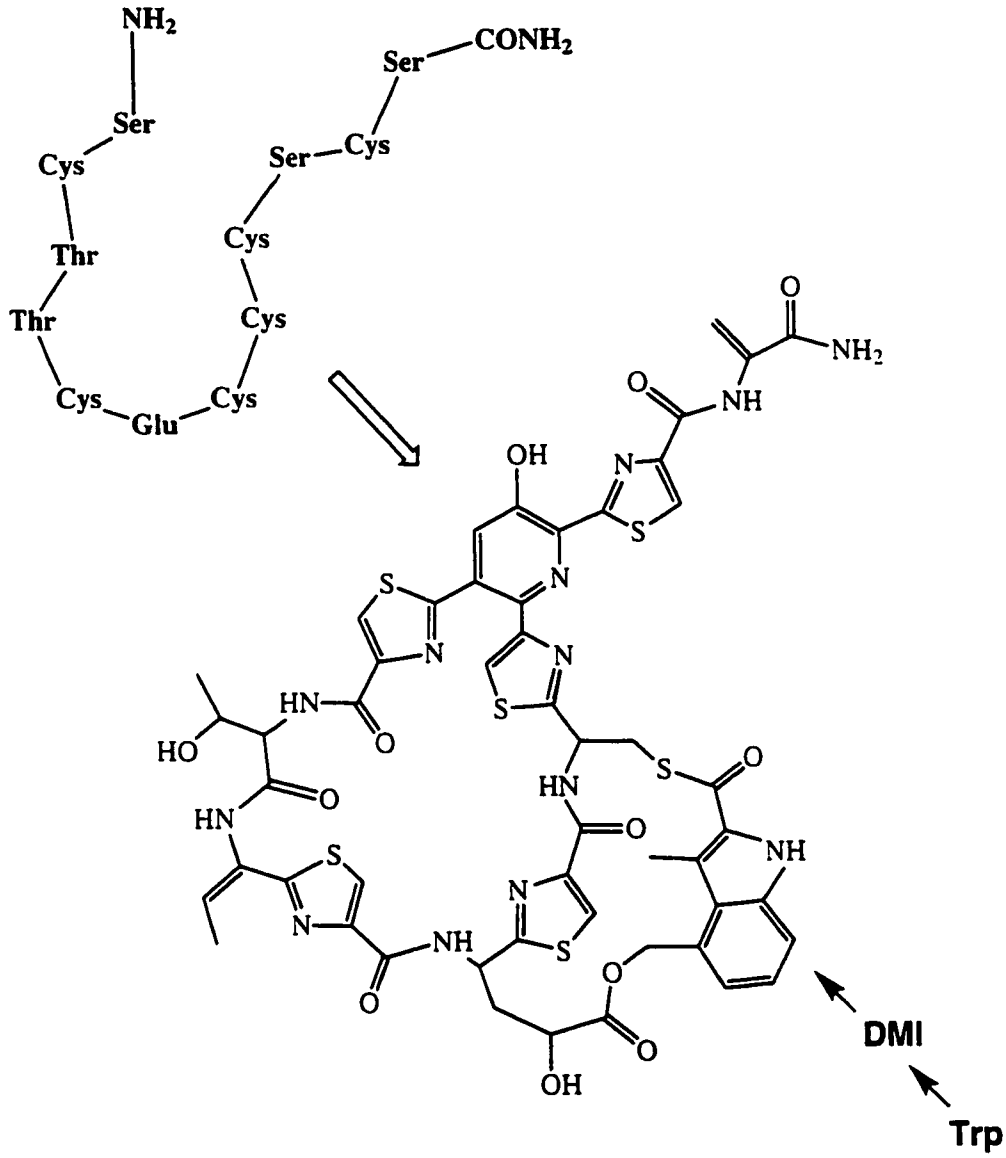
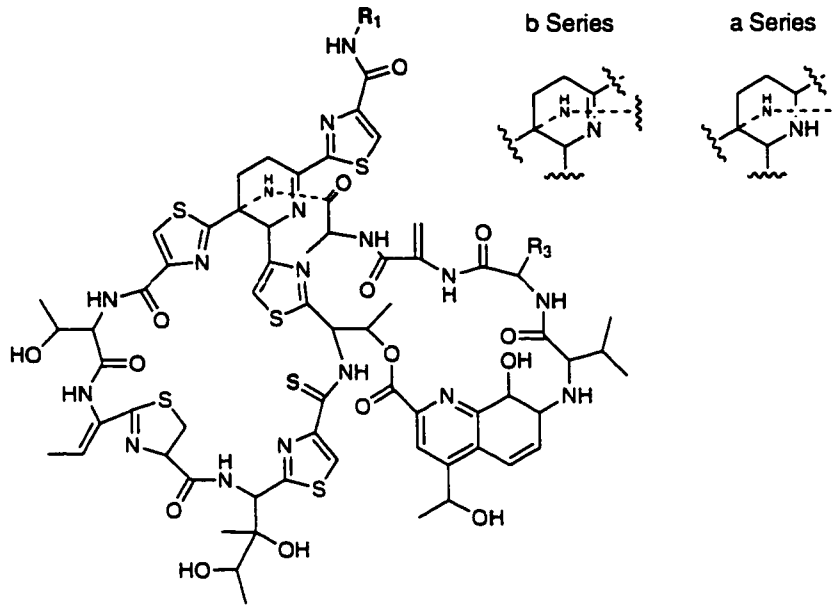
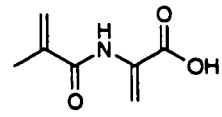
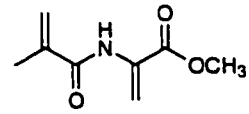
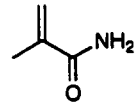


Figure 1.2: The thiopeptide antibiotic nosiheptide, a closely related structure to thiostrepton. Precursor amino acids are shown above, and the origin of the dimethyl indolic acid moiety (DMI) is shown bottom right.



Thiopeptin $B_{a,b}$ R_1 =  R_3 = ala

$A_{1a,b}$ R_1 =  R_3 = ala

$A_{4a,b}$ R_1 =  R_3 = ala

$A_{3a,b}$ R_1 = H R_3 = ala

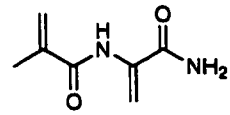
Siomycin R_1 =  R_3 = deala

Figure 1.3: The thiopeptins and siomycin. The thiopeptin b series contains a tetrahydropyridine ring, in the a series the same ring has been hydrogenated to the hexahydropyridine ring.

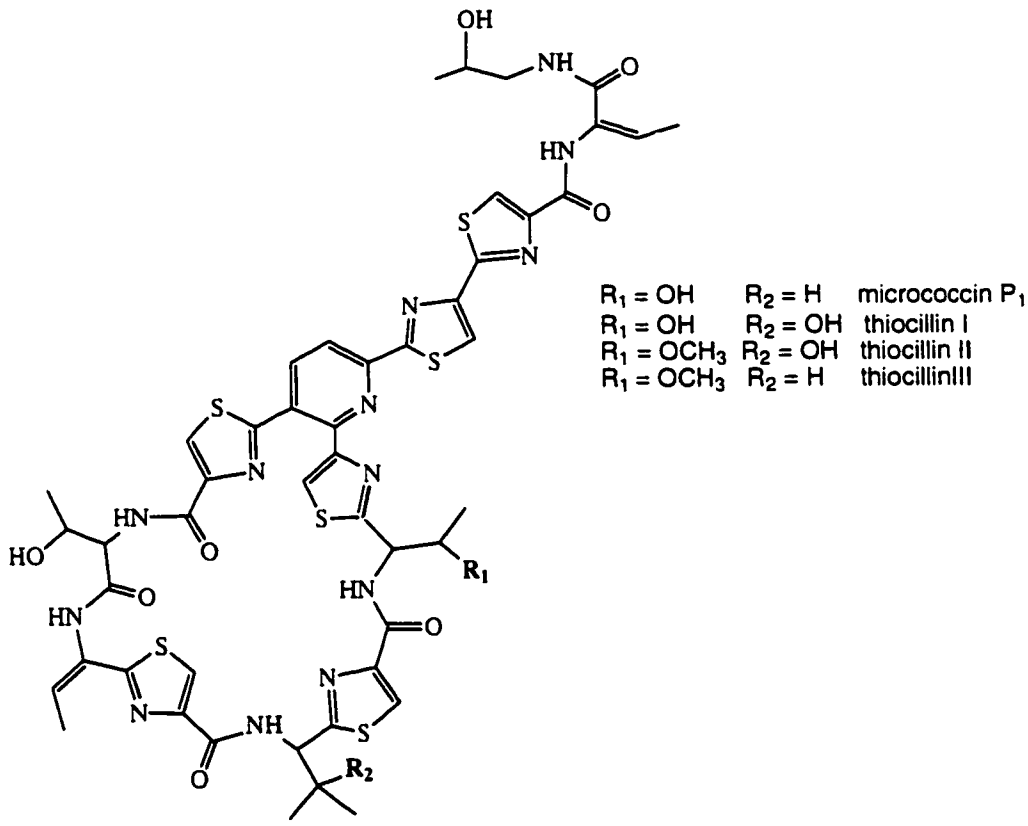


Figure 1.4: Micrococcin and the thiocillins

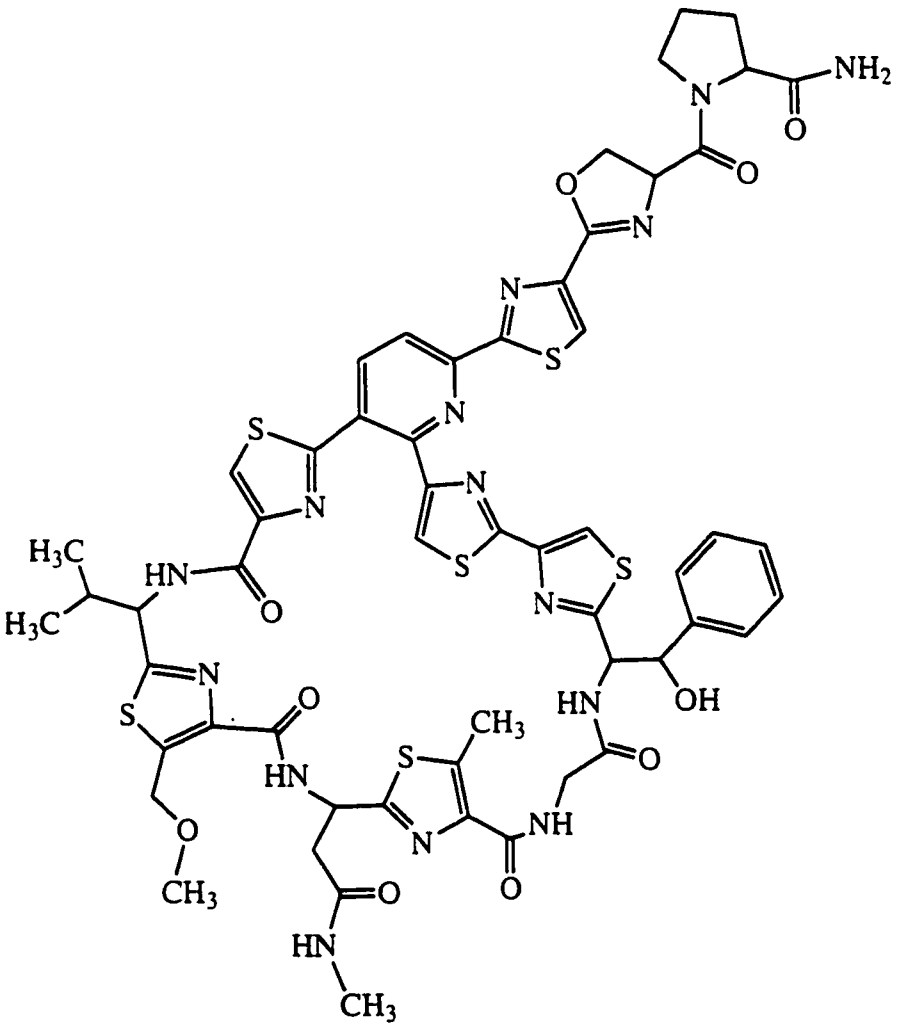


Figure 1.5: GE2770

micrococcin / thiocillins

ser cys thr thr cys val cys thr cys ser cys cys thr

nosiheptide

ser cys thr thr cys glu cys cys cys ser cys ser

thiostrepton

ile ala ser ala ser cys thr thr cys ile cys thr cys ser cys ser ser

siomycin

val ser ser ala ser cys thr thr cys ile cys thr cys ser cys ser ser

thiopeptins

val ala ser ala ser cys thr thr cys ile cys thr cys ser cys ser ser

val ala ser ala ser cys thr thr cys ile cys thr cys ser cys ser

val ala ser ala ser cys thr thr cys ile cys thr cys ser cys

Figure 1.6: Linear peptide precursors to the thiopeptide antibiotics showing the high degree of sequence identity of the precursor amino acids.

Figure 1.7: Subtilin and nisin, ribosomally produced peptide antibiotics. Precursor peptides are given, with modified amino acids underlined and in bold face, and lanthionine bridge locations shown. Amino acids are modified by the dehydration of serine to 2,3-didehydroalanine, threonine to 2,3-didehydrobutyrine, and the addition of the cysteine sulfhydryl to these unsaturated systems, yielding the lanthionine bridges. The gene clusters are represented below the precursor sequence.

Nisin gene	Subtilin gene	Activity
A	S	Preprotein
nisB, nisC	spaB, spaC	Peptide modification
nisE, nisF, nisG	spaF, spaG	Self protection
nisI	spaI	Lipoprotein involved in immunity
nisK	spaK	Regulation
nisP	spaP	Proteolysis of preprotein
nisR	spaR	Regulation
nisT	spaT	Transport

Subtilin

WKSETLCTPGCVTGALQTCFLQTLTCNCKISK

Nisin

ITSITLCTPGCKTGALMGCNMKTATCHCSIHVSK

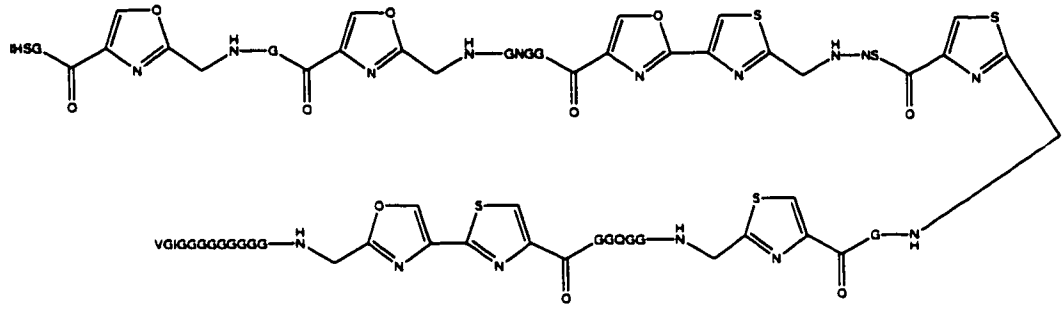


Figure 1.8: Microcin B17, a ribosomally biosynthesized peptide antibiotic containing thiazole and oxazole ring systems.

Acyladenylation and attachment to enzyme

$$E^1 - S^1H + A^1 + MATP^{-2} \rightleftharpoons (A^1AMP)E^1 - S^1H + MPP_i^{-2} \rightarrow E^1 - S^1A^1 + AMP^{-2}$$

$$E^2 - S^2H + A^2 + MATP^{-2} \rightleftharpoons (A^2AMP)E^2 - S^2H + MPP_i^{-2} \rightarrow E^2 - S^2A^2 + AMP^{-2}$$

Condensation

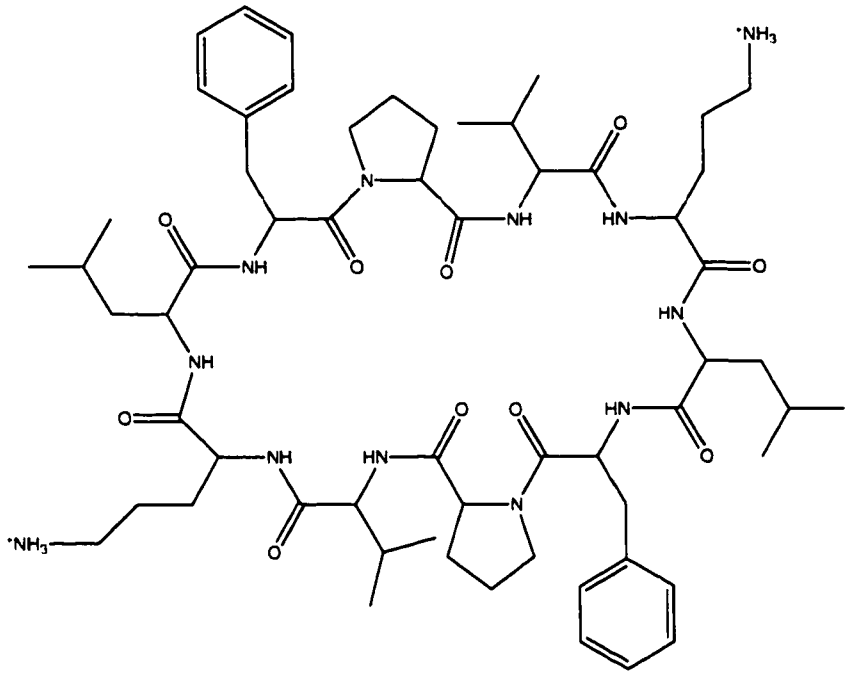
$$E^1 - S^1A^1 + E^2 - S^2A^2 \rightarrow E^2 - S^2A^2A^1 + E^1 - S^1H$$

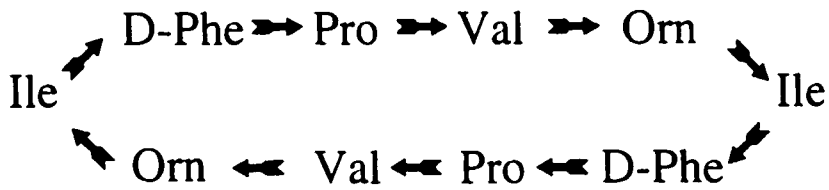
Figure 1.9: The general model for peptide elongation using the thiotemplate system. $E^1 - S^1H$ and $EE^2 - S^2H$ are two consecutive activation domains with the phosphopantetheine arm attached. M is an appropriate divalent metal, and A^1 and A^2 are two consecutive amino acids in the peptide.

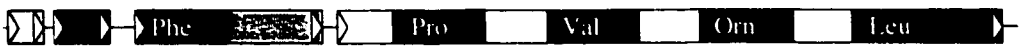
Domain	Core	Consensus Sequence
adenylation	A1	L(TS)YxEL
	A2	LKAGxAYL(VL)P(LI)D
	A3	LAYxxYTSG(ST)TGxPKG
	A4	FDxS
	A5	NxYGPTE
	A6	GELxIxGxG(VL)ARGYL
	A7	Y(RK)TGDL
	A8	GRxDxQVKIRGxRIELGEIE
	A9	LPxYM(IV)P
	A10	NGK(VL)DR
condensation	C1	SxAQxR(LM)(WY)xL
	C2	RHExLRTxF
	C3	MGGxISDG(WV)S
	C4	YxD(FY)AVW
	C5	(IV)GxFVNT(QL)(CA)xR
	C6	(HN)QD(YV)PFE
	C7	RDxSRNPL
thiolation	T	DxFFxxLGG(HD)S(LI)

Figure 1.10: The genetic organization of a peptide synthetase module, showing the conserved domain regions and their consensus amino acid sequences. [33]

Figure 1.11: Gramicidin S, a cyclic peptide antibiotic produced by *Bacillus subtilis*. The amino acid connectivity is shown below the chemical structure, and a representation of the biosynthetic gene cluster at the bottom showing the modular construction of peptide antibiotics.







adenylation domain (aa: amino acid substrate)

thioesterase (domain or distinct protein)

4'-PP-transferase

thiolation domain

epimerisation domain

condensation domain

Chapter 2

PREVIOUS STUDIES ON THE BIOSYNTHESIS OF THIOSTREPTON

2.1 Structural determination

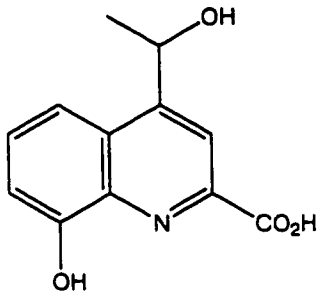
The structure of thiostrepton was elucidated through a combination of studies on various degradation products and X-ray crystallography. Thiostrepton was subjected to acid hydrolysis, and the resulting products were characterized by their proton NMR spectra, IR spectra, and by comparison with known compounds [8][16][6]. These studies gave the structures of many of the building blocks of thiostrepton (figure 2.1), but did not yield sufficient information to connect the building blocks and yield the final structure [7].

Crystals of thiostrepton were grown and the X-ray diffraction pattern observed. This experiment confirmed the degradation studies and, along with previous results, yielded the complete structure of thiostrepton and its conformation in crystal form [2].

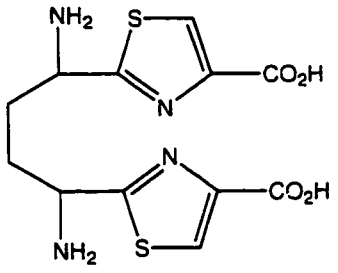
The majority of the protons in thiostrepton were assigned in an NMR study using chemical shift values, coupling patterns, and selective decoupling experiments. The X-ray crystal structure of the antibiotic was used extensively to do this. Some initial ^{13}C NMR assignments were done by Tori *et al.* [70]. Some corrections of the

(-)-Alanine
(-)-Threonine
(-)-Isoleucine
(+)-Cysteine

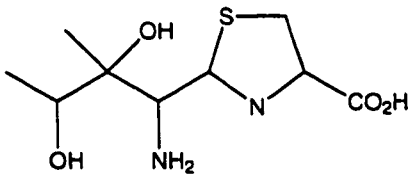
Pyruvic Acid
Ammonia



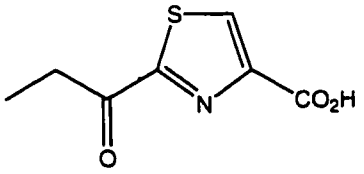
4-(α-Hydroxyethyl)-8-hydroxyquinaldic acid



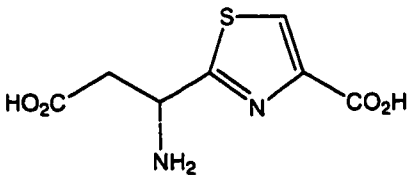
Thiostreptoic acid



Thiostreptine



2-Propionylthiazole-4-carboxylic acid



2-(1-Amino-2-carboxyethyl)thiazole-4-carboxylic acid

Figure 2.1: Products obtained upon hydrolysis of thiostrepton. Two equivalents of alanine and pyruvic acid were found, four to five equivalents of ammonia, and less than one equivalent of (+)-cysteine.

earlier study and assignment of the remaining signals were completed by Hensens and Albers-Schönberg. They used ^{1}H–^{13}C one and two bond coupling, the disappearance of coupling to amide protons in protic solvents, and comparison with model compounds [20]. The final assigments, with further refinements of the Hensens and Albers-Schönberg study, were done by Mocek *et al.* [39] using two dimensional NMR on unlabeled and biosynthetically multiply labeled thiostrepton samples.

2.2 The biosynthetic pathway

The precursors of thiostrepton have been elucidated by feeding of isotopically labeled compounds [40], showing that the antibiotic is derived entirely from natural amino acids. There are 18 total precursor amino acids, including tryptophan, which gives rise to the hydroxyethylquinaldic acid moiety (HEQ). These amino acids are then highly modified, presumably by a wide variety of enzymes, to yield the antibiotic.

2.2.1 Amino acid modifications

Many of the modified amino acids in thiostrepton are of obvious origin. The dehydroalanine and butyrine residues come from the dehydration of serine and threonine, respectively. The thiostreptine residue, which is made from isoleucine, presumably arises by dehydrogenation, followed by epoxidation and ring opening. The tetrahydropyridine ring is believed to be formed by the cycloaddition of two dehydroalanine residues, which are known to be derived from serine (figure 2.2).

The formation of the dihydrothiazole and thiazole rings has been postulated to result from the nucleophilic attack of the cysteine sulfhydryl group on the carbonyl

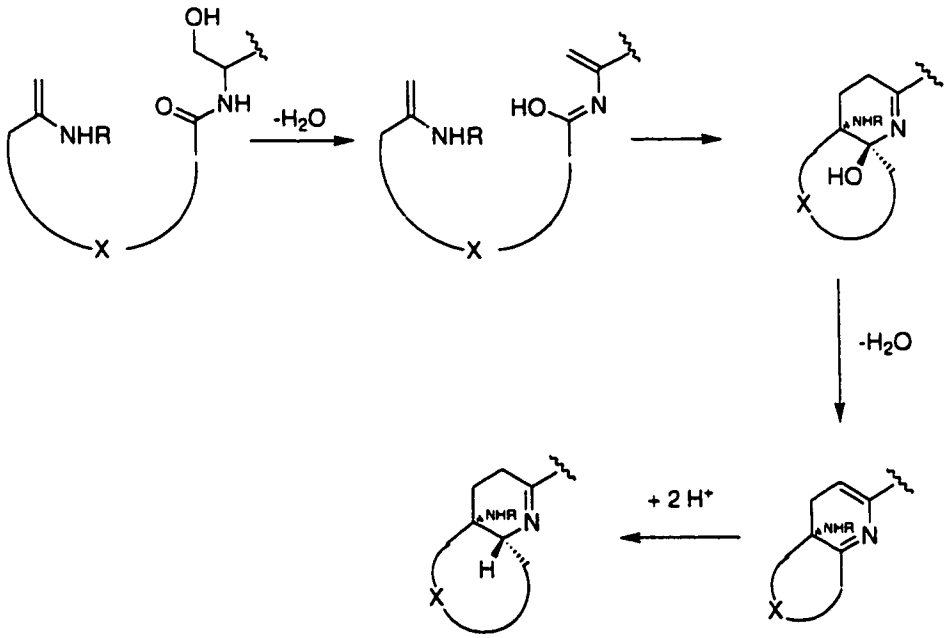


Figure 2.2: Formation of the tetrahydropyridine moiety of thiostrepton. The dehydration of two serine residues to dehydroalanine is postulated to initiate the process.

Table 2.1: Amino acid precursors of thiostrepton

Residue	*Occurrences*	*Precursor*
Dehydroalanine	3	Serine
Thiazole	4	Cysteine
Dihydrothiazole	1	Cysteine
Butyrine	1	Threonine
Tetrahydropyridine	1	2 Serines
Hydroxyethylquinaldic Acid	1	Tryptophan
Threonine	2	Threonine
Isoleucine	1	Isoleucine
Alanine	2	Alanine
Thiostreptine	1	Isoleucine

carbon of the next amino acid, followed by dehydration (figure 2.3). Thiazole is then formed by oxidizing the dihydrothiazole ring. This mechanism has been studied in the biosynthesis of microcin B17 (figure 1.8) in detail, and the gene responsible for the reaction has been cloned. The peptide antibiotic bacitracin, which has been shown to be produced using the nonribosomal thiotemplate mechanism, contains thiazoline rings that have been postulated to be formed by an activity integrated into the condensation domain of the module activating the amino acid preceding the cysteine residue [58].

Less is known about the biosynthesis of the hydroxyethylquinaldic acid moiety

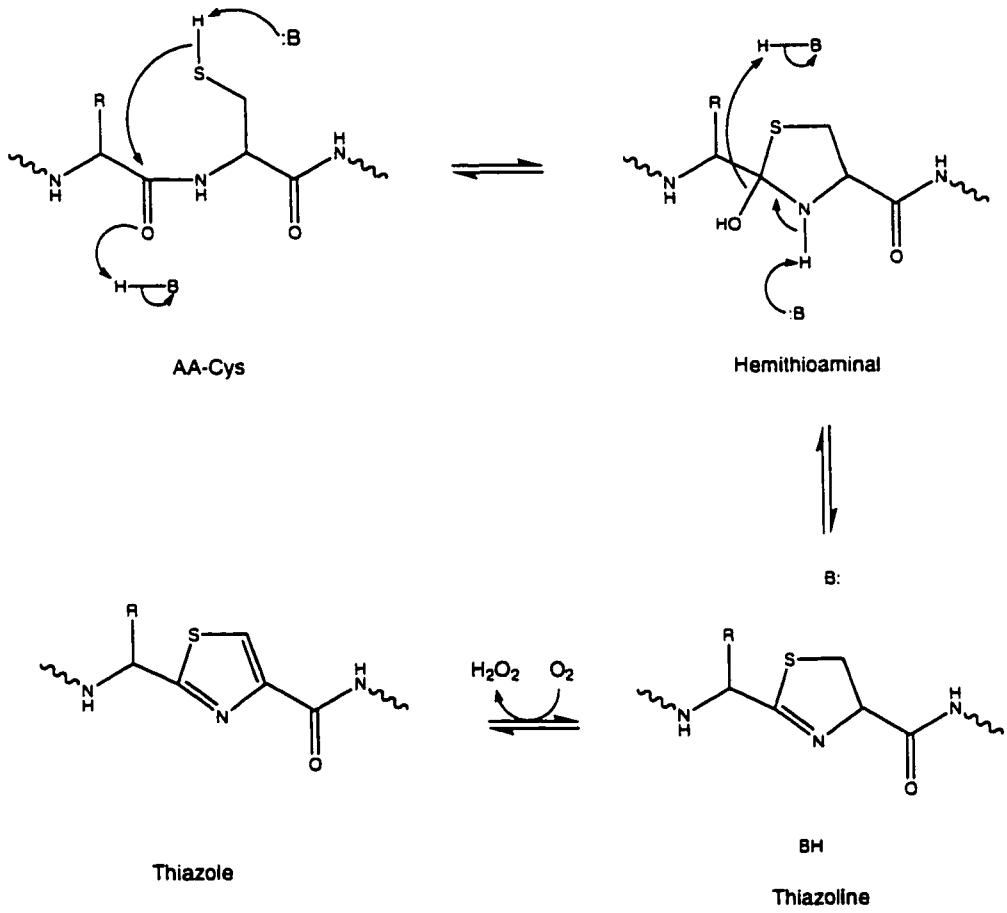


Figure 2.3: Formation of the thiazole rings

(HEQ). It is known to be derived from the amino acid tryptophan and the first step, a methylation of the indole two position, has been studied [81]. This reaction requires S-adenosylmethionine as the methyl donor and has been shown to proceed with net retention of configuration of the methyl group.

The enzyme catalyzing the methylation of the indole two position was partially characterized in a cell-free system [18]. The reaction was shown to require S-adenosylmethionine and L-tryptophan, and produce 2-methyltryptophan. The apparent K_m value for S-adenosylmethionine was measured to be 120 μM. The purification of this enzyme proved to be prohibitively difficult. The enzyme has a very short half life (2 hours at 30°C), and does not survive ammonium sulfate precipitation, dialysis, membrane ultrafiltration, or column chromatography under a variety of conditions [18]. The apparent K_m of tryptophan was not determined since the endogenous tryptophan cannot be removed from the cell-free extract without denaturing the enzyme. While the loss of activity after desalting could be explained by the separation of the enzyme from a necessary, but loosely bound, cofactor, recombination of the two fractions failed to restore enzymatic activity.

When (S)-[1',2'-$^{13}C_2$,*indole*-^{15}N]tryptophan was fed to producing cells, the thiostrepton produced showed one bond ^{13}C-^{15}N NMR coupling to C2, and two bond ^{13}C-^{15}N NMR coupling to C1. This demonstrates that it is the N1 to C2 bond that is broken in the rearrangement from 2-methyltryptophan to the quinaldic acid derivative, and that this rearrangement is intramolecular [40]. Other than this experiment, the mechanism for the further transformation of 2-methyltryptophan into HEQ has not been elucidated, though two mechanisms have been proposed in the literature.

In the first mechanism [53] (figure 2.4), the amino group of 2-methyltryptophan is transaminated forming an alpha keto acid. The five membered ring is activated by the attack of an electrophile, forming an imine bond which is readily hydrolyzed to open the ring. The six membered ring is then formed by the nucleophilic attack of the amine nitrogen on the alpha keto carbon, followed by dehydration. Re-aromatization of the ring is accomplished by having the electrophile leave as the reduced X^-. There is precedence for this mechanism in the work of van Tamelen and Haarsted [75] who showed that 2-methyltryptophan is oxidized to 4-acetylquinoline-2-carboxylic acid in the presence of hypochlorite in approximately 20% yield. The methyl ketone which is formed must be reduced to yield free HEQ.

The second proposed mechanism [64] (figure 2.5) differs from the first in that instead of an electrophile attacking the ring, an enzyme-bound protonated base donates a proton to the ring, which is removed to form an enol near the end of the reaction. This is postulated to isomerize directly to form HEQ.

Once HEQ is formed, an enzyme adenylates the acid to activate the carboxyl group towards nucleophilic attack from the threonine hydroxyl group. The enzyme that catalyzes this reaction has been partially purified and characterized [53]. Apparent K_m values for *rac*-HEQ, *R*-HEQ, and *S*-HEQ were determined to be 3 μM, 31 μM, and 1.9 μM, respectively, showing that the enzyme has a high preference for the *S* isomer of HEQ. The much lower activity seen in the incubation with *R*-HEQ was attributed to a small amount of the *S* isomer in the sample. A somewhat lower affinity was measured for ATP, at a K_m of 300 μM. A series of HEQ and tryptophan homologues were synthesized to test the substrate specificity of the HEQ acyl adeny-

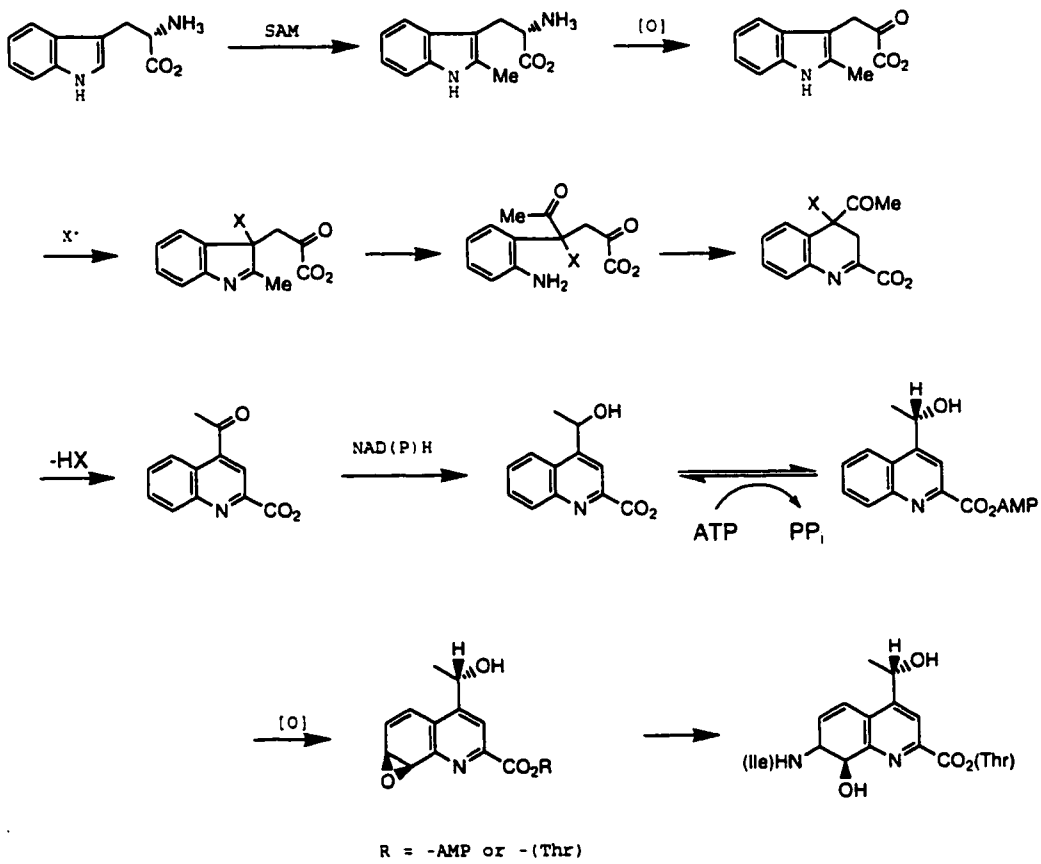


Figure 2.4: First proposed mechanism for the formation of HEQ from tryptophan.

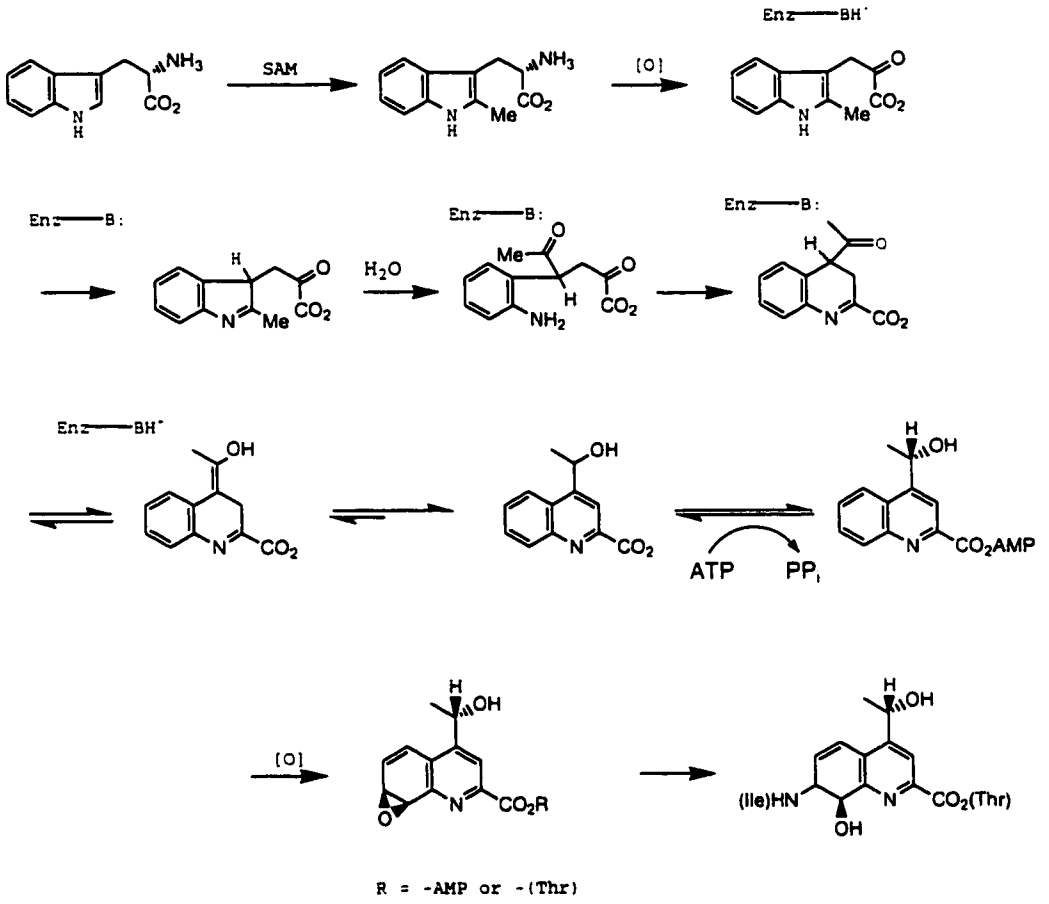


Figure 2.5: Second proposed mechanism for the formation of HEQ

lase, and of these structures 4-(hydroxymethyl)quinoline-2-carboxylate was the only other substrate to be sucessfully adenylated by the enzyme, with an apparent K_m of 80 μM, and a V_{max} about 60% lower than that measured for *rac*-HEQ. SDS-PAGE analysis of protein that had been photolabeled using [α-^{32}P]-8-azido-ATP, and native gel filtration chromatography both indicated a molecular weight between 45 and 50 kDa for the enzyme.

The second attachment of HEQ to the ring is presumably achieved by the epoxidation of the carbocycle, followed by ring opening by the amino group of the N-terminal isoleucine.

2.3 Attempts to find the biosynthetic gene cluster

Several standard approaches are used to locate a biosynthetic gene cluster in a bacterial genome. Most of these take advantage of the observation that for all known bacterial secondary metabolites the genes coding for the enzymes which are responsible for the production of the metabolite are clustered on the genome. Finding one pathway gene conveniently allows one to identify the remaining biosynthetic genes by examining the surrounding region.

2.3.1 Random mutagenesis

The earliest approach to finding gene clusters involved generating random mutants by treating cells with chemicals that are known to cause DNA mutations, or by irradiation of the cells with ultraviolet light. Surviving colonies were then screened for mutants lacking the production of the metabolite that is being investigated. This

was an excellent approach for finding the gene cluster of metabolites that were easily assayed, but if the assay was not rapid this method could become prohibitively time consuming. The early uses of this technique took advantage of colored metabolites, where the lack of production is visibly evident and the screening of thousands of cultures is feasible. One can find the gene that has been mutated by transforming the mutant with a cosmid library and screening for reconstituted production. Unfortunately production of thiostrepton must be monitored by TLC or HPLC as the antibiotic does not change the appearance of cultured cells. Chemical mutants of *S. azureus* were generated first using nitrous acid as the mutagen. While non-producing mutants were isolated, they were found to revert spontaneously to the wild type phenotype. When another set of mutants was generated using MNNG as the mutagen a total of 10 non-producing mutants of *S. azureus* were isolated. These mutants were later found to have lost large portions of their genomic DNA, making them not useful for finding the biosynthetic gene cluster.

2.3.2 Oligonucleotide probing

An oligonucleotide probe was developed [79] based on the sequence predicted for the peptide precursor of thiostrepton, assuming that it is ribosomally produced, and the GC bias of Streptomycete DNA. The oligonucleotide probe was designed from a sequence of the putative peptide precursor that had a low genetic degeneracy. Thiostrepton's amino acid sequence contains a number of serine residues, which have a 6 fold coding degeneracy, significantly decreasing the chance of having a positive result. Cysteine has only a two fold coding degeneracy, so designing the probe based on a

sequence containing the largest number of cysteine residues increases the chances. Isoleucine has a three fold degeneracy, lower than the average. Most other amino acids found in thiostrepton have a four fold degeneracy in the genetic code. The peptide chosen as a basis for the probe, and the coding sequence used as the probe were:

Cys	Thr	Thr	Cys	Ile	Cys	Thr	Cys
TGC	ACC	ACC	TGC	ATC	TGC	ACC	TGC
	G	G				G	

This probe was used in a Southern hybridization experiment, and although a wide range of different conditions were employed no clear hybridization was observed. This experiment, however, suffered from lack of a positive control.

2.3.3 Chloramphenicol ribosomal inhibition assay

Although not a genetic approach, the following experiment was done to confirm the above Southern hybridization result. Chloramphenicol is an antibiotic which inhibits the ribosomal synthesis of proteins in bacteria [82]. Cultures of *S. laurentii*, *S. azureus*, and *S. actuosus* were grown and chloramphenicol was introduced during the production phase. After addition of chloramphenicol the cells were exposed to ^{35}S labeled cysteine. Radiolabel incorporated into both total protein and the antibiotic was measured after culturing for sufficient time to produce more of the antibiotic. It was found that in cultures of *S. actuosus* and *S. azureus* the difference in incorporation of the radiolabel into total protein was fairly small between the cultures subject to chloramphenicol and the control cultures which had no chloramphenicol. *S. laurentii*

showed very little incorporation of the radiolabel when exposed to chloramphenicol, which indicates a dramatic reduction in ribosomal protein synthesis.

When the produced antibiotics were analyzed it was seen that thiostrepton production in *S. azureus* and nosiheptide production in *S. actuosus* was not significantly affected by the presence of chloramphenicol, which suggests that the antibiotics are not ribosomally produced. The production of thiostrepton by *S. laurentii*, however, was almost completely stopped by the addition of chloramphenicol, indicating the possibility that the antibiotic may be ribosomally produced. Alternatively the antibiotic could be produced nonribosomally, but an enzyme in the pathway degrades rapidly and cannot be replaced in the presence of chloramphenicol [79].

At this stage there are no clear indications as to whether or not the antibiotic is produced ribosomally, but the evidence suggests that nonribosomal production is more likely.

2.4 The resistance gene

In prokaryotic systems it has been found in all cases investigated to date that the gene cluster encoding the biosynthesis of an antibiotic also contains a gene or genes conferring resistance to that antibiotic. Sometimes there are multiple resistance genes, or even multiple resistance mechanisms, in a producing organism. Not all of the resistance genes are necessarily clustered with the biosynthetic genes, but at least one has always been found in the cluster. It has been hypothesized that this trait is preserved since if a major mutation occurs in the bacteria which removes the resistance gene, that same mutation would very likely abolish production of the antibiotic, allowing

the mutant to survive. Many research projects have taken advantage of this property to find antibiotic biosynthesis gene clusters in prokaryotic organisms.

The thiostrepton resistance gene from *S. laurentii* was cloned by screening a plasmid library of *S.laurentii* genomic DNA with the nosiheptide resistance gene from *S. actuosus* and the thiostrepton resistance gene from *S. azureus*. A 1.5 kb DNA fragment was isolated and thought to contain the resistance gene. This fragment was cloned and used to transform *S. lividans* TK24, a thiostrepton sensitive streptomycete. Cultures of the transformed *S. lividans* were found to be resistant to thiostrepton *in vivo*. The 1.5 kb fragment was sequenced and found to contain an 810 nt open reading frame which codes for a protein with high homology to both the *S. azureus* thiostrepton resistance gene (68% nucleotide identity), and the *S. actuosus* nosiheptide resistance gene (70% nucleotide identity) [64].

Cosmid libraries were generated for *S. laurentii* in both the *E. coli* cosmid vector SuperCos (Stratagene, Inc.) and the *E. coli*-streptomycete shuttle cosmid vector pOJ446 [5]. Cosmid clones which contained the fragment of DNA conferring resistance to thiostrepton were identified by Southern hybridization. Overlapping cosmids were used to develop a restriction site map of a 70 kb region which contains the resistance gene. No further resistance genes have been isolated from this organism.

Since no other thiostrepton resistance mechanisms have been detected in *S. laurentii* it was reasonable to assume that the thiostrepton biosynthetic genes are located within this mapped region. With this in mind, a DNA library of randomly generated fragments was prepared, and a number of these were sequenced. When these random sequence fragments were mapped, and compared with sequences deposited in the

databases, it was found that the thiostrepton resistance gene is located in a cluster of ribosomal protein operons. This cluster includes regions homologous to rpoC, EF-Tu, EF-G, L22, L30/S5, and secY in an arrangement consistent with known clusters of ribosomal protein operons in *E. coli*, *Mycobacterium leprae*, and *Micrococcus luteus* [64] (figure 2.6). No sequence homolgies to known nonribosomal peptide synthetases were detected.

To prove conclusively that this region does not contain any genes responsible for the biosynthesis of thiostrepton it is necessary to disrupt a gene in this region and show that the resulting mutant does not produce thiostrepton. pKC1132 (figure 2.7) is a plasmid that is used to generate mutants in Streptomycetes by single crossover homologous integration. It contains a replication origin for *E. coli*, allowing easy construction and production of the mutational vector, but does not contain the replication origin for Streptomycetes. If the plasmid does not integrate into the genome of the transformed Streptomycete there is no origin present to allow the plasmid to replicate during cell division. This does not allow the bacteria to pass the traits, including the selective marker that confers resistance to an antibiotic, that are coded on the DNA of the plasmid on to their offspring. Integration occurs by having the region in the vector that originated from the genomic DNA bind to its complementary strand on the genomic DNA. When the genome is being replicated there is a small but finite chance that the DNA polymerase will move over to the plasmid, replicate it, then move back onto the genomic DNA and continue the replication. This integrates the DNA of the vector into the genome of the next, and further, generations. It also has the useful side effect of interrupting a gene if a fragment of that gene was used

as the homologous insert in the vector (figure 2.8). It is possible to have an insertion event which does not inactivate the gene. If the DNA inserted in pKC1132 contains an end of the gene of interest, the insertion will yield two intact copies of the gene. This will only inactivate the cluster if regulation or transcription of genes downstream of the gene of interest are interrupted. If the genes on pKC1132 are in the same orientation as the insert, the insertion will probably not interrupt transcription of the remaining biosynthetic genes. If it is in the opposite orientation transcription will likely be disrupted.

Direct PEG mediated transformation of *S. laurentii* protoplasts with plasmids using the Hopwood protocol [23] has been unsuccessful, as has electroporation. In order to introduce DNA successfully into the cells it was necessary to do a conjugation experiment. Conjugation is the mechanism by which bacteria exchange genetic information, and involves the direct transfer of plasmid DNA from one cell to another by mating. This process was first observed to occur between Gram-negative bacteria, but was thought to be unlikely between Gram-negative bacteria and Gram-positive bacteria. The latter was proven not to be true [72], and the process was found also to work with Streptomycetes [34]. The genes responsible for this mechanism have since been found to be the oriT gene, a 760 bp fragment isolated from the natural conjugational vector RK2 which must be present in the transferred DNA, and the *tra* gene which must be carried by the donor bacterium.

DNA from the region surrounding the resistance gene was digested and ligated into pKC1132, and the resulting plasmids were transferred from *E. coli* to *S. laurentii* by conjugation. Apramycin resistant mutants generated in this way were grown

under conditions favoring thiostrepton production, and assayed for the presence of the antibiotic (section 3.2, page 47). The mutants were found to produce thiostrepton at approximately wild type levels, with the exception of two mutants that were found to be non-producers. These were shown to be insertions in regions containing the genes encoding ribosomal proteins S10 and spc, leading to the hypothesis that the lack of production in these mutants is due to pleiotropic effects, rather than the disruption of a thiostrepton biosynthetic gene [65]. This result very strongly indicates that the genes responsible for producing thiostrepton are not located in the vicinity of the cloned resistance gene. To date, experiments to delete the known thiostrepton resistance gene in order to find a possible second resistance gene have proven unsuccessful, and no viable mutants have been isolated. This has been postulated to be due to the bacteria having become dependent upon the rRNA methylation, making the disruption of this thiostrepton resistance gene a fatal mutation.

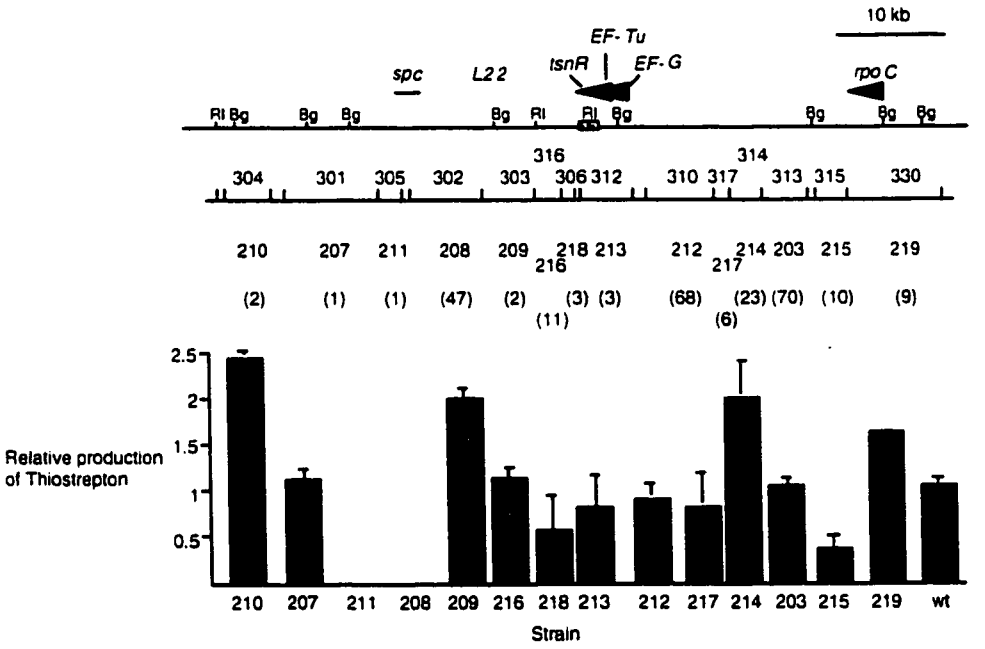


Figure 2.6: Map of the DNA surrounding the known thiostrepton resistance gene in *Streptomyces laurentii*. The bar graph below the map represents the production profile of the mutants generated by single crossover insertional mutation. The non-producing mutants were disrupted in the region containing an SPC operon.

Figure 2.7: pKC1132 is a gene disruption plasmid used in streptomycete research. OriT is the origin of conjugal transfer. ApR is the Apramycin resistance gene. The multi cloning site (MCS) is located within the LacZα gene which provides blue/white selection. pUC is the *Escheria coli* replication origin. A piece of genomic DNA is inserted into the multi cloning site to generate the gene disruption plasmid.

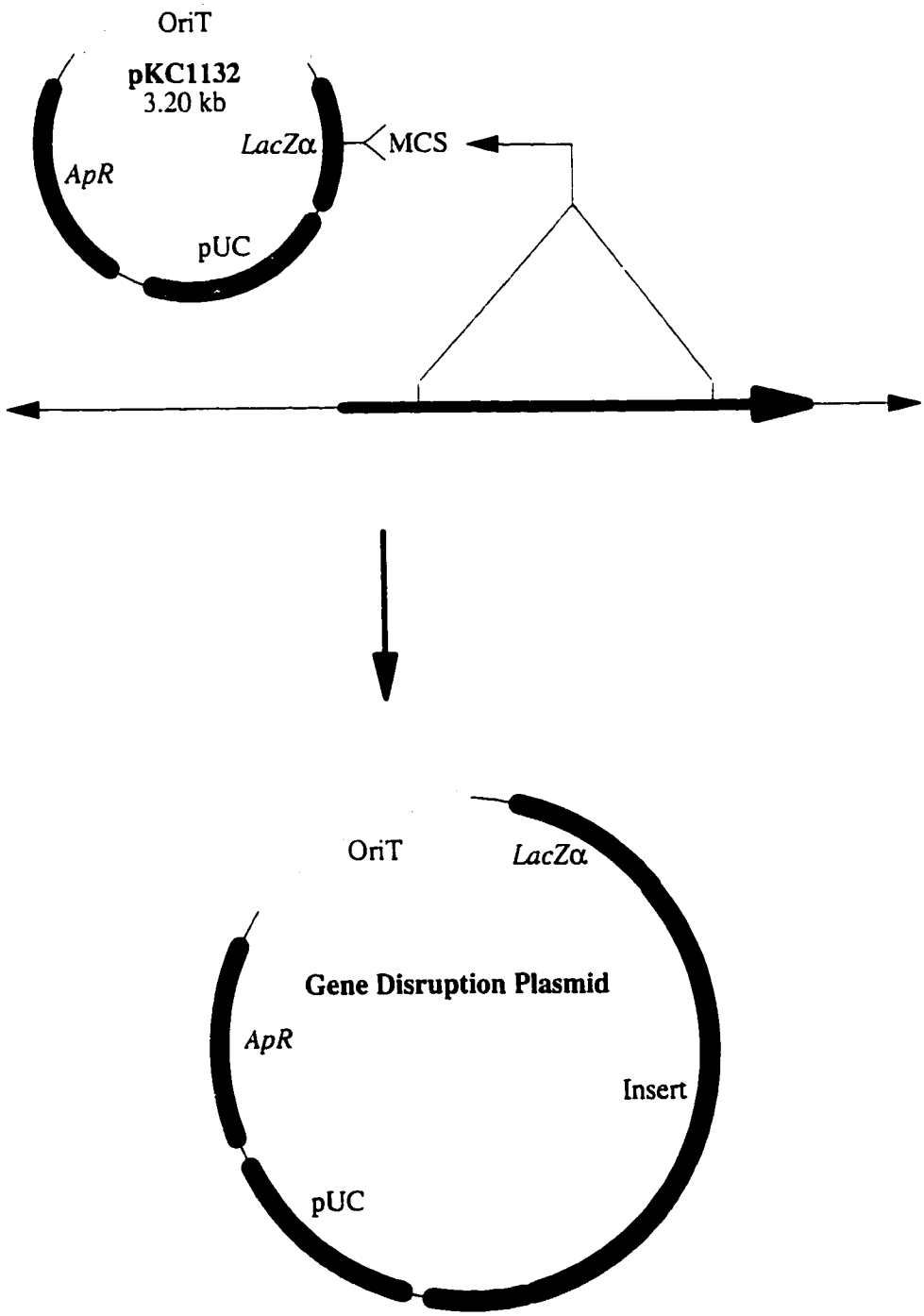
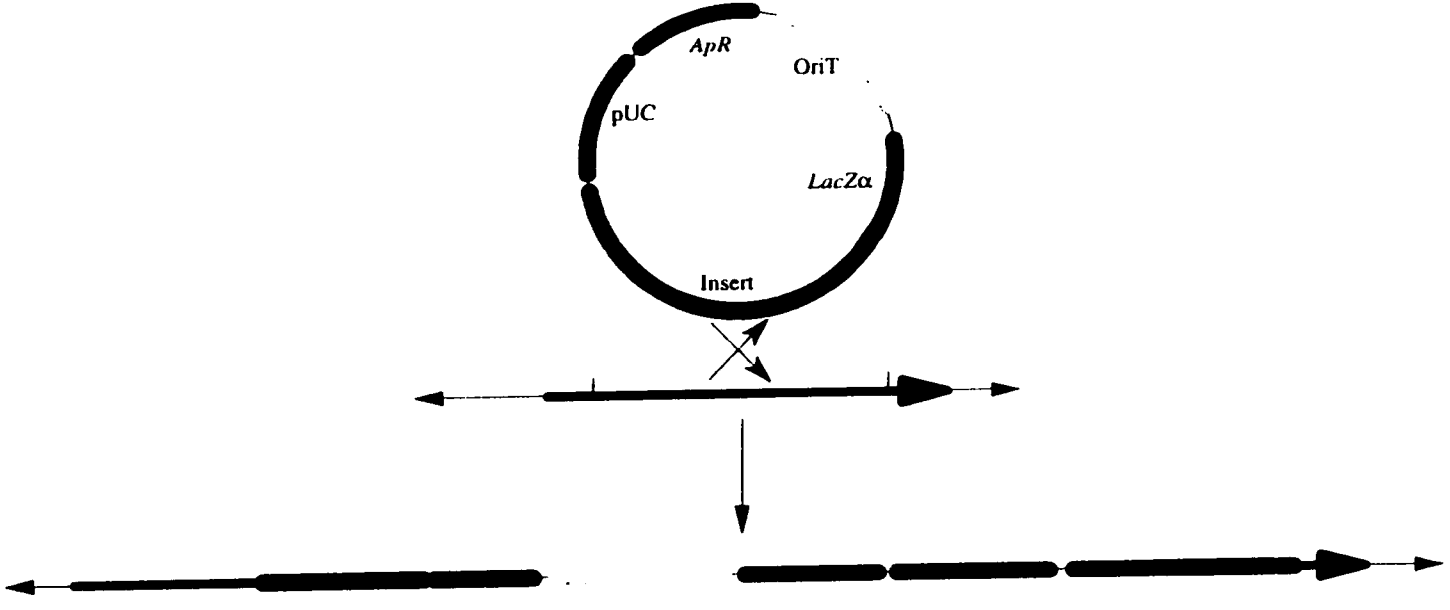

OriT
pKC1132
3.20 kb
LacZα
MCS
ApR
pUC
OriT
LacZα
Gene Disruption Plasmid
ApR
Insert
pUC

Figure 2.8: Single crossover insertional mutation. The DNA polymerase makes a mistake, and incorporates the DNA of the gene disruption plasmid into the genome at the homologous region.


ApR
OriT
pUC
LacZα
Insert

Chapter 3

RESULTS

3.1 Synthesis of enantiomerically pure HEQ and analogues

An initial goal of this project was to synthesize enantiomerically pure *R*-HEQ, and to re-synthesize several of the HEQ analogues that were used in the specificity study in order to fully characterize the structures for publication.

R-HEQ was synthesized from quinoline in a procedure involving nine synthetic steps (figure 3.1). Quinoline was converted to *N*-benzoyl-2-cyano-1,2-dihydroquinoline using chemistry developed by Reissert et al [55] in a reaction requiring benzoyl chloride and potassium cyanide. The product of this reaction was hydrolyzed by the aqueous hydrobromic acid/acetic acid system of Davis [13], yielding the hydrobromide salt of quinoline-2-carboxylic acid that was then readily converted to the methyl ester with thionyl chloride and methanol.

Introduction of an acetyl group at the 4-position of the quinoline nucleus was achieved by a modification of the homolytic aromatic substitution chemisty developed by Minisci et al [38][37]. An $FeSO_4$ system was used to generate MeCO· radicals from acetaldehyde, which under acidic conditions add to the protonated quinoline ring specifically at C-4, since the 2-position is blocked by the carbomethoxy group. Near quantitative yields were obtained when acetaldehyde was present in excess (used as

the solvent). Reduction of the ketone with $NaBH_4$ in methanol affords *rac*-HEQ in high yield. Hydrolysis of the methyl ester was achieved in a two phase basic system using NaOH in THF/H_2O. A sample of *rac*-HEQ was reserved for complete spectral analysis.

The synthesis of stereomerically pure *R*-HEQ was performed by converting methyl-HEQ to the butyrate derivative, and incubating this compound with lipase isolated from *Candida cylindracea*. This lipase shows a preference for the *R* isomer, and if the reaction is worked up before more than 10 percent of the butyrate has been hydrolyzed the resulting alcohol is highly enriched in the *R* isomer. This sample of enriched alcohol was derivatized using 1-(-)-camphanic acid chloride, and recrystallized. The NMR spectrum of the resulting diastereomer showed none of the *S*-HEQ derivative.

Several analogues of HEQ were re-synthesized in order to obtain final characterization data for publication. Methyl 4-hydroxymethyl-2-quinoline carboxylate was synthesized from methyl 2-quinoline carboxylate using Popp's modification [51] of Reissert-type chemistry [55]. This compound can be oxidized using the Swern method [32] to yield methyl 4-formyl-2-quinoline carboxylate.. The benzoyl and acetyl derivatives of methyl HEQ were formed by reaction of benzoyl chloride or acetyl chloride, respectively, with methyl HEQ.

The free acids of methyl 4-acetyl-2-quinoline carboxylate and methyl 4-methoxy-2-quinoline carboxylate were readily produced using a two phase water/THF basic hydrolysis (figure 3.2).

The NMR signals originating from the aromatic protons of HEQ had not been unequivocally assigned. In order to do so, HMQC and NOESY NMR spectra were

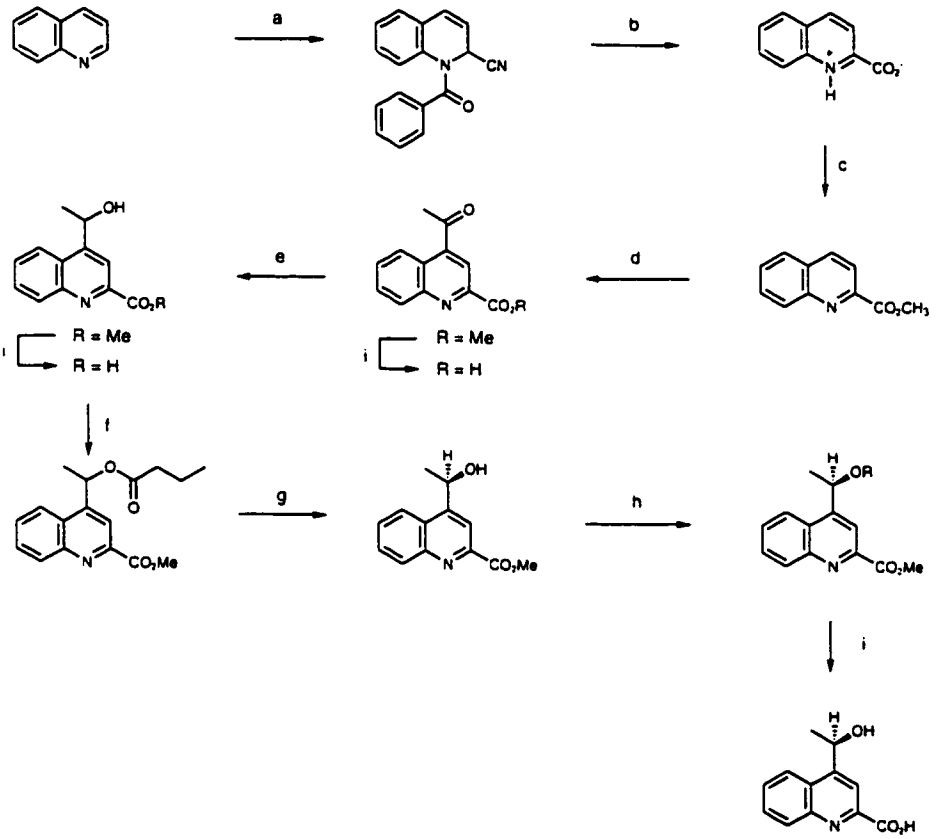


Figure 3.1: Synthesis of the *R* isomer of 4-(1-hydroxyethyl)-2-quinoline carboxylic acid. R = 1-(-)-camphanate. (a) PhCOCl, KCN, CH_2Cl_2; (b) HOAc, HBr, Δ; (c) NaOH, then $SOCl_2$, MeOH, Δ; (d) CH_3CHO, H_2O_2, $FeSO_4$, TFA; (e) $NaBH_4$, MeOH; (f) C_3H_7COCl, pyridine, CH_2Cl_2; (g) *Candida cylindracea* lipase, pH 7, DMSO-H_2O; (h) 1-(-)-camphanic acid chloride, pyridine, CH_2Cl_2; (i) NaOH-THF/H_2O.

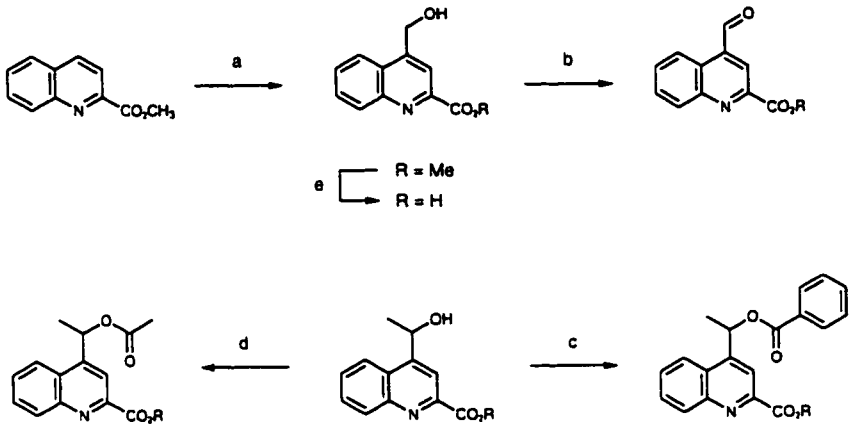


Figure 3.2: Synthesis of HEQ analogues. (a) TFA, H_2O_2, $FeSO_4$, MeOH. δ: (b) DMSO, $(COCl)_2$, $N(Et)_3$, CH_2Cl_2; (c) DEAD; $PhCO_2H$; Ph_3P, CH_2Cl_2; (d) CH_3COCl, pyridine, CH_2Cl_2; (e) NaOH-THF/H_2O.

obtained for the acid in D_2O. An NOE signal was seen from the methyl group to three of the aromatic protons, one of which had been assigned as the proton in the quinoline 3 position, and the other two can now be safely assigned to the protons in the quinoline 6 and 7 positions based on their coupling patterns. With positions 6 and 7 known, the protons in the 8 and 9 positions were readily assigned. HMQC correlations were observed from the newly assigned protons to the carbons directly bonded to them, finishing the assignment of the molecule.

3.2 Assay of mutants in the region surrounding the resistance gene

To conclude the study of the region surrounding the known thiostrepton resistance gene the mutants that were generated by direct insertion needed to be grown under conditions in which the wild type organism is known to produce thiostrepton, and the resulting cultures assayed for the production of thiostrepton. Two cultures of each mutant strain were grown. The mutant cultures were extracted with chloroform and assayed using a reverse phase HPLC system developed by Pei Zhou [80] which clearly separates the antibiotic from other extracted components. As described in section 2.4 only two of the mutant types were shown to be thiostrepton non-producing (figure 2.6), and both had been shown to be mutations in ribosomal protein operons [65]. In order to be sure that the peak observed by HPLC was indeed thiostrepton, fractions were collected from the end of the column and spotted onto antibiotic disks. These showed clear zones of no growth when suspended in soft agar that had been inoculated with *Staphylococcus aureus* and *S. lividans* (both thiostrepton sensitive) and incubated. *S. laurentii*, on the other hand, showed no growth inhibition under the same conditions. This and comparison with a thiostrepton standard confirmed our results.

Figure 3.3: Assay of thiostrepton production by strains of *S. laurentii* that contained mutational inserts in the region of the genome surrounding the known resistance gene.

A: HPLC profiles:

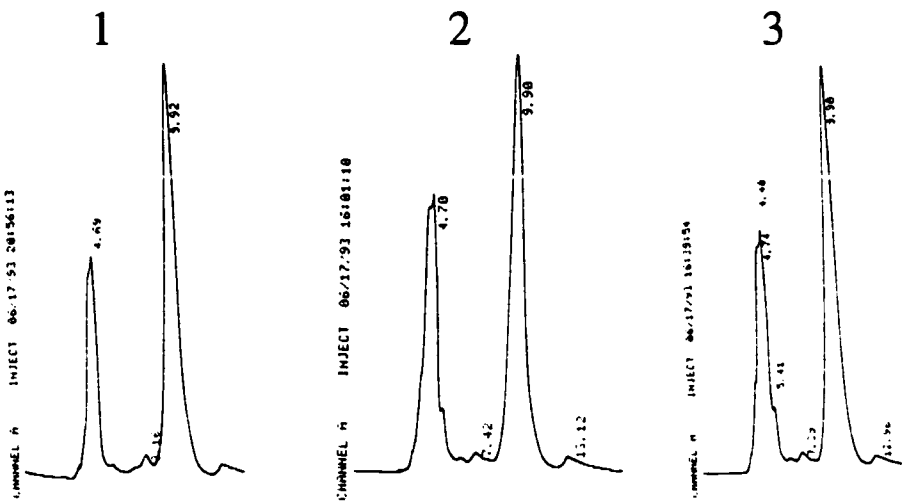
1. Standard thiostrepton, 150 μg/ml
2. Extracted wild type *S. laurentii* culture
3. Extracted mutant *S. laurentii* TS203 culture

The eluate corresponding with the peak at 9.9 minutes was collected, and antibiotic disks were spotted with 2.5, 5, 10, and 20 μl (counterclockwise from top right on each plate) and placed onto LB medium. Soft agar containing bacteria were used to overlay the plates containing the disks, which were then incubated at 30°C for two days.

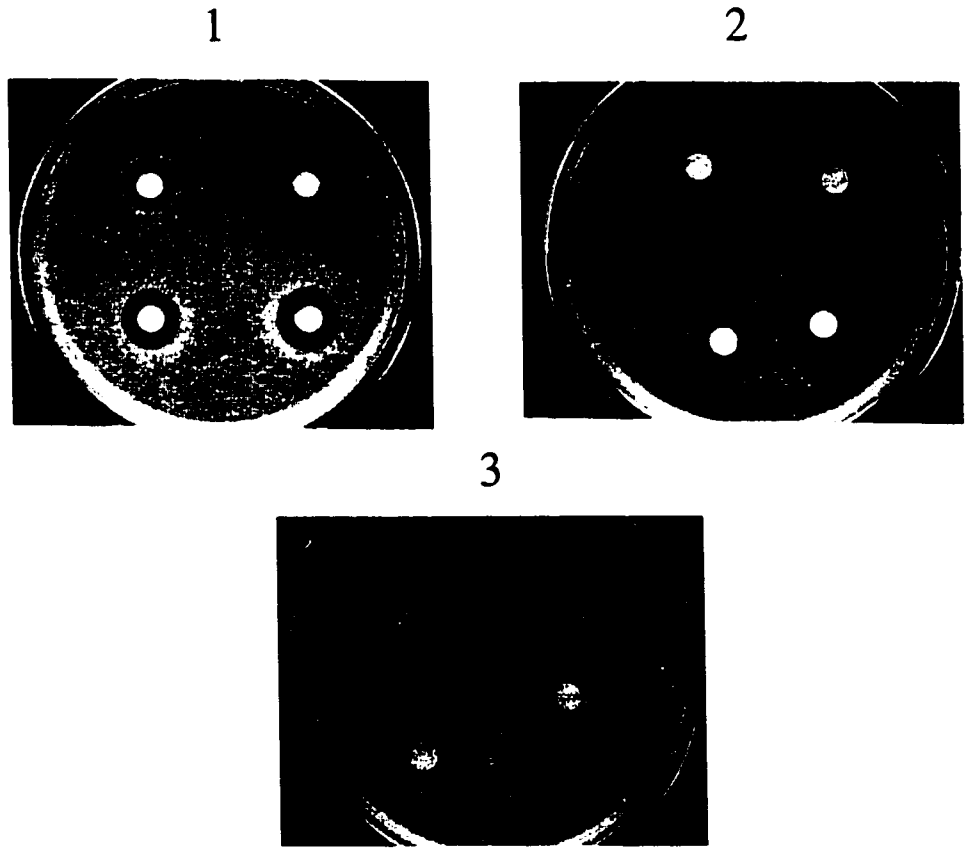
B: Bacterial growth inhibition by TS203 product

1. *S. aureus*
2. *S. lividans*
3. *S. laurentii*

A



B



3.3 Examining possible peptide synthetases

The data indicating whether the biosynthesis of thiostrepton proceeds through a ribosomal or a nonribosomal process were somewhat inconclusive (section 2.3.3). Studies were done to determine if there are any nonribosomal peptide synthetases present in *S. laurentii* cell-free extracts.

The reactions in the known nonribosomal peptide synthetase systems all require ATP to activate the amino acid prior to condensation with the following amino acid. The ribosomal mechanism uses GTP in the same role, allowing one to differentiate between ribosomal amino acid activation and nonribosomal amino acid activation.

These enzymes are capable of catalyzing a reaction of an amino acid and ATP to the acyl adenylate of the amino acid and pyrophosphate (figure 1.9). The presence of these enzymes can be detected by either the exchange of radiolabeled pyrophosphate into ATP in the presence of the substrate amino acid, or by measuring directly the binding of a radiolabeled amino acid to the peptide via the phosphopantetheine cofactor.

Cell-free extract was prepared and proteins precipitated with varying concentrations of ammonium sulfate. Radiolabeled pyrophosphate was exchanged into ATP in the presence of serine in the peptides precipitating between 40% and 65% ammonium sulfate saturation. These active fractions were loaded onto a DEAE cellulose anion exchange column and eluted at different salt concentrations. Serine, threonine, and ornithine (an unnatural amino acid which is unlikely to be activated by this system) activation was tested for each of the fractions, and it was found that the 500 mM NaCl eluate activated both serine and threonine, but did not activate ornithine. No

other active fractions were observed.

When this fraction was filtered through a 300 kDa cutoff ultrafiltration membrane it was found that none of the activities were observable in the buffer that had passed through the membrane, giving a lower boundary for the size of the enzyme or enzymes responsible for the activity at larger than 300 kDa. This active fraction was assayed for activation by incubation with ATP and radiolabeled amino acid, precipitation of the proteins in the reaction with TCA, careful washing of the pellet, and resuspension of the pellet in formic acid for scintillation counting. This should leave only enzyme bound amino acid to be counted. This assay was performed using ^{14}C labeled serine, threonine, and alanine, and ^{35}S labeled cysteine. All reactions were also performed in the absence of ATP as a negative control. Significant binding was observed in each of these cases. Addition of nonlabeled amino acid was performed to see if any inhibition was observed. Serine and threonine were each not affected by the addition of the other, or by the addition of alanine. Alanine activation was shown to be inhibited by the addition of unlabeled serine.

These results indicate that there is at least one peptide synthetase, at 300 kDa or larger, present in the cell-free extract. While there is no guarantee that the enzyme or enzymes responsible for the observed activity are involved in thiostrepton production, it is another observation supporting nonribosomal biosynthesis of the antibiotic.

Table 3.1: Putative peptide synthetase activities observed in a partially purified cell-free extract. Cell-free extract was fractionated with ammonium sulfate, and products precipitating at 40-65% ammonium sulfate saturation collected and pooled. After resuspension and desalting the resulting solution was applied to a DEAE cellulose anion exchange column and eluted with a NaCl step gradient, with the activity eluting at 500 mM. This fraction was loaded onto an ultrafiltration membrane with a size cutoff of 300 kDa. No activity was observed in the buffer that passed through the membrane. The buffer that remained above the filter contained all of the activity, indicating a size for the enzyme or enzymes greater than 300 kDa. Values reported for pyrophosphate exchange are the ratio of observed radioactivities between the reaction with amino acid substrate to the reaction without. Values reported for amino acid binding are the ratio of measured radioactivities for the reactions with ATP and without ATP. Inhibition studies were performed using cold serine, threonine, and alanine. The only observed inhibition was that of serine inhibiting alanine binding. n.d. = not determined.

Amino acid	^{32}P PP_i exchange	Amino acid binding	Binding inhibition
Serine	2.3	4.4	none
Threonine	3.1	1.7	none
Ornithine	0.9	n.d.	n.d.
Alanine	n.d.	3.4	serine
Cysteine	1.6	1.5	n.d.

3.4 Searching for the gene cluster

After these early studies, the focus of the project shifted to directly searching for the biosynthetic cluster in *S. laurentii*. Several approaches were undertaken in this study to find the cluster. Southern hybridization with a known Streptomycete nonribosomal peptide synthetase encoding gene against *S. laurentii* genomic DNA was used as an initial approach to finding the gene cluster. A reverse genetics approach in which an enzyme known to be on the pathway is purified, the amino acid sequence determined, and that sequence used to devlelop Southern probes or PCR primers to identify the encoding gene in a cosmid library was employed using the partially characterized HEQ adenylase discussed in section 2.2.1, page 32. Finally, a method that had been recently developed using the highly conserved core amino acid sequences found in nonribosomal peptide synthetases to design PCR primers for amplification of fragments of these genes from the genome was used.

3.4.1 Southern probing using homologous DNA

The traditional method used to find secondary metabolite gene clusters directly, and the first approach taken in this study, is to synthesize a Southern hybridization probe from a DNA template containing a gene that has a function similar to that gene which one is trying to isolate. This approach takes advantage of the high degree of concordance found in the primary amino acid sequence of enzymes that catalyze the same general reaction. The disadvantage of this approach is that there are multiple ways to encode for a given protein sequence, with as many as six possible codon sequences for each amino acid [82]. Streptomycete DNA is known to have a very high

GC ratio, with as much as 75% of Streptomycete DNA consisting of GC pairs, with the GC content in the third position of codons above 90% [41]. This significantly decreases the coding degeneracy found in Streptomycetes, but results in a low likelihood of non-Streptomycete DNA showing significant hybridization to Streptomycete genes of interest. When an appropriate probe has been synthesized, it is first used to hybridize against restriction endonuclease digested genomic DNA. If a significant hybridization is observed, giving a clear band on autoradiography film, the probe can be used in a colony hybridization experiment against a cosmid library that has been generated from the species' genomic DNA. Each cosmid in the library contains a randomly generated 50-80 kb fragment of the genome. *E. coli* cells, transduced by the cosmid library in such a manner so as to have no more than one cosmid infecting a given cell, are hybridized to the probe using colony hybridization [59]. Colony hybridization is a method based on Southern hybridization where colonies of bacteria are grown on a DNA binding membrane. Once grown to sufficient size, the colonies are lysed on the membrane, the DNA released by the lysing bound to the membrane, and the membrane washed to remove the cell debris. This can then be used as in a Southern hybridization experiment, and colonies showing hybridization to the probe can be isolated and the cosmid DNA purified. The general result is a number of overlapping cosmids, each containing the segment of DNA hybridizing to the probe.

Most of the peptide antibiotics being studied are not being produced by Streptomycetes or other highly GC rich species, and would not generate probes with sufficient homology to a Streptomycete gene. However, the δ-(*L*-α-aminoadipyl)-*L*-cysteinyl-bis-*D*-valine synthetase (ACVS) gene, a nonribosomal peptide synthetase gene in-

volved in the biosynthesis of penicillin in *Streptomyces clavuligerus* [25] had just been isolated by the Jensen lab at the University of Alberta, and a fragment of the gene known to contain conserved domains was kindly provided by them.

Randomly primed DNA fragments were generated using this template, and labeled by the incorporation of [α^{32}P]dCTP during synthesis. This probe was then used to hybridize against restriction endonuclease digested genomic DNA from *S. laurentii* and *S. azureus*, thiostrepton producers, and *S. actuosus*, the nosiheptide producer. As a control the gene fragment from *S. clavuligerus* that had been used as the template to generate the probe was hybridized against as well. After hybridization to the probe, the membrane was washed first at low stringency. At this stringency the probe bound non-specifically to all of the digested genomic DNA, as well as to the positive control. A second wash with a slightly higher stringency was performed which removed virtually all of the probe from the digested genomic DNA while retaining the hybridization to the positive control (figure 3.4). With this lack of specific hybridization it is highly unlikely that the probe would be useful in screening a cosmid library. At the time these experiments were conducted this gene was the only Streptomycete nonribosomal peptide synthetase gene isolated, so hybridization using other genes was not performed.

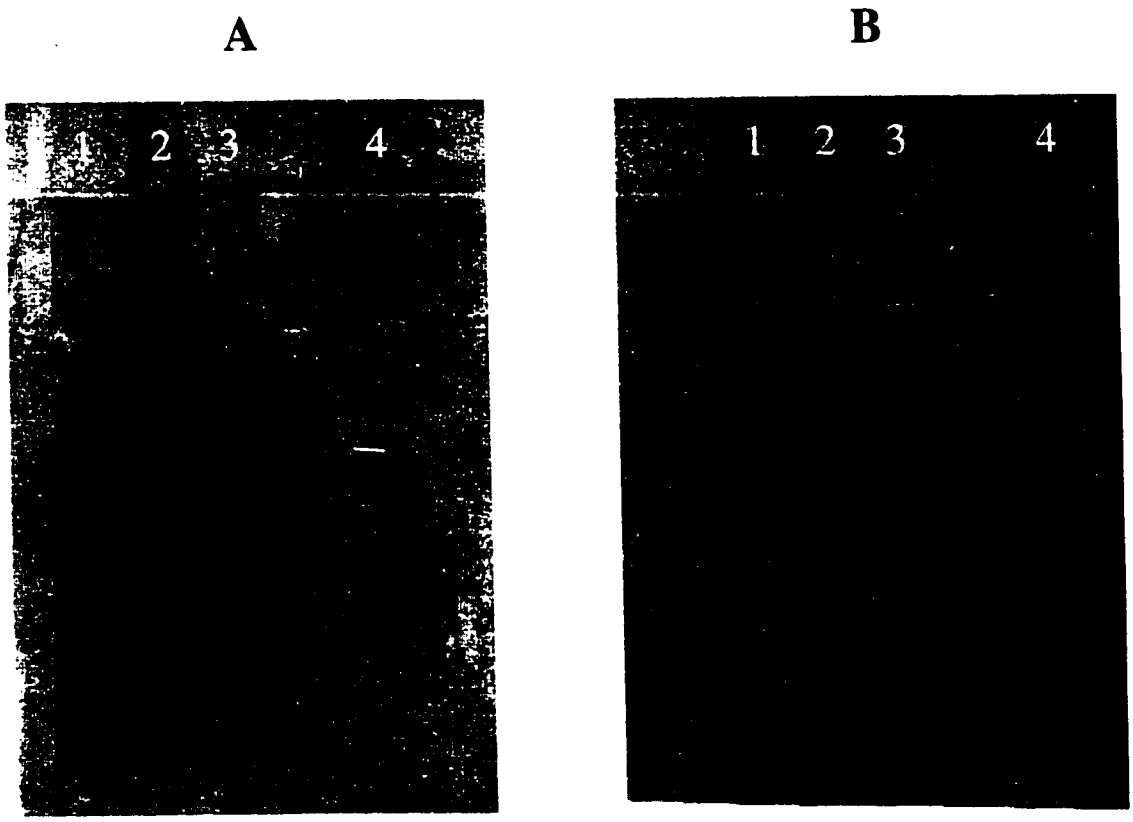


Figure 3.4: Southern hybridization of the probe synthesized using the ACVS gene fragment against restriction endonuclease-digested genomic DNA, and the original template as a positive control. Panel A shows the autoradiography film of the membrane after washing at low stringency (2xSSPE, 0.1% SDS at 50°C). Panel B shows the autoradiography film of the membrane after washing at slightly higher stringency (2xSSPE, 0.1%, 60°C).

lane	Sample
1	*S. azureus* genomic DNA, digested with *Bam*HI
2	*S. actuosus* genomic DNA, digested with *Bam*HI
3	*S. laurentii* genomic DNA, digested with *Bam*HI
4	XP6 plasmid from *S. clavuligerus* containing part of the ACVS gene, insert excised with *Bam*HI

3.4.2 The HEQ adenylase

Another approach previously undertaken in studies to find biosynthetic gene clusters is reverse genetics. An enzyme known to be on the biosynthetic pathway is purified and a partial amino acid sequence determined. This sequence can be used to design and synthesize an oligonucleotide which is labeled and then used as a Southern hybridization probe against a cosmid library in order to find the gene encoding the enzyme.

Two thiostrepton biosynthetic enzymes have been identified and partially purified. The tryptophan methyl transferase described in section 2.2.1, page 28, catalyzes the methylation of the indole two position of tryptophan, and the HEQ adenylase (section 2.2.1, page 32), which catalyzes the formation of the acyl adenylate of HEQ.

The HEQ adenylase was chosen as the target enzyme for a number of reasons. Firstly, the tryptophan methyl transferase was found to be very unstable, and probably membrane bound [18]. Initial studies on the HEQ adenylase included some partial purification work, and the assay for activity was shown to be rapid and straightforward.

Assaying for HEQ adenylase activity takes advantage of the reversibility of the activation reaction. The reaction

$$HEQ + ATP \rightleftharpoons HEQ - adenylate + PP_i$$

is capable of exchanging radioactivity from ^{32}P-labeled pyrophosphate into ATP. An assay which takes advantage of this and the strong binding of ATP to activated charcoal was developed by Keller *et al.* [26] in the study of the incorporation of 3-methyl-4-hydroxyanthranilic acid into the antibiotic actinomycin in *Streptomyces*

chrysomallus. This method was further refined for *S. laurentii* by Smith [64]. This assay allows for rapid screening of large numbers of fractions for activity.

Initial studies confirmed the work of Smith [64] who found that the adenylase does not bind appreciably to DEAE cellulose, but that passing the cell-free extract through this anion exchange resin provides a 5-10 fold purification, and that the adenylase precipitates between 35% and 65% concentrated ammonium sulfate.

A problem noticed early in studies of this enzyme is its low stability. At pH 6.8, the stability maximum, in degassed phosphate buffer containing DTT, the half life of the activity was found to be about eight hours at 4°C. The stability drops noticeably at higher or lower pH, and in TRIS buffer. Glycerol appears to enhance the stability of the enzyme, and was included in buffers used in the initial purification of the cell-free extract and some of the chromatography steps.

The first approach taken to purify the adenylase used fast protein liquid chromatography (FPLC) columns after the initial DEAE cellulose and ammonium sulfate purification steps.

Hydrophobic interaction chromatography separates proteins based on their tendency to form aggregates with nonpolar compounds in a polar solvent. This reduces the entropic decrease created by the formation of a structural cavity in the solvent around the nonpolar solute. The more nonpolar the solute the more likely the formation of this aggregate with the nonpolar chromatography matrix. Structure forming salts like ammonium sulfate increase this interaction, and aid in the binding of the protein to the resin. The column material can either be a long alkyl chain attached to the resin, or the resin can be linked to an aromatic group via an alkyl chain. The

enzymes are generally bound to the column at a high ammonium sulfate concentration, then eluted by either a gradient or stepwise decrease in the concentration of the structure forming salt [48].

Cell-free extract was partially purified by a negative binding step using DEAE cellulose, then the proteins which precipitated between 35% and 65% ammonium sulfate saturation were collected. This pellet was resuspended and desalted into 50 mM phosphate buffer, pH 6.8, 1 mM DTT (buffer A) containing 30% ammonium sulfate. The resulting solution was injected onto a Phenyl Superose hydrophobic interaction FPLC column. The activity was found in fractions eluting at 0% ammonium sulfate concentration, indicating a fairly hydrophobic protein (figure 3.5).

The next chromatographic step involved the use of a strong anion exchange resin. Proteins bind to this column by the interaction of negatively charged groups on the peptide binding to quaternary ammonia ions covalently linked to the column resin. Elution is generally achieved by a salt gradient increasing the ionic strength of the buffer. Active fractions from this column were pooled and the buffer exchanged by gel filtration to 100 mM TRIS buffer, pH 8.0, 1 mM DTT (buffer B). By using a high pH the enzyme should have a higher total negative charge due to deprotonation of acidic functional groups. Unfortunately, as stated before the stability of the enzymatic activity is markedly lower at this high pH. This sample was loaded onto a Mono-Q anion exchange column that had been equilibrated with the same buffer and the activity eluted with a KCl gradient. The enzyme eluted at fairly low salt concentration, in fractions that showed a low absorbance at 254 nm (figure 3.6). This is a very effective purification step, which somewhat offsets the loss of activity observed at the higher

pH.

As rapidly as possible, the active fractions that were collected from the anion exchange column were pooled and exchanged back into buffer A by gel filtration. The Mono-S strong cation exchange FPLC column binds to positive charges on the protein via sulfate groups covalently attached to the resin. It was found that the activity bound to the column at pH 6.8, allowing for chromatography at the stability optimum. The protein mixture was applied to the column, which had been equilibrated with buffer A, and eluted with a KCl gradient. From the activity of the eluted fractions it appears that about a third of the activity did not bind to the column, but that which did was only eluted at fairly high KCl concentration (330 mM) (figure 3.7). No absorbance at 254 nm was observed in this experiment, probably due to low protein concentration.

The products of these columns were separated on an SDS-PAGE gel which showed no bands in the expected size range in the active fractions from the Mono-S column (figure 3.8). Apparently this protein existed in low enough concentration that it is unobservable by Coomassie Blue staining, but still showed strong activity. Some attempts were made to run larger amounts of partially purified cell-free extract by multiple injections and consolidation of active fractions, but invariably the activity degraded before completion of this purification scheme.

Clearly another bulk purification step is needed before moving to FPLC. In affinity chromatography a compound to which the enzyme is known to bind is covalently linked to a resin in hopes that the attachment to the resin does not interfere with the binding of the enzyme to this structure. The covalently attached molecule can be a

substrate for the enzyme, or be similar enough to a substrate that the enzyme may still bind. An HEQ affinity column was synthesized in an earlier study [63]. This column was shown to not have any affinity for the adenylase under the conditions tested. This leaves ATP as the most likely candidate for binding. Kinetic studies have shown that the apparent K_m of the adenylase for HEQ is about ten fold smaller than that for ATP (section 2.2.1, page 32), hopefully also indicating that the binding is less selective and that the presence of a linker arm to a resin does not inhibit binding the protein as was observed with the HEQ affinity resin. There are a variety of different commercial ATP affinity resins with different attachments to the molecule available, increasing the chance that an adenylase binding resin will be found.

Three ATP affinity columns were assayed for HEQ adenylase binding. They consisted of ATP bound to the column matrix through either the N6 position, the C8 position, or the ribose hydroxyl groups of ATP via a 9-11 carbon alkyl linker arm. The resins were suspended in buffer A and degassed, then test columns were poured with an approximately 200 μl settled volume. Aliquots of partially purified cell-free extract, containing about 0.5 mg total protein, which had been passed through DEAE cellulose and fractionated with ammonium sulfate, were loaded onto the columns. The columns were washed with buffer A, then with buffer A containing 10 mM ATP, followed by buffer A containing 100 mM ATP. Proteins present in the fractions were precipitated with ammonium sulfate before assaying for activity to remove the ATP present in the elution buffers. The affinity column containing ATP attached via N-6 and the affinity column with ATP attached via the ribose hydroxyl groups showed no affinity for the enzyme, but the column with ATP attached through C8 retained

a fair amount of emzyme activity (figure 3.9).

The profile of the affinity column with ATP attached via C8 showed enough retention to warrant further investigation. A larger column was constructed (1 ml) and onto it was loaded cell-free extract from two liters of cell culture, partially purified by DEAE cellulose and an ammonium sulfate fractionation. This column was capable of retaining all of the activity in the sample until eluted with 100 mM ATP (figure 3.10). When the sample is concentrated and desalted with TCA and analyzed using SDS-PAGE significant purification is evident (figure 3.10). One major disadvantage of using this column is the high loss of activity observed. Surprisingly, even with the eight fold reduction in activity observed in the purified extract an eleven fold purification of the protein is achieved by this chromatography step. This loss of activity, however, proved to be problematic in future purification steps. Purification to a distinct band of the expected size was not achieved by FPLC after the ATP affinity step before all activity was lost, despite several attempts and an increase in the total amount of cell culture used in the purification.

Another fast bulk purification step is needed early in the protocol. In hopes of improving on the ATP affinity chromatography step Affi-Gel Blue, a dye affinity matrix, was screened for adenylase binding. Unfortunately this resin either deactivates the adenylase during chromatography, or binds it too tightly to elute as no activity was observed in any of the eluted fractions.

A bulk anion exchange resin that will bind the adenylase was sought to use in place of the negative binding experiment done with DEAE cellulose. DEAE Sephacel was screened for binding the adenylase at several pH's from 6.8 to 7.8 in 50 mM

phosphate buffer containing 10% glycerol and 1 mM DTT. No binding was observed in this pH range, and it was decided that the rapid loss of activity at even higher pH values would make the step unsuitable early in the purification scheme. A slightly better purification was obtained by running the cell-free extract through this column as opposed to the DEAE cellulose negative binding step.

Since the protein does not bind easily to anion exchange columns, the next logical step is to test for binding to a cation exchange resin. CM Sepharose is a cation exchange resin that interacts with proteins via a carboxymethyl group covalently linked to the resin. This resin was assayed for adenylase binding at pH's ranging from 5.8 to 6.8. It was found that on a small scale binding was first observed at pH 6.4, and the entire activity was bound at pH 6.0. Subsequent experiments used the column at pH 5.8 to ensure full retention of activity.

Cell-free extract was partially purified by passing through DEAE Sephacel and ammonium sulfate fractionation. The pellet was resuspended in buffer C (50 mM phosphate pH 5.8, 10% glycerol, 1 mM DTT) and desalted by dialysis. This solution was applied to an equilibrated CM Sephadex column and eluted with a NaCl step gradient, with the activity eluting in the 200 mM salt wash. This fraction was applied to the ATP affinity column directly. Unfortunately, little of the activity bound to the affinity column in this buffer, with the majority of the activity being found in the flow through. This could be due to the high salt concentration, or the presence of glycerol. Some activity was measured in the 100 mM ATP wash, although analysis of these samples by SDS-PAGE showed no protein bands in the expected size range (figure 3.11).

The lack of binding to anion exchange resins, and the extreme pH required for binding to the CM Sephadex cation exchange column together imply that the protein is most likely quite hydrophobic. The strong binding to the Phenyl Superose FPLC column supports this conclusion as well. Phenyl Sepharose 6 (high substitution) bulk hydrophobic interaction resin was screened for adenylase binding. It was found that the protein completely bound to the column, and that the activity remained bound to the column even when eluted with buffer A, and that buffer A with 20% glycerol was required to elute the activity from the column. An advantage of this chromatographic step is that it can be used immediately following an ammonium sulfate precipitation without desalting as the loading buffer contains ammonium sulfate, allowing for rapid buffer exchange before the column.

Cell-free extract was passed through DEAE Sephacel, yielding a seven fold purification as expected. The flow through was fractionated by ammonium sulfate precipitation, and the proteins precipitating between 35% and 65% saturated ammonium sulfate resuspended in buffer A containing 1 M ammonium sulfate and loaded on the Phenyl Sepharose column. After washing the column with buffer A the activity was eluted with buffer A containing 20% glycerol. These two steps together afforded a 130 fold purification. The protein was precipitated with ammonium sulfate, resuspended in buffer C, and desalted by gel filtration. The resulting solution was loaded onto the CM Sephadex column and eluted with a salt gradient, giving a thirty fold purification in this step. At this point, however, the protein concentration was again too low to give clearly visible bands at any size by Coomassie Blue staining after SDS-PAGE analysis (figure 3.12).

A purification scheme needs to be designed allowing rapid purification of the protein with minimum degradation. Step gradients on bulk resins afford rapid protein purification. It has been observed that glycerol enhances the stability of the protein, and that significant activity is lost during ammonium sulfate precipitation. Dialysis degrades the protein due to the long time required to exchange buffers. As the protein is considerably less stable at pH values higher and lower than 6.8, a scheme designed to minimized buffer exchanges, and maintain the pH at 6.8 would minimize degradation of the activity during purification.

A much stronger bulk cation exchange resin had recently been developed and marketed under the name SP Sepharose. The enzyme was demonstrated to bind to this resin easily in buffer A containing 10% glycerol, with the activity eluting from the column when washed with 200 mM KCl.

This allows for the design of a purification that keeps the protein at the optimal pH and does not involve any buffer exchange steps. Cell-free extract was passed through DEAE Sephacel, and precipitated directly with 65% saturated ammonium sulfate. The pellet was resuspended in buffer A containing 1 M ammonium sulfate and loaded onto the Phenyl Sepharose column which was then washed with buffer A. The protein was then eluted with buffer A containing 20% glycerol. This fraction was applied directly to the SP Sepharose column, and the active fraction precipitated with TCA. A quarter of this sample, and samples saved from the other purification steps, were analyzed by SDS-PAGE, and an easily identified protein band of the correct size was observed (figure 3.13).

It was decided that performance of an Edman [78] degradation was worthwhile to

determine the amino acid sequence of the N-terminus of this protein band. Half of the remaining TCA precipitate was separated by SDS-PAGE and transferred to a PVDF membrane. The membrane was lightly stained and the band excised for submission for N-terminal protein sequencing. The sequencing reaction was successful, giving 14 amino acids at high confidence levels, and another ten less certain cycles (figure 3.14).

An oligonucleotide was designed based on this sequence and the codon preference of *Streptomycete* DNA [41] (figure 3.16). This fragment was end labeled using $\gamma-^{32}$P labeled ATP and T4 kinase [59], and hybridized against *Bam*HI restriction endonuclease digested *S. laurentii* genomic DNA. No specific hybridization was observed to any fragments. This experiment suffers, however, from lack of a positive control.

In another attempt to find the gene encoding the sequenced protein the remaining TCA precipitated protein solution was transferred to a PVDF membrane and submitted for N-terminal sequencing of protease generated fragments, which should give internal sequence information of the protein. Of the fragments analyzed, only two gave useable amino acid sequences (figure 3.15). Oligonucleotides were designed based on these sequences to use as reverse PCR primers (figure 3.16). By requiring the match of two primers instead of just one oligonucleotide the specificity of this approach should be considerably higher than the Southern hybridization.

These primers were used to amplify fragments from *S. laurentii* genomic DNA using *Pfu* polymerase, *Pfu* Turbo polymerase, and *Taq2000* polymerase (Stratagene, Inc.) (figure 3.17). Several annealing temperatures and DNA polymerases were tested for amplification, and it was found that at 60°C the *Taq2000* polymerase cleanly amplified a 350 bp product when the T-28 reverse primer was used, and a 200 bp

product with the T-38 reverse primer. These fragments were purified, reamplified, and cloned as described in section 3.4.3, page 88.

Plasmids containing inserts of the appropriate size were isolated and the inserts sequenced using PCR cycle sequencing. It would be expected that the larger fragment, amplified using the T-28 reverse primer, would contain the sequence of the product amplified using the T-38 reverse primer, but this was not observed. Only the product generated by amplification with the T-38 primer contained both primers in the same reading frame. This product showed some homology to aminopeptidases, but not to any known acyl adenylases (figure 3.18). The product obtained using the T-28 primer only showed significant homology to transcriptional regulators (figure 3.19). In either case it is unlikely that this protein is the acyl adenylase.

Since this was the major product in the expected size range after the latest purification attempt, and previous purification attempts had yielded active fractions that did not show any proteins visible by Coommassie Blue staining, it is likely that this enzyme is present in very low concentrations in the cell-free extract. This possibility, especially when combined with the low stability of the protein, makes purifying this enzyme for eventual protein sequencing not a viable route to finding the thiostrepton biosynthetic cluster.

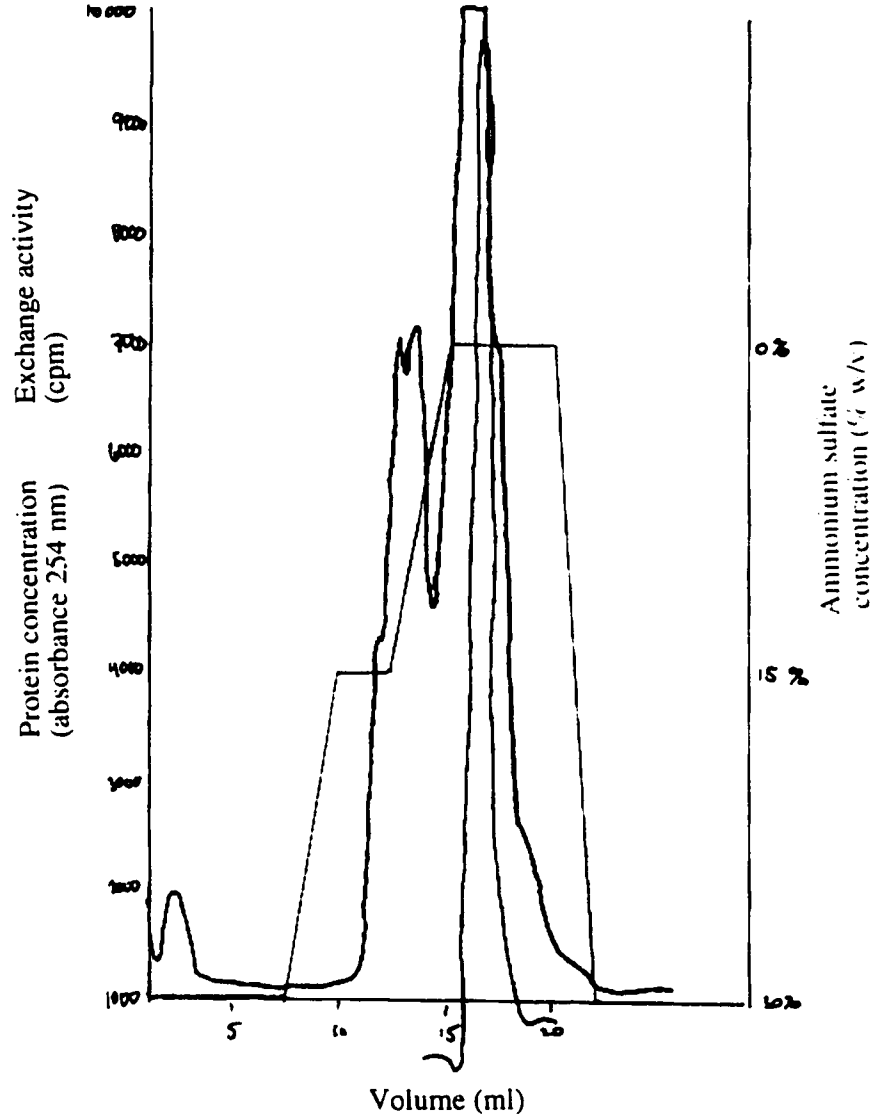


Figure 3.5: Results from the Phenyl Superose hydrophobic interaction FPLC column. Cell free extract had been purified using a negative binding purification step using DEAE cellulose followed by a 35% to 65% ammonium sulfate saturation. The sample was exchanged to buffer A containing 30% w/v ammonium sulfate and loaded on a column equilibrated with the same buffer. The activity was eluted with a gradient from 30% to 0% w/v ammonium sulfate in buffer A. The blue line is the absorbance at 254 nm, a rough indicator of protein concentration. The black line indicates PP_i exchange activity in the presence of HEQ. The red line represents the ammonium sulfate concentration.

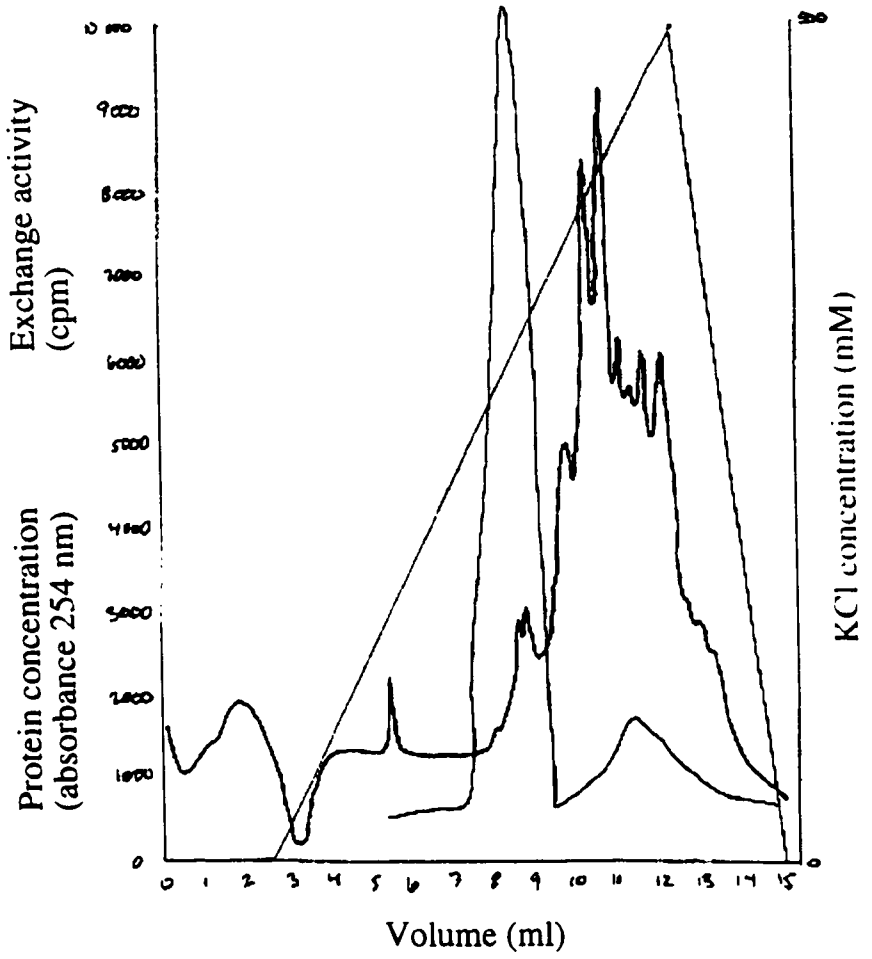


Figure 3.6: Results of the Mono-Q FPLC column on the active fractions from the Phenyl Superose column. The protein binds weakly to the resin, even at a highly elevated pH. Blue indicates absorbance at 254 nm, red the concentration of KCl in the buffer, and black the pyrophosphate exchange activity.

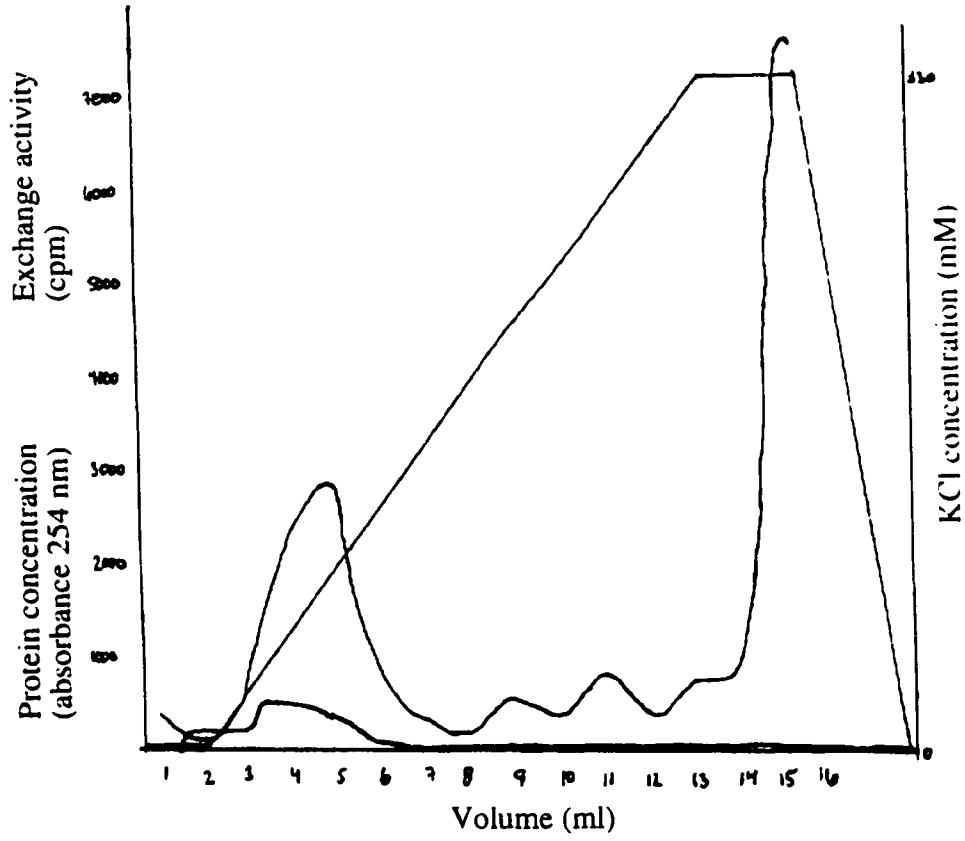


Figure 3.7: Results of chromatography using the Mono-S strong cation FPLC column. Active fractions from the Mono-Q column were pooled, and the buffer exchanged to buffer A by gel filtration. The resulting solution was loaded onto a Mono-S column equilibrated with the same buffer and eluted with a KCl gradient. About a third of the activity did not bind to the column, with the remainder eluting at about 330 mM KCl. No absorbance was observed at 254 nm, possibly due to low protein concentration. Blue represents the absorbance at 254 nm, red the concentration of KCl in the buffer, and black the pyrophosphate exchange activity.

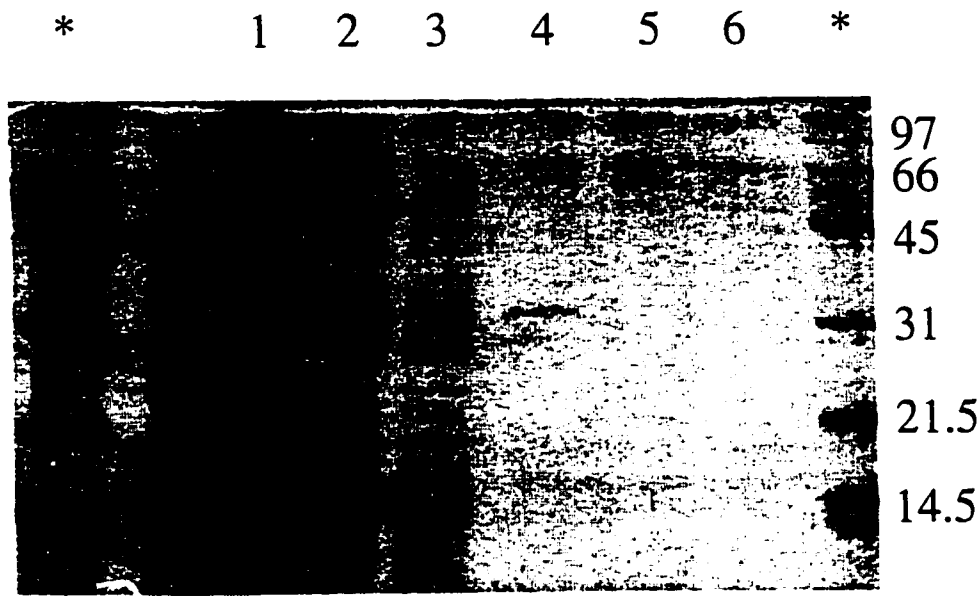
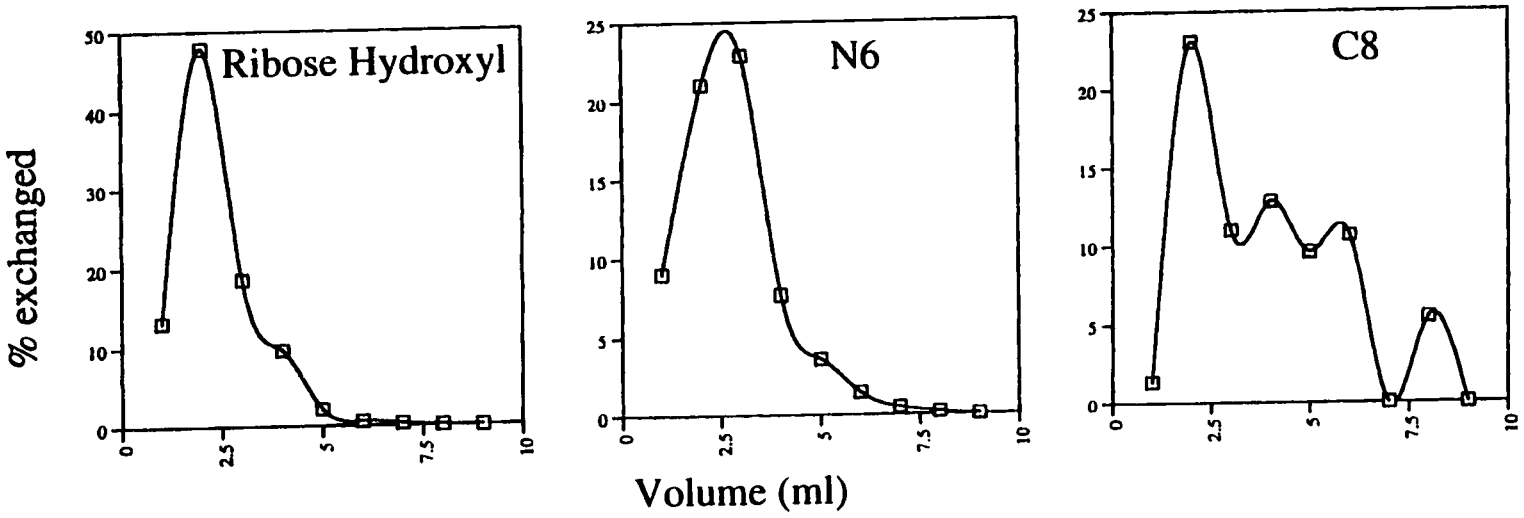


Figure 3.8: SDS-PAGE of active fractions from the FPLC purification. Lanes marked with an asterix are molecular weight standards. Lane 1 is the combined active fractions after the Phenyl Superose column. Lane 2 is the combined active fractions after the Mono-Q column. Lanes 3, 4, 5, and 6 are fractions from the Mono-S column. Lanes 3 and 4 are from the early fractions containing activity that did not bind to the column. Lane 5 is from a fraction that showed no activity, and lane 6 is from the active fraction that was eluted at 330 mM KCl. The gel was stained with Coommassie Blue. No protein bands of the expected size were observed in the active fraction binding to the Mono-S column.



$$percent\ exchanged = \frac{cpm_{positive} - cpm_{negative}}{total\ activity\ in\ reaction}$$

Figure 3.9: Results of the screening of the three ATP affinity columns. Columns tested had ATP bound to the column by a 9-11 carbon alkyl chain, with the linker arm attached to ATP at either the N-6 or C8 positions, or to the ribose hydroxyl groups. The C8 attached column was the only one to show affinity for the adenylase, with some of the activity staying strongly bound to the column requiring elution with 10 mM ATP. Activity is reported as percent of the pyrophosphate exchanged into ATP, calculated by the equation given below the diagrams. The negative control is performed by assaying without HEQ in the reaction, giving the background exchange activity. The total activity is the amount of radioactivity, experimentally determined, in the pyrophosphate added to the reaction.

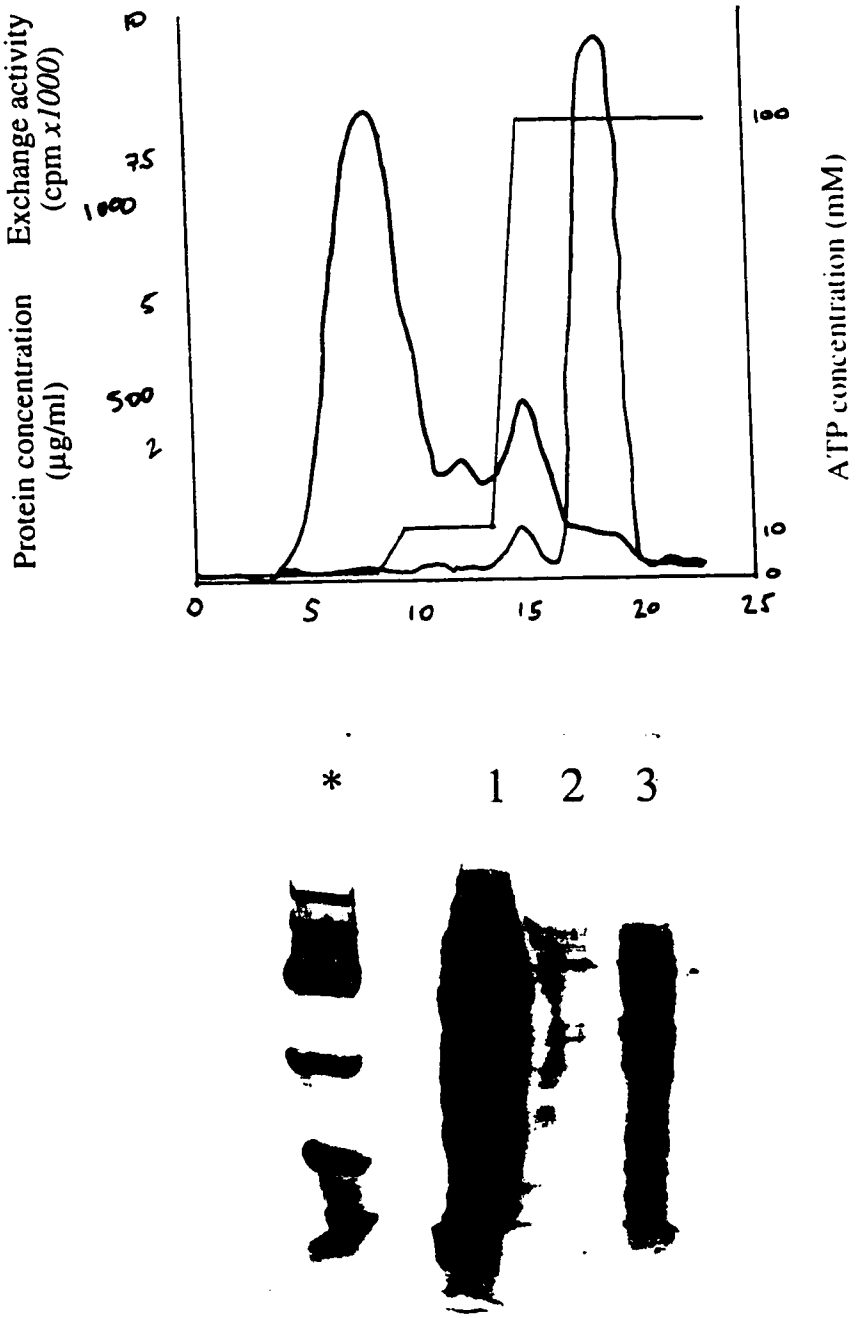


Figure 3.10: After testing different ATP affinity chromatography resins, the resin with ATP attached via the C8 position was chosen to do a full scale test. All observable activity bound to the column, and was not eluted until the column was washed with 100 mM ATP. In the upper graph red represents ATP concentration (mM), black represents exchange activity (cpm x 1000), and blue the protein concentration in μg/ml. Lane 1 of the SDS-PAGE gel is the sample loaded onto the affinity column, lane 2 is the unconcentrated active fraction, and in lane 3 the active fraction was concentrated using TCA precipitation.

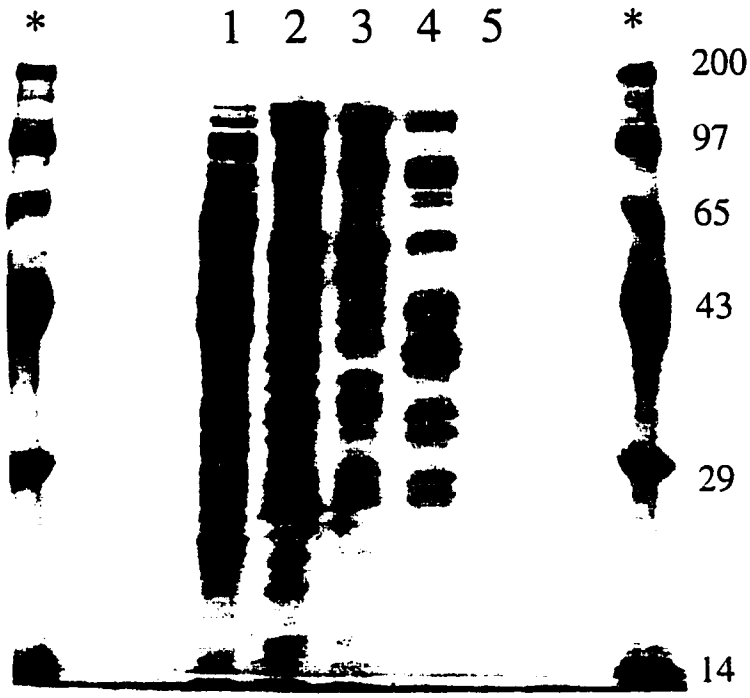


Figure 3.11: SDS-PAGE analysis of active fractions collected from a purification attempt using CM Sephadex. Lane 1 is the remaining protein after passing through DEAE Sephacel and ammonium sulfate fractionation. Lane 2 is the active fraction after exchanging the resuspended ammonium sulfate pellet into buffer C, and chromatography through CM Sephadex using a NaCl step gradient. Lane 3 is the flow through fraction resulting from the direct application of the active CM Sephadex fraction to the ATP affinity column. This fraction contained approximately 75% of the activity. Lane 4 is the result of eluting the ATP affinity column with 10 mM ATP, in which no activity was observed. Lane 5 is the 100 mM wash of the ATP affinity column. About 25% of the activity was observed in this fraction. There are no protein bands of the expected size observed in this lane, so the purification was not continued beyond this point.

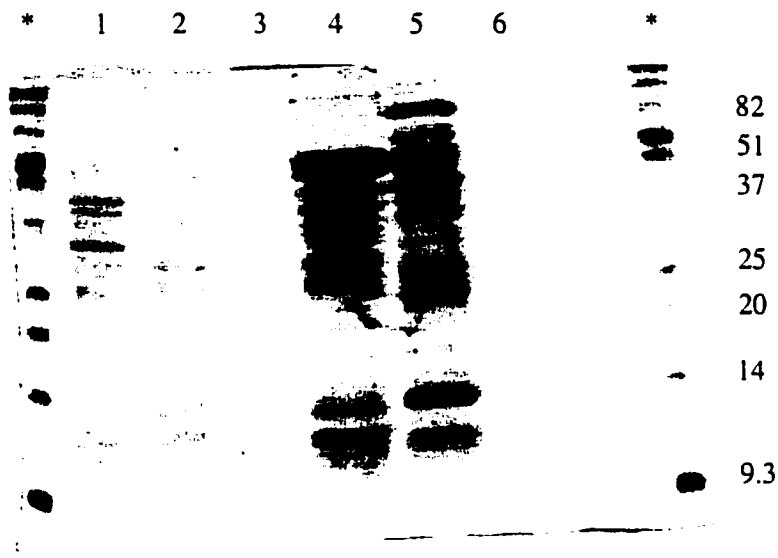


Figure 3.12: SDS-PAGE analysis of purification attempt incorporating the use of the Phenyl Sepharose bulk hydrophobic interaction chromatography resin.

Lane 1 Cell-free extract
Lane 2 Flow through from DEAE Sephacel column
Lane 3 Proteins precipitating at 35% saturated ammonium sulfate
Lane 5 Proteins precipitating between 35% and 65% saturated ammonium sulfate
Lane 5 20% glycerol eluate from the Phenyl Sepharose column
Lane 6 Active fraction after CM Sephadex chromatography

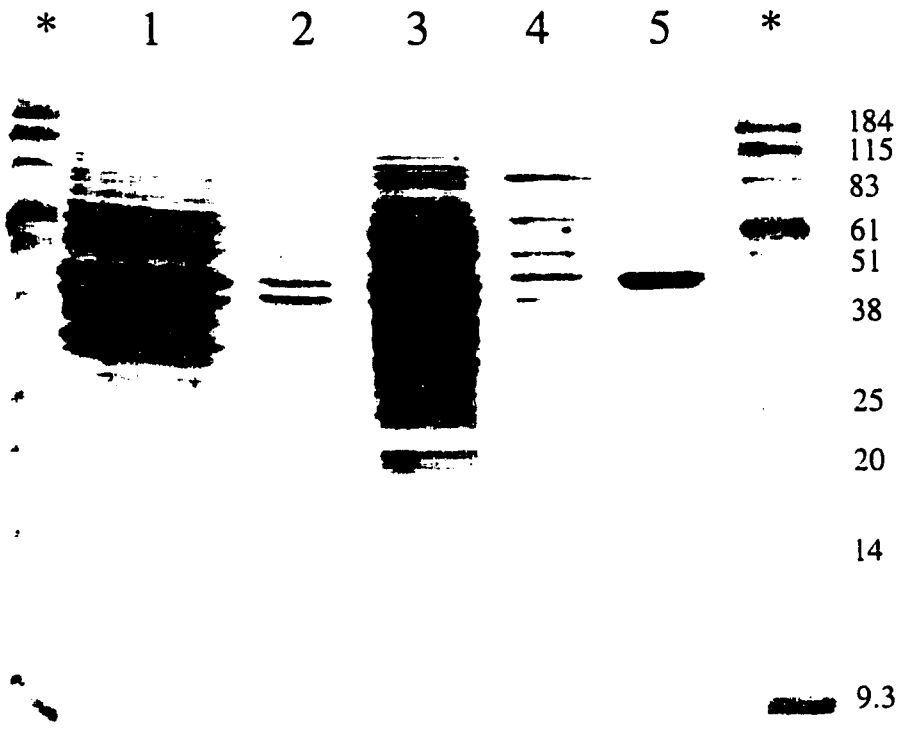


Figure 3.13: SDS-PAGE analysis of purification employing SP Sepharose. The band at around 50 kDa in lane 5 was transferred to a PVDF membrane, excised, and submitted for sequencing.

Lane1 Cell-free extract
Lane2 Flow through from DEAE Sephacel column
Lane3 Proteins precipitating at 65% saturated ammonium sulfate
Lane4 20% glycerol wash from Phenyl Sepharose column
Lane5 Active fraction eluting from the SP Sepharose column

Cycle	Run #1	Run #2	Comments
1	G	G	
2	G	G	
3	L	L	
4	G	G	
5	A	A	
6	-	-	blank, usually implies cysteine
7	L	L	
8	-	-	also probably cysteine
9	Q	Q	
10	A	A	
11	E	E	
12	P	P	
13	V	V	
14	L	L	
15	-	-	peaks are beginning to get weak
16	G	G	weak signals
17	S	A	
18	-	P	
19	A	A	
20	-	G	
21	V	V	
22	V	V	
23	T	T	
24	E	E	run stopped

Figure 3.14: N-terminal protein sequence from partially purified protein shown in figure 3.13. Blank cycles are generally due to a cysteine residue, which only appears if the sample is reduced and alkylated prior to analysis. Good sequence data was obtained for the first fourteen cycles, with weaker data for another ten.

Cycle	T-28	Comment	T-38	Comment
1	G		A	
2	T		D	slightly weak
3	S		G	
4	F		-	likely cysteine
5	V		V	
6	R		G	
7	A		A	
8	G		P	very weak
9	G	slightly weak		
10	G	very weak		
11	L			
12	T			
13	-	possibly cysteine		
14	R	very weak		

Figure 3.15: Results of internal peptide sequencing of protease generated fragments T-28 and T-38.

Forward

L	C	Q	A	E	P	V	L
CTS	TGS	CAG	GCS	GAG	CCS	GTS	CTS

Reverse T-28

F	V	R	A	G	G	G	L
AAG	CTS	GCS	CGS	CCG	CCG	CCG	GAS
SAG	SCC	SCC	SCC	SGC	SCG	SAC	GAA

Reverse T-38

A	D	G	C	V	G	A	P
CGS	CTS	CCG	ACS	CAS	CCG	CGS	GGS
SGG	SGC	GCC	SAC	SCA	GCC	STC	SGC

Figure 3.16: Primers designed based on the protein sequence (figures 3.14 and 3.15) obtained by N-terminal and internal sequencing done on the protein band observed in figure 3.13. The two DNA sequences given for each of the reverse primers are the coding sequence, followed by the primer sequence which is the reverse complement of the coding sequence. S represents either G or C in the primer sequences.

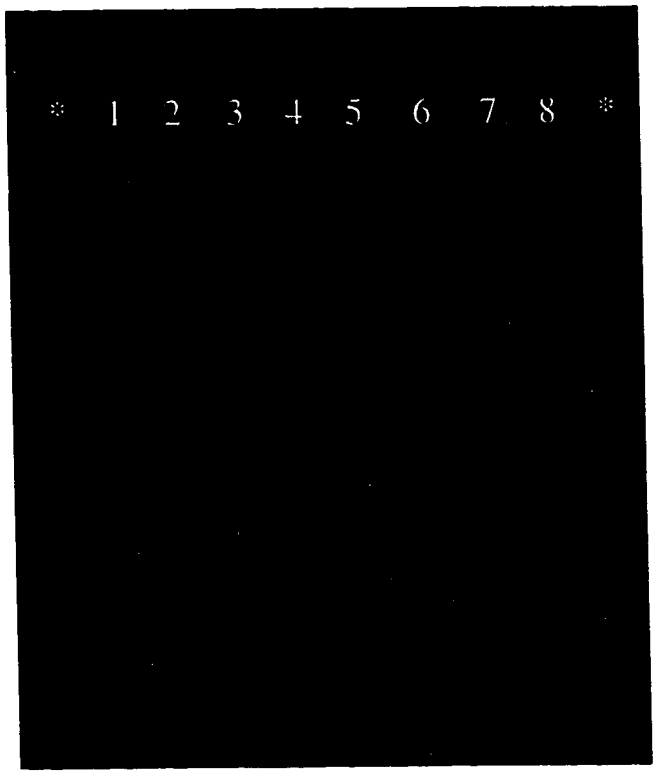


Figure 3.17: PCR amplification using primers from figure 3.16 to amplify fragments from *S. laurentii* genomic DNA.

Lane	Polymerase	Reverse Primer
1	*Pfu*	T-28
2	*Pfu*	T-38
3	*Pfu* Turbo	T-28
4	*Pfu* Turbo	T-38
5	*Taq2000*	T-28
6	*Taq2000*	T-38
7	*Taq2000* + 1 mM $MgCl_2$	T-28
8	*Taq2000* + 1 mM $MgCl_2$	T-38

```
                                                                  Score    E
Sequences producing significant alignments:                       (bits) Value

pdb|1B65|A  Chain A, Structure Of L-Aminopeptidase...              47   5e-05
dbj|BAA29147.1|  (AP000001) 361aa long hypothetica...              34   0.39
emb|CAB49008.1|  (AJ248283) D-AMINOPEPTIDASE. [Pyr...              34   0.39

 pdb|1B65|A Chain A, Structure Of L-Aminopeptidase

           D-Ala-EsteraseAMIDASE FROM
           Ochrobactrum Anthropi Score = 47.3 bits (110), Expect = 5e-05
 Identities = 32/102 (31%), Positives = 44/102 (42%), Gaps = 11/102 (10%)

Query: 6   PVLLTNTYAVGPVHRGVAEWVRGHRADMAPE----WLLPVVTET-DGHLNDIHGPAVRPW 60
           PV++TNT+ +G  H     W+    A          W++PVV ET DG LNDI+G  V
Sbjct: 102 PVVITNTHGIGMAHHATVRWMVDRYASTYQTDDFLWIMPVVAETYDGALNDINGFPVTEA 161

Query: 61  QAEPVLG------VEQGLCAQGHQLALKTGGIGPGAGERAVQ 96
             L       V++G C  G  +       G G    R V+
Sbjct: 162 DVRKALDNVASGPVQEGNCGGGTGMITYGFKGGTGTASRVVE 203


 dbj|BAA29147.1| (AP000001) 361aa long hypothetical D-aminopeptidase

           [Pyrococcus horikoshii]
           Length = 361

 Score = 34.4 bits (77), Expect = 0.39
 Identities = 31/119 (26%), Positives = 55/119 (46%), Gaps = 15/119 (12%)

Query: 6   PVLLTNTYAVGPVHRGVAEWVRGHRADMAPEW--LLPVVTE-TDGHLNDIHGPAV-RPWQ 61
           P++LTNT ++G    G+ +++      D+       + P+V E  D +LNDI G  V R
Sbjct: 95  PIILTNTLSIGTAVEGLLDYILEENEDIGVTTGSVNPLVLECNDSYLNDIRGRHVKREHV 154

Query: 62  AEPVLGVEQGLCAQGHQLALKTGGIGPGAGERAVQPFRGVLRPDRGVVGEVTGQLVLLG 120
            E +    ++            + G +G G G   A + F+G +      +V E+ G+    +G
Sbjct: 155 VEAIKRADED---------FEEGAVGAGTGMSAFE-FKGGIGSASRIV-EIEGKKYTVG 202


 emb|CAB49008.1| (AJ248283) D-AMINOPEPTIDASE. [Pyrococcus abyssi]
           Length = 361

 Score = 34.4 bits (77), Expect = 0.39
 Identities = 31/119 (26%), Positives = 55/119 (46%), Gaps = 15/119 (12%)

Query: 6   PVLLTNTYAVGPVHRGVAEWVRGHRADMAPEW--LLPVVTE-TDGHLNDIHGPAV-RPWQ 61
           P++LTNT ++G    G+ ++V      D+       + P+V E  D +LNDI G  V R
Sbjct: 95  PIVLTNTLSIGTAVDGLLDYVLSENEDIGVTTGSVNPLVLECNDSYLNDIRGRHVKREHV 154

Query: 62  AEPVLGVEQGLCAQGHQLALKTGGIGPGAGERAVQPFRGVLRPDRGVVGEVTGQLVLLG 120
            E +    ++            + G +G G G   A + F+G +      +V E+ G+    +G
Sbjct: 155 VEAIKNAKED---------FEEGAVGSGTGMSAFE-FKGGIGSSSRIV-KIEGRSYTIG 202
```

Figure 3.18: Homologies found to the PCR product amplified from *S. laurentii* genomic DNA using the T-38 reverse primer. The strongest homologies to known enzymes are to aminopeptidases. No homologies to acyl activating proteins were observed. The query sequence is that which was submitted to the database, and the subject sequence is the database sequence that it is similar to. The score is an indication of the similarity of the two sequences. The expect value is a statistical number representing the likelihood that the homology could be found by random chance.

```
                                                                  Score     E
Sequences producing significant alignments:                       (bits)  Value

emb|CAB12196|  (Z99106) similar to transcriptional re...           55   1e-07
pir||A29606  hypothetical protein J12 (mmr 5' region)...           46   5e-05
gi|2983797  (AE000737) transcriptional regulator (Ars...           38   0.011

 emb|CAB12196| (Z99106) similar to transcriptional regulator (ArsR family)
          [Bacillus subtilis]
          Length = 104

 Score = 54.7 bits (129), Expect = 1e-07
 Identities = 29/66 (43%), Positives = 43/66 (64%), Gaps = 1/66 (1%)

Query: 19 HHPDVDEIALPRVLFALSEPLRLTMVRALSERGE-VDSLDLGPDLPRSTLTHHTSLLRES 77
          +HP+ + + L +VL ALS+PLRL +V+ L+E  E   S      + +STL+HH  +LRES
Sbjct: 5  NHPETETLQLTKVLHALSDPLRLELVKQLAEAKEKTCSTCADVQVAKSTLSHHFKVLRES 64

Query: 78 GVVFVR 83
          G+  VR
Sbjct: 65 GIAQVR 70


 pir||A29606 hypothetical protein J12 (mmr 5' region) - Streptomyces
          coelicolor plasmid SCP1 gi|1196874 (M18263) unknown
          protein [Plasmid SCP1] gi|225819|prf||1314204A gene
          mmr ORF [Streptomyces sp.]
          Length = 111

 Score = 46.1 bits (107), Expect = 5e-05
 Identities = 27/68 (39%), Positives = 41/68 (59%), Gaps = 5/68 (7%)

Query: 16 RTYHHPDVDEIALPRVLFALSEPLRLTMVRALSERGE---VDSLDLGPDLPRSTLTHHTS 72
          R   HPD D I L  VL AL +P+R ++VR L++  E      + D+   + RST THH
Sbjct: 9  RITDHPDADAITLQGVLDALVDPVRRSIVRQLAKAPEDIACGTFDI--TVSRSTGTHHFK 66

Query: 73 LLRESGVV 80
          +LR++G++
Sbjct: 67 VLRQAGII 74


 gi|2983797 (AE000737) transcriptional regulator (ArsR family) [Aquifex
          aeolicus]
          Length = 102

 Score = 38.3 bits (87), Expect = 0.011
 Identities = 23/55 (41%), Positives = 32/55 (57%), Gaps = 2/55 (3%)

Query: 28 LPRVLFALSEPLRLTMVRALSERGEVDSLDLGPDLPRS--TLTHHTSLLRESGVV 80
          L R+ +ALSEP RL MV+ L ER E+   D       S   ++ H  +LRE+G+V
Sbjct: 6  LARIFYALSEPKRLCMVKLLLEREELCVCDFMRIFKESQPKISFHLKVLREAGLV 60
```

Figure 3.19: Homologies found to the PCR product amplified from *S. laurentii* genomic DNA using the T-28 reverse primer. The strongest homologies to known proteins were to transcriptional regulating enzymes. No homologies to acyl activating proteins were observed.

3.4.3 PCR approach

Marahiel and coworkers developed a novel method for cloning nonribosomal peptide synthetase encoding genes [73] in which polymerase chain reaction (PCR) primers are designed based on the conserved domain regions found in this family of enzymes (figure 1.10). In the earliest examples, cores A2 and A3 were used to design primers. While this approach was successful, it was found that in order to confirm that the resulting PCR products were indeed fragments of nonribosomal peptide synthetases, it was necessary to first clone the gene from the genome, then sequence the region flanking that which had been amplified to find other cores. Further studies involved designing primers based on cores A7 and T. The PCR product resulting from this experiment contains three cores not encoded by the primers, and allows confirmation of correct clones by direct sequencing of the PCR product.

The high percentage of GC pairs in Streptomcyete DNA precluded direct usage of the primers designed by Marahiel. New primers, based on both the conserved amino acid sequences observed for cores A7 and T and the strong GC preference of Streptomycete DNA, were designed (figure 3.20). This choice of primers should yield a product five hundred base pairs in length containing cores A8, A9, and A10. When these fragments are end-sequenced, core A8 should be most readily apparent, due both to its size and its proximity to core A7. Sequencing technology at this time was such that achieving a read sufficiently long to see homology to core A9 would be unlikely, but observation of core A10 was a possiblity.

The first PCR reaction undertaken used mechanically sheared genomic DNA from *S. azureus*, *S. actuosus*, and *S. laurentii* as templates, and *Taq* polymerase to catalyze

Core A7 alignments		Core T alignments	
SrfA1	germy rtgdlvkw.t	SrfA1	sevgvtdnff slggdsikgi
grsa_bacbr	GEKLY KTGDQARWLS	grsa_bacbr	EKIGIKDNFY ALGGDSIKAI
Bac2	gervy rtgdlarwls	Bac2	eqigindhff digghslkaf
bacC	gerly ktgdlakwlp	bacC	eksgindhff emgghslkaa
FenE	germy ktgdlarwlp	FenE	gpvgihdnff drgghslkat
SrfA2	geriy rtgdlarwlp	SrfA2	vkagvtdnff migghslkam
grsb_bacbr	NERMY RTGDLARWLP	grsb_bacbr	SQIGIQDNFF SLGGHSLKAI
virg	gtrmy rtgdlarwag	virg	skvgihedff dlgghsllat
Consensus	GERMY RTGDLARWLP	Consensus	-K-GI-DNFF DLGGHSLKAI
Primer	RTGDLARW	Primer	NFF PLGGH

Primer Design:

Core A7 forward primer:

R	T	G	D	L	A	R	W
CGS	ACS	GGS	GAC	CTS	GCS	CGS	TGG

Core T reverse primer:

N	F	F	P	L	G	G	H
AAC	TTC	TTC	CCS	CTS	GGS	GGS	CAC
GTG	SCC	SCC	SAG	SGG	GAA	GAA	GTT

Figure 3.20: Designing the first set of PCR primers. Alignments of core A7 and core T of some representative nonribosomal peptide synthetases are shown above with their consensus, and the amino acid sequence upon which the primers were based. The primer sequence is shown for the forward primer, and the coding sequence as well as the primer sequence is shown for the reverse primer. S represents either G or C in the primer sequences.

Gene	**Antibiotic**	**Gene**	**Antibiotic**
SrfA1	Surfactin	grsb_bacbr	Gramicidin S
SrfA2	Surfactin	FenE	Fengycin
grsa_bacbr	Gramicidin S	virg	Virginiamycin

the reaction. In an experiment with fairly high stringency, discreet band formation was observed in all three cases. The expected 500 bp products were observed in each experiment, with *S. actuosus* showing an additional 450 bp fragment, and *S. azureus* showing an additional 750 bp fragment. The products formed using *S. laurentii* were fainter than the other two, but were clearly observable (figure 3.21). The larger than expected PCR product could be explained by the inclusion of another activity in the adenylation domain. This has been observed in several nonribosomal peptide synthetase systems, including racemase activity in the first enzyme in the gramicidin S pathway (figure 1.11), and N-methyl transferase in the cyclosporin biosynthetic pathway [28].

Optimization of the PCR reaction using *S. laurentii* DNA as the template was done by varying the concentration of DMSO, $MgCl_2$, and the annealing temperature until clear band formation was observed. At this stringency both a 450 bp band and a 500 bp band was observed. The 500 bp band was excised from the gel, and the DNA purified (figure 3.22). This purified DNA was used as a template to re-amplify the fragment, yielding enough DNA to clone the fragment into a plasmid vector for sequencing. It was found that re-amplification proceeded most cleanly using *Pfu* polymerase, a DNA polymerase with exonuclease proofreading activity that generates truly blunt ended PCR products (figure 3.23).

The PCR products from the *Pfu* polymerase amplification were purified. Initial attempts to clone these blunt ended fragments into the *Eco*RV cloning site of the pBluescript KS^- plasmid vector proved unsuccessful. Use of the pCRScript KS^- blunt end cloning vector, however, readily yielded vector containing insert in the

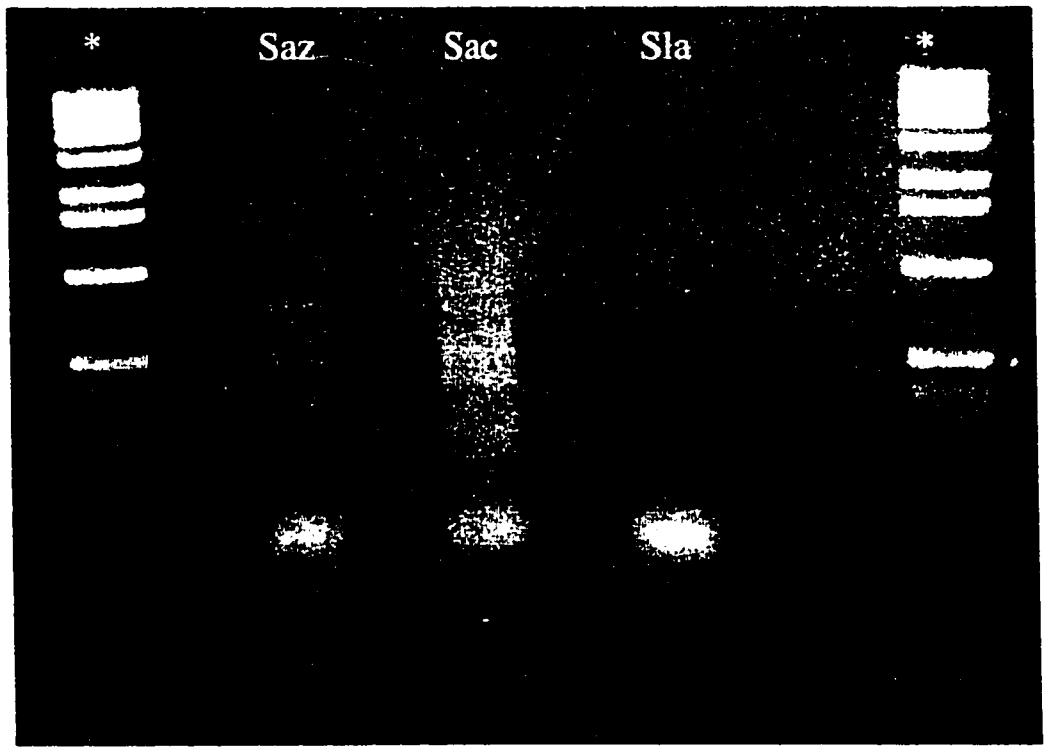
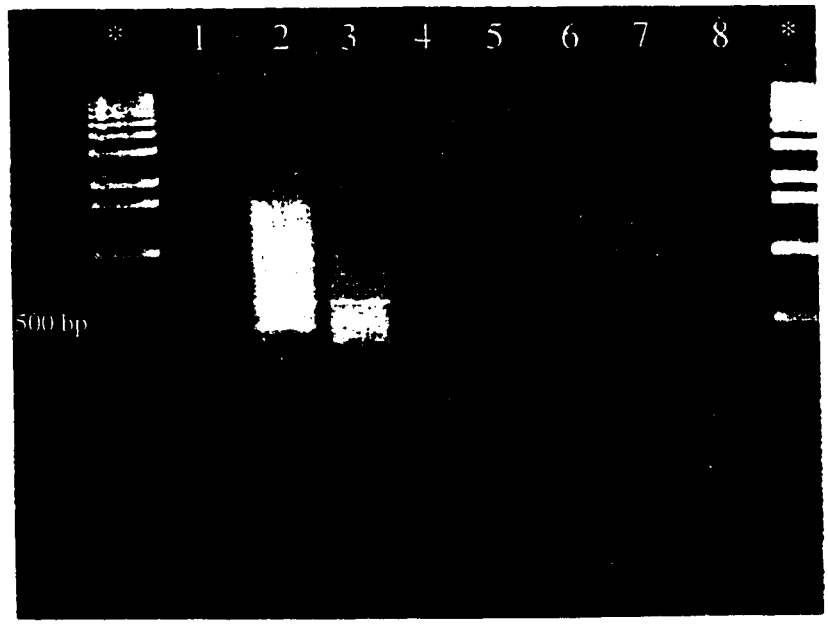


Figure 3.21: Initial PCR reaction, using mechanically sheared *S. azureus*, *S. actuosus*, and *S. laurentii* as templates. Band formation is clearly evident for both *S. azureus* and *S. actuosus*, with the expected 500 bp band as well as one other band each. The band amplified from *S. laurentii* is fainter, but has the expected size.



PCR Conditions

Lane	$MgCl_2$ concentration	DMSO concentration	Enzyme
1	0.15 mM	0%	Taq Polymerase
2	0.15 mM	5%	Taq Polymerase
3	0.65 mM	5%	Taq Polymerase
4	1.15 mM	5%	Taq Polymerase
5	1.65 mM	5%	Taq Polymerase
6	2.15 mM	5%	Taq Polymerase
7	2.65 mM	5%	Taq Polymerase
8	3.15 mM	5%	Taq Polymerase

Figure 3.22: Optimization of PCR conditions in *S. laurentii*. The annealing temperature was 60°C, and magnesium and DMSO concentrations were varied in order to gain higher selectivity. The lanes marked with an asterix are the DNA size standard, with the 500 bp band labeled. Lane 1 was chosen as the most selective, and the 500 bp band excised for reamplification.

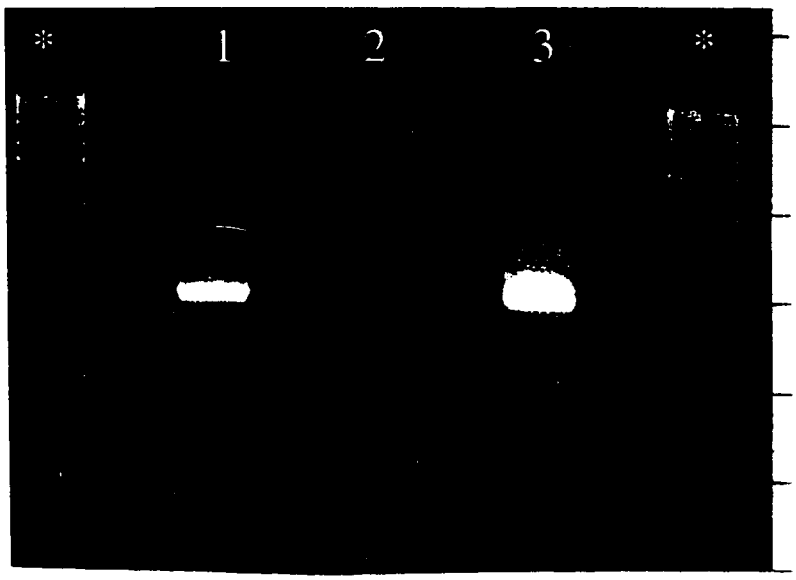


Figure 3.23: Reamplification of the 500 bp PCR product from figure 3.22 using three different polymerases. The *Taq* polymerase supplied from Gibco, Inc. (lane 2) did not show any amplification, probably due to age of the enzyme. The Boeringher-Mannheim *Taq* polymerase (lane 3) showed strong, but somewhat non-specific, re-amplification. Stratagene's *Pfu* polymerase (lane 1) showed fairly clean re-amplification with high yield

majority of colonies examined. The pCRScript KS$^-$ blunt end cloning system utilizes a pBluescript KS$^-$ based vector that has had a *Srf*I endonuclease restriction site engineered into the multi cloning site of the vector. *Srf*I is an eight base cutting restriction endonuclease which generates blunt ends. When the ligation reaction is performed, this enzyme is added to the reaction mixture. If the vector religates without an insert, the endonuclease that is present re-digests the vector. When an insert is ligated to one arm of the vector, the restriction site is destroyed, allowing the construct to remain untouched. This allows for a much higher percentage of colonies containing the plasmid ligated to an insert. Digestion with the restriction

endonuclease *Pvu*II excises a fragment of the plasmid vector, containing the insert and 450 bp of vector DNA, allowing for rapid screening for positive clones.

Four clones containing appropriately sized inserts were selected for sequencing. Three of the four clones showed sequence homology to known nonribosomal peptide synthetases when translated into protein sequence and compared to databases using the BLAST algorithm [1]. Initial examination of the sequences showed that two of the sequences, samples Q and 1H, were identical, while sample X was unique. All samples showed homology to core A8, while sample X also showed homology to core A10 (figure 3.24). Examination at a later date with more sophisticated alignment tools showed that sample X is identical to samples Q and 1H, but the sequencing gel had a region of compression, distorting the data.

A cosmid library of *S. laurentii* genomic DNA had been constructed in the cosmid vector pOJ446 (section 2.4, page 36). This vector contains both the pUC *E. coli* origin of replication, and the SCP2* Streptomycete origin of replication, as well as the OriT gene, which allows conjugal transfer of this cosmid between *E. coli* and Streptomycete bacteria. This library was screened by colony hybridization with randomly primed, [α^{32}P]dCTP labeled fragments synthesized using the PCR product 1H as a template. Twelve cosmid clones showed strong hybridization to this probe, and four of these were selected for further characterization. Southern hybridization to these five clones after digestion with the restriction endonuclease *Bam*HI yielded strong hybridization to a 23 kb fragment present, at least in part, in all of the cosmids. pOJ446 is known for causing rearrangements in insert DNA, so to simplify mapping of the region flanking the hybridizing fragment, a library which had been constructed

Figure 3.24: Alignment of the sequence of the three PCR products with sequences in the protein database. Alignments were obtained using the BLASTX program [1], which compares a six frame translation of a nucleotide sequence to the major protein databases. The homology to core A7 is due to the forward primer sequence, and half of the homology to the T core is from the reverse primer sequence.

Sequences producing significant alignments:	Score bits)	E Value
gi|1075778|pir||JX0340 gramicidin S synthase 2 - Bac...	60	3e-09
gi|2522214 (AF023465) fengycin synthetase FenE [Baci...	59	6e-09
gi|2052277|emb|CAA72310| (Y11547) Virginiamycin S sy...	57	1e-08
gi|2501965 (AF021262) daptomycin biosynthetic protei...	57	2e-08
gi|2623773 (AF004835) tyrocidine synthetase 3 [Brevi...	56	3e-08
gi|1771242|emb|CAA65394| (X96558) lysobactin synthet...	56	3e-08
gi|2982196|gb|AAC06348| (AF007865) bacitracin synthe...	56	3e-08

gi|1075778|pir||JX0340 gramicidin S synthase 2 -

Bacillus brevis gi|511490|dbj|BAA06146|
(D29676) gramicidin S synthetase 2 [Brevibacillus brevis]
Length = 4450

Score = 59.7 bits (142), Expect = 3e-09
Identities = 28/39 (71%), Positives = 35/39 (88%), Gaps = 1/39 (2%)
Frame = +1

```
Query: 13   GDLARWTEAGVLEFVGRGD-QVKVRGYRIELGEVESALLQ 129
            GDLA+W   G++E+VGR D QVKVRGYRIELGE+ESA+L+
Sbjct: 3963 GDLAKWRSDGMIEYVGRVDEQVKVRGYRIELGEIESAILE 4002
```

Score = 54.7 bits (129), Expect = 9e-08
Identities = 26/38 (68%), Positives = 32/38 (83%), Gaps = 1/38 (2%)
Frame = +1

gi|2522214 (AF023465) fengycin synthetase FenE [Bacillus subtilis]
Length = 2554

Score = 58.6 bits (139), Expect = 6e-09
Identities = 27/37 (72%), Positives = 32/37 (85%), Gaps = 1/37 (2%)
Frame = +1

```
Query: 13   GDLARWTEAGVLEFVGR-GDQVKVRGYRIELGEVESAL 123
            GDLARW   G +E+VGR  DQVK+RGYR+ELGE+ESAL
Sbjct: 1878 GDLARWLPDGTIEYVGRVDDQVKIRGYRVELGEIESAL 1915
```

gi|2052277|emb|CAA72310| (Y11547) Virginiamycin S synthetase

[Streptomyces virginiae]
Length = 1997

Score = 57.4 bits (136), Expect = 1e-08
Identities = 29/37 (78%), Positives = 31/37 (83%), Gaps = 1/37 (2%)
Frame = +1

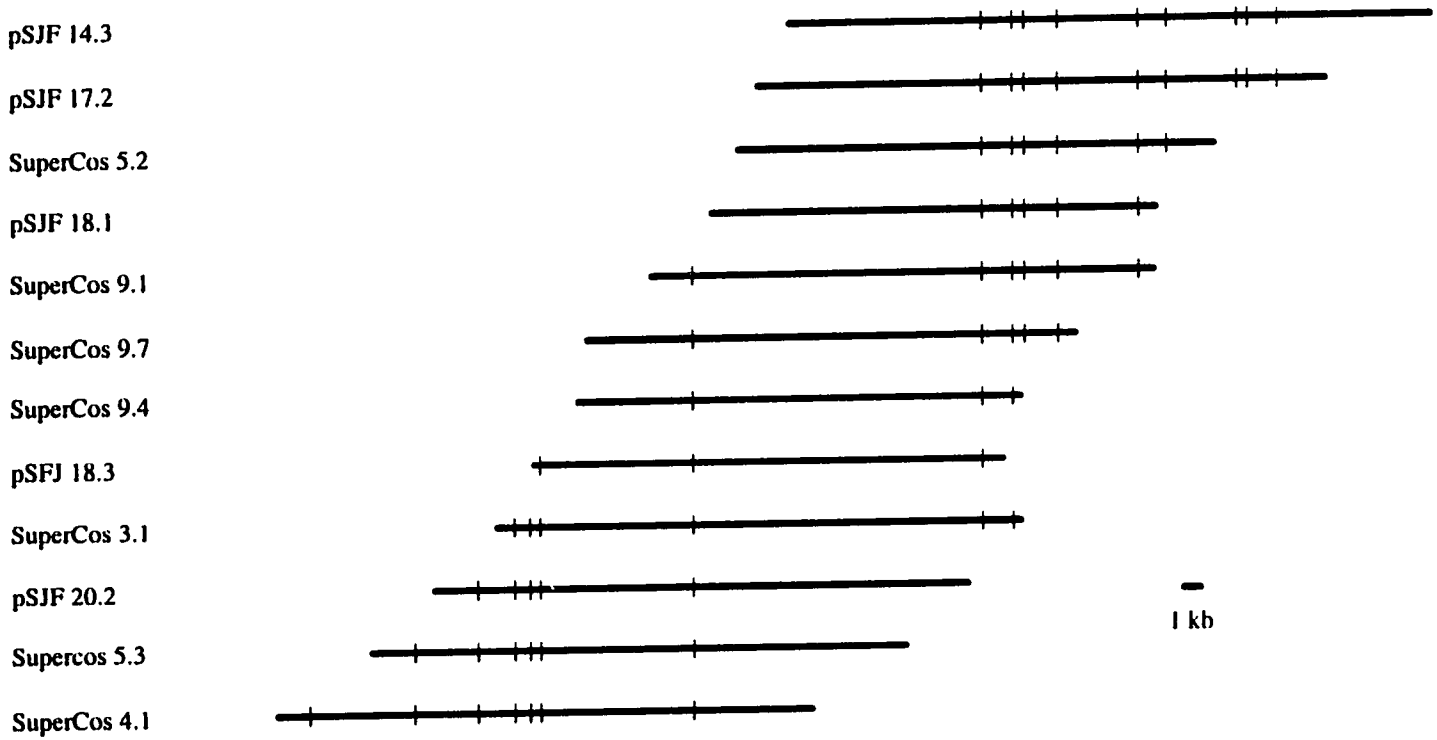
```
Query: 13  GDLARWTEAGVLEFVGRGD-QVKVRGYRIELGEVESAL 123
           GDLARW  +GVLEF GR D QVKVRG+RIE GEVES L
Sbjct: 440 GDLARWAGSGVLEFAGRADHQVKVRGFRIEPGEVESVL 477
```

in the SuperCos cosmid vector was screened as well. Of fifty strongly hybridizing colonies eight were selected for analysis. Both restriction endonuclease digestion and Southern hybridization gave the same results as were seen with the clones isolated from the pOJ446 library. These results indicated that this is the only region on the *S. laurentii* genome containing DNA strongly homologous to the probe.

The *Bam*HI restriction endonuclease digestion sites were mapped on these overlapping cosmids. This was done by matching of single and double restriction endonuclease digestion patterns, hybridization of the clones with the 1H Southern probe, and by synthesizing Southern probes from four of the *Bam*HI restriction endonuclease generated fragments, sized 12, 6.5, 2.5, and 5.5 kb, and examining their hybridization to restriction endonuclease digested cosmids. The resulting map (figure 3.25) encompasses a total of approximately 70 kb of DNA surrounding the 23 kb fragment which hybridizes to the 1H probe.

The *Bam*HI fragments resulting from the digestion of these cosmids were subcloned into the pBluescript KS^- plasmid vector (figure 3.26). The clone containing the 23 kb fragment which hybridized strongly with the 1H probe, now designated pSJF114, was selected for further analysis. This clone was digested with the restriction endonucleases *Not*I, *Kpn*I, and *Sac*I, and mapped (figure 3.27). The products of digestion by the restriction endonuclease *Not*I were subcloned into the $pBCKS^-$ plasmid vector and the resulting constructs subjected to sequence analysis. Sequencing the ends of these fragments was accomplished using the method of Shine and Delgarno [59], yielding a clear read of one to three hundred bases per reaction. Analysis of the sequence data was performed by translating the sequence into the six possi-

Figure 3.25: Cosmid map of the region containing the fragment that shows strong hybridization to the 1H probe. Vertical lines represent cleavage sites for the restriction endonuclease *Bam*HI. Cosmids designated pSJF were isolated from the pOJ446 library, and those designated Supercos originate from the SuperCos library. Approximately 70 kb of DNA were mapped. The sizes of the *Bam*HI generated fragments are given in figure 3.26


pSJF 14.3
pSJF 17.2
SuperCos 5.2
pSJF 18.1
SuperCos 9.1
SuperCos 9.7
SuperCos 9.4
pSFJ 18.3
SuperCos 3.1
pSJF 20.2
Supercos 5.3
SuperCos 4.1
1 kb

ble protein sequences that it could encode and comparing these protein fragments to the protein databases using the BLAST [1] sequence comparison program. Three subclones, pSJF202, pSJF203, and pSJF204, showed significant homology to nonribosomal peptide synthetases, but none of the highly conserved core sequences were found. Southern hybridization using the 1H probe showed strong hybridization to pSJF203 and pSJF204, and noticeable hybridization to pSJF202. With these results in hand it was decided to take a more detailed look at pJSF203 and pSJF204. Internal deletion and subcloning of restriction endonuclease digestion products from these plasmids yielded several smaller constructs (figure 3.28), which were end sequenced as before. Homologies to nonribosomal peptide synthetase conserved cores were observed in both cases, and found to have the same orientation. Further sequence analysis of pSJF205 and pSJF206 also showed regions homologous to nonribosomal peptide synthetase core regions (figure 3.29). When the positions of these core regions are examined on the map of pSJF114 it appears that this 23 kb fragment contains at least five peptide synthetase modules, indicating that it may be a part of the thiostrepton biosynthetic cluster. Another interesting homology found on pSJF211 was a strong similarity to the NshC gene. The NshC gene is proximal to the nosiheptide (figure 1.2) resistance gene in *S. actuosus* [30] and was originally thought to be involved in nosiheptide biosynthesis. NshC shows homology to 4'-phosphopantetheine transferases, a gene required for the production of peptides by the thiotemplate mechanism [29]. Later studies have shown by deletion of this gene from the genome that NshC is not necessary for the production of antibiotic [68]. This homology further indicates that this region is involved in thiopeptide antibiotic biosynthesis. While gene

homology can give indications that this cluster is responsible for the production of thiostrepton, conclusive proof requires mutating a gene in the cluster on the genome, and demonstrating that this mutant does not produce the antibiotic.

The *Not*I fragments obtained from the digestion of pSJF114 were ligated into pKC1132, the insertional mutation plasmid vector discussed in section 2.4. These plasmids were then used to transform *E. coli* cells, and transferred by conjugation to *S. laurentii*. One colony from each exconjugant was isolated, grown in production conditions, and screened by HPLC for the presence of thiostrepton by the same method used to analyze mutations in the region surrounding the thiostrepton resistance gene (section 3.2, page 47). Each of these were shown to produce thiostrepton at or near wild type levels. Since pSJF203 has been shown to contain a peptide synthetase domain, twelve exconjugants were picked and grown under production conditions. Of these twelve, three were found to not produce thiostrepton. The production levels of thiostrepton can vary significantly from flask to flask, so to prove that these are truly non-producing genotypes the cultures were grown again. These second cultures of the three initial non-producers each were found to produce thiostrepton at or near wild type levels, indicating a producing phenotype.

One possible explanation for this result could be that the plasmid DNA did not integrate at the homologous section of genomic DNA, or that the wild type strain was able to generate an apramycin resistant mutant. In order to prove that the insertion took place, and that it occured at the expected location, genomic DNA was isolated from the mutant strains and hybridized against a probe synthesized from pKC1132. When the genomic DNAs of the mutant strains were digested with

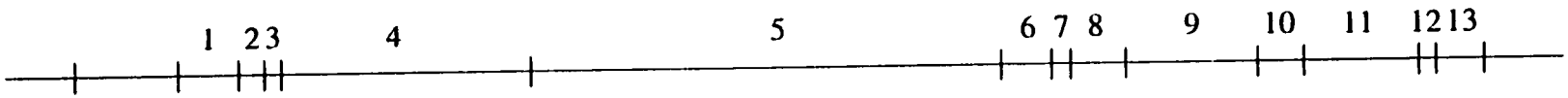
the restriction endonuclease *Not*I a 3.2 kb hybridizing fragment, the cleanly excised pKC1132 vector, was observed. In order to determine that the insertion occurred in the correct location the genomic DNA was digested with the restriction endonuclease *Kpn*I, which does not cut the pKC1132 vector, but has two known recognition sites on pSJF114 (figure 3.27). The expected size for this fragment is 6.4 kb, 3.2 kb from pKC1132, and another 3.2 kb from the genome. A 6.5 kb hybridizing fragment was observed, indicating that the integration occurred at the expected location (figure 3.30).

This result strongly indicates that these peptide synthetase modules are not necessary for the biosynthesis of thiostrepton. The possiblity does exist, however, that this is a now inactive part of the cluster. If one examines the structure of nosiheptide (figure 1.2 and 1.6) it appears that it may be produced by a partially inactivated and slightly modified thiostrepton biosynthetic cluster. Similarly, thiostrepton may be produced by a partially inactivated cluster encoding a larger antibiotic. With this thought in mind the region surrounding pSJF114 was examined by end sequencing. Sequencing technology by this time had advanced to the use of dye terminated reactions, and much longer and more reliable sequence reads. The ends of the cloned *Bam*HI digestion products (figure 3.26) were sequenced and compared to the protein databases.

Sequence data from pSJF111, on one of the ends of the mapped region, show strong homology to many geranylgeranyl pyrophosphate synthetases and farnesyl pyrophosphate synthetases. It is extremely unlikely that the biosynthesis of thiostrepton involves any terpenoid intermediates, so one can safely assume that this is outside

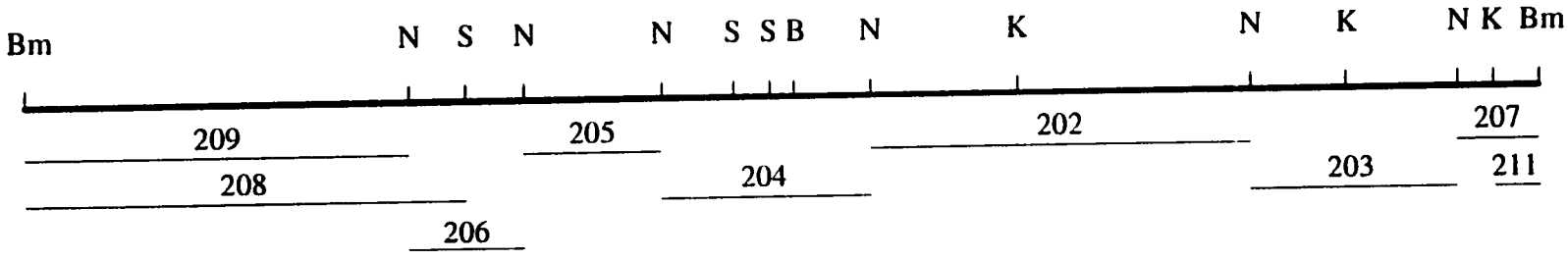
the cluster. Sequence data from pSJF121, the *Bam*HI fragment adjacent to the 203 end of pSJF114, showed weak homology to peptide synthetases, and in a confirming experiment showed weak hybridization to a peptide synthetase Southern hybridization probe, indicating at least one nonribosomal peptide synthetase module located on the fragment. Southern hybridization against restriction endonuclease digested DNA from pSJF121 gives evidence of at least two modules located here. On the other end of pSJF114, the fragment proven to contain peptide synthetase modules, pSJF118 shows homology to several oxidoreductases. pSJF119 is similar to known enoyl pyruvyl transferases. pSJF105 contains regions homologous to oxidoreductases of several types, and most interesting, pSJF108 contains the beginning of an open reading frame which is highly homologous to several known methyl transferases. A full map of these homologies and the peptide synthetase modules found on pSJF114 is shown in figure 3.31.

While many of these activities could be a part of a putative thiostrepton biosynthetic cluster, the absence of any further peptide synthetase modules is disappointing at best. When taken in context with the insertional mutation experiments the most likely conclusion to be drawn from these results is that this cluster does not encode genes responsible for the biosynthesis of thiostrepton.



#	clone	size(kb)	#	clone	size(kb)
1	pSJF111	3.1	8	pSJF117	2.5
2	pSJF113	0.8	9	pSJF116	6.5
3	pSJF112	1.4	10	pSJF105	2.0
4	pSJF121	12	11	pSJF103	5.5
5	pSJF114	23	12	pSJF108	0.7
6	pSJF118	2.5	13	pSJF107	2.5
7	pSJF119	1.0			

Figure 3.26: Map of the region surrounding the hybridizing fragment. Vertical lines indicate the presence of a recognition site for the restriction endonuclease *Bam*HI. The table gives the size of the *Bam*HI digestion products that were subcloned into pBluescript KS^-. Mapping experiments conducted by Ya-Fen Jiang.



clone	size (kb)	clone	size (kb)	clone	size (kb)
pSJF202	6.0	pSJF205	2.2	pSJF208	7.0
pSJF203	3.2	pSJF206	1.8	pSJF209	6.0
pSJF204	3.1	pSJF207	1.2	pSJF211	0.5

Figure 3.27: Map of pSJF114, the 23 kb fragment which shows strong hybridization to the peptide synthetase probe. The restriction endonuclease digestion sites are indicated as follows: Bm = *Bam*HI, B = *Bgl*II, N = *Not*I, S = *Sac*I. Subcloned fragments are indicated below the map, and described in the table. Mapping experiments conducted by Ya-Fen Jiang

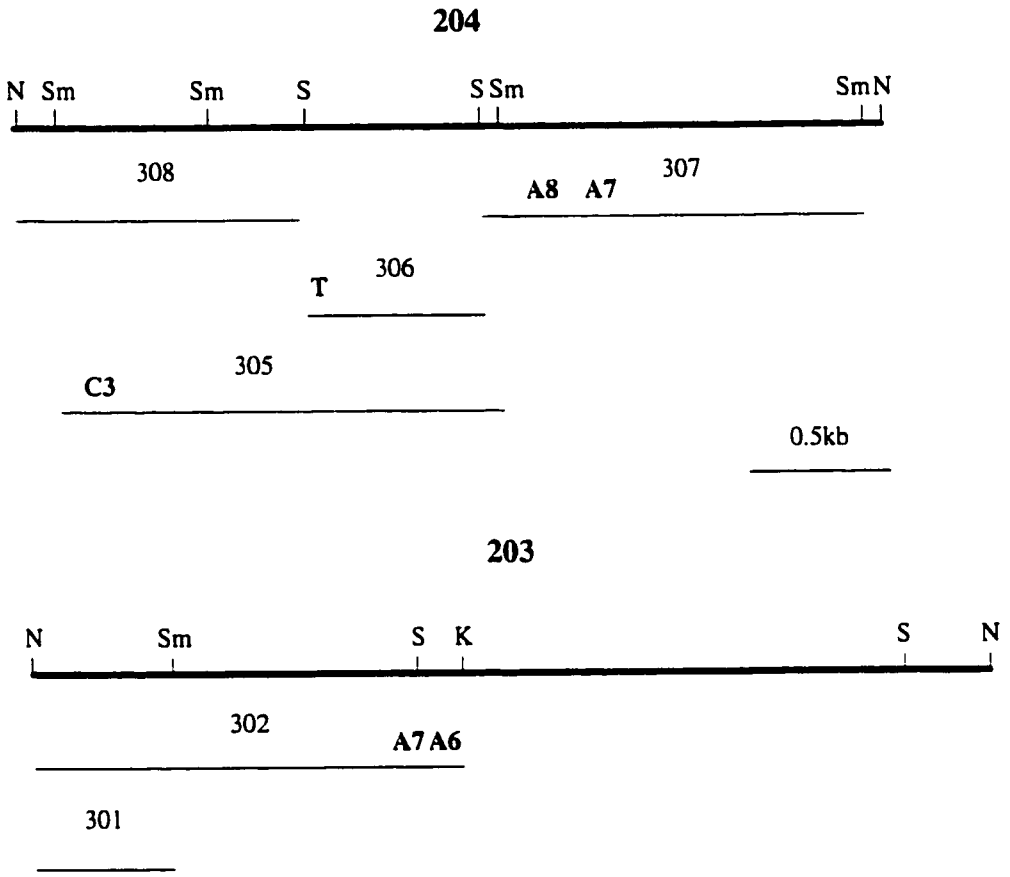


Figure 3.28: Map of subclones derived from pSJF203 and pSJF204. Restriction endonuclease digestion sites are indicated as follows: K = *Kpn*I; N = *Not*I; S = *Sac*II; Sm = *Sma*I. Nonribosomal peptide synthetase conserved core regions found by end sequencing of these subclones are indicated. Mapping experiments conducted by Ya-Fen Jiang

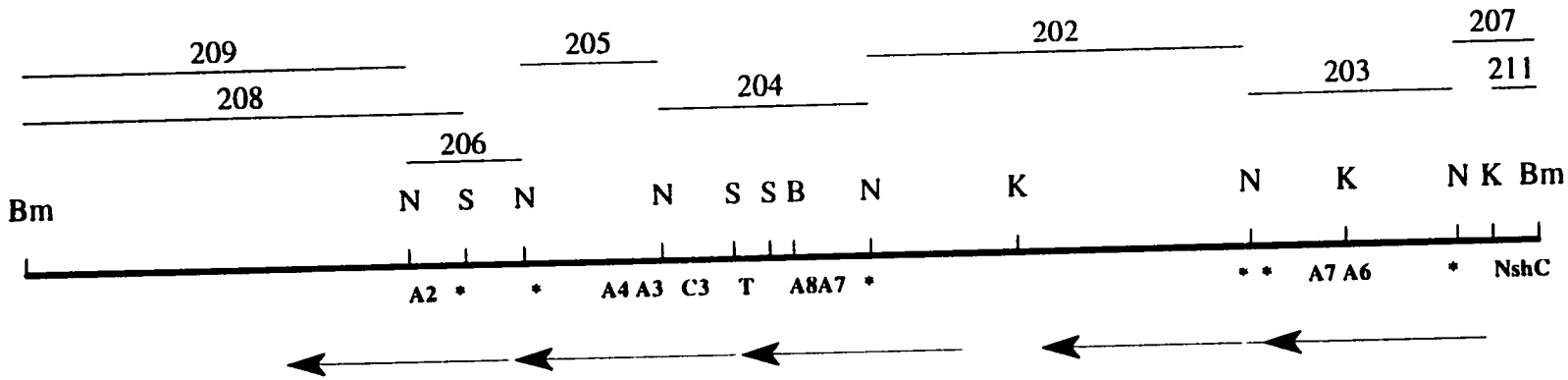
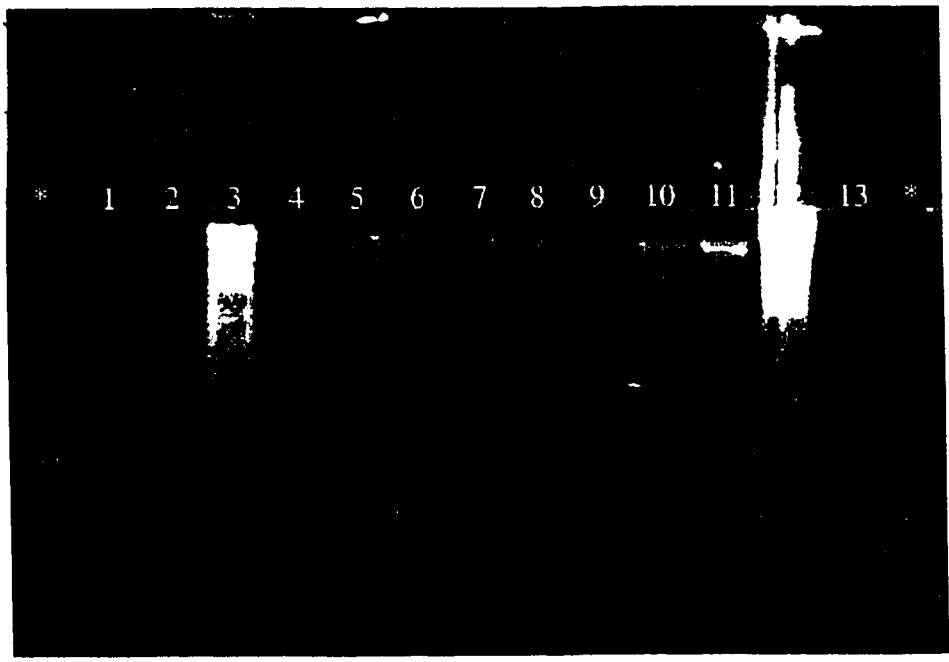


Figure 3.29: Map of pSJF114 showing homologies to nonribosomal peptide synthetases. Regions with high similarity to nonribosomal peptide synthetase core regions are indicated by the name of the core, and regions which showed homology to peptide synthetases, but not to the core regions, are indicated by an asterix. Location of the nonribosomal peptide synthetase domains found by sequencing are indicated by the bold arrows. The region shown to be homologous to the *S. actuosus* NshC gene is labeled as well.

Figure 3.30: Southern hybridization demonstrating integration of the insertional mutation vector pSJF203' (the insert from pSJF203 cloned into pKC1132). Six strains generated by the putative integration of pSJF203' showing resistance to apramycin were grown, and their genomic DNA isolated. The DNA was digested with the restriction endonucleases *Not*I, and *Kpn*I. Wild type genomic DNA, digested with the restriction endonuclease *Bam*HI, was also used to determine that no pKC1132-hybridizing DNA is present in the unmodified genome. These samples were hybridized against a probe synthesized from the vector pKC1132. The *Not*I digested genomic DNA contained a 3.2 kb hybridizing fragment (the excised vector), as expected, and the *Kpn*I digested genomic DNA contained a 6.4 kb hybridizing fragment, which is in agreement with integration at the correct site.

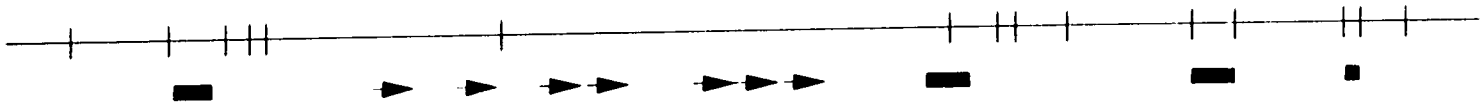
	Source	Enzyme		Source	Enzyme
1	*S. laurentii* pSJF203'-1	*Not*I	**7**	*S. laurentii* pSJF203'-1	*Kpn*I
2	*S. laurentii* pSJF203'-2	*Not*I	**8**	*S. laurentii* pSJF203'-2	*Kpn*I
3	*S. laurentii* pSJF203'-3	*Not*I	**9**	*S. laurentii* pSJF203'-3	*Kpn*I
4	*S. laurentii* pSJF203'-4	*Not*I	**10**	*S. laurentii* pSJF203'-4	*Kpn*I
5	*S. laurentii* pSJF203'-8	*Not*I	**11**	*S. laurentii* pSJF203'-8	*Kpn*I
6	*S. laurentii* pSJF203'-9	*Not*I	**12**	*S. laurentii* pSJF203'-9	*Kpn*I
			13	*S. laurentii*	*Bam*HI

Generation of mutant vectors and Southern hybridization experiment conducted by Ya-Fen Jiang.


* 1 2 3 4 5 6 7 8 9 10 11 13 *
⇐ 12 kb
⇐ 6 kb
⇐ 3 kb
⇐ 1.7 kb
⇐ 500 bp


* 1 2 3 4 5 6 7 8 9 10 13 *
1 2 3
⇐ 6.4 kb
⇐ 3.2 kb
⇐ 500 bp

Figure 3.31: Map of the homologies found by end sequencing the *Bam*HI fragments surrounding pSJF114 which had been cloned into pBluescript KS^-. End sequence data was obtained from all of the plasmids listed in figure 3.26. Reads were between 150 and 450 bp in length.



peptide synthetase modules found or confirmed by sequencing (orientation known)

peptide synthetase modules found by Southern blots only (orientation unknown)

end sequence data from this fragment shows very high homology to geranyl geranyl pyrophosphate synthases and farnesyl pyrophosphate synthases

end sequence data from these fragments show homology to oxidoreductases

end sequence data shows homology to enoyl pyruvyl transferases

total sequence of this fragment has the beginning of an open reading frame with homology to several types of methyl transferases

3.4.4 Other peptide synthetase clusters in S. laurentii

A survey was done to determine if there are any further clusters of peptide synthetases present in *S. laurentii*. As several new clusters of peptide synthetases had been discovered by this time, and we had some information about peptide synthetases from this species, new primers based on the same cores as before were synthesized (figure 3.32).

PS-TGD forward primer:

G	T	X	M	Y	R	T	G	D
GGC	ACS	SGS	ATG	TAC	CGC	ACS	GGC	GAC C

PS-LGG reverse primer:

D	X	D	F	F	X	L	G	G
GCC	GCC	SAG	SNY	GAA	GAA	GTC	STC	GTC

Figure 3.32: Second set of PCR primers. Letters represent the standard one letter abbreviations for amino acids and nucleic acids, with the exception of S, Y, and N in the primer sequence. S represents a mixture of G and C. Y represents G, C, or A. N can be any base. The amino acid sequence for the reverse primer is the sequence in the order found on the peptide.

These primers should theoretically amplify a larger fraction of the peptide synthetases present in the species being studied since there is higher degeneracy and they are based on a better understanding of the codon preference of *S. laurentii* and other Streptomycetes.

PCR reactions using these new primers and the older primers allowed us to compare their efficacy. The new primers amplified a 500 bp product strongly and cleanly,

as well as a less strong, but still clear, product approximately 550 bp in length (figure 3.33).

The entire reaction was reamplified using *Pfu* DNA polymerase and the resulting products cloned into pCRScript KS^-. Seventy five plasmids were screened for insert, and eighteen were found to contain inserts 500 bp or larger. When these eighteen insert containing plasmids were sequenced from both ends it was found that eleven contained regions homologous to nonribosomal peptide synthetase core regions. Of these eleven sequences, four were redundant, yielding seven unique sequences.

In order to determine which of the seven unique sequences is from the known cluster of nonribosomal peptide synthetases in *S. laurentii*, and which are from previously unknown clusters, each of these inserts were used to synthesize Southern probes that were hybridized to restriction endonuclease digested *S. laurentii* genomic DNA. As the *Bam*HI and *Not*I digestion sites have been well characterized in the region encompassing the known peptide synthetase modules they were chosen to digest the DNA. If a probe hybridizes to a *Bam*HI fragment of the same size as pSJF114 (23 kb) or pSJF121 (12 kb), it is likely derived from the known cluster of nonribosomal peptide synthetase modules. Results from this hybridization experiment indicated that pEPSY002, pEPSY007, pEPSY045, pEPSY059, and pEPSY063 originate from pSJF114, the 23 kb *Bam*HI fragment containing most of the known modules. As the technique used does not allow for clear separation of fragments larger than 15 kb, it was necessary to confirm this experiment. The pEPSY063 probe was hybridized to pSJF17.2, pSJF18.3, pSJF20.2, and *S. laurentii* genomic DNA digested with *Bam*HI. The probe hybridized strongly to the 23 kb *Bam*HI fragment, with a weaker hy-

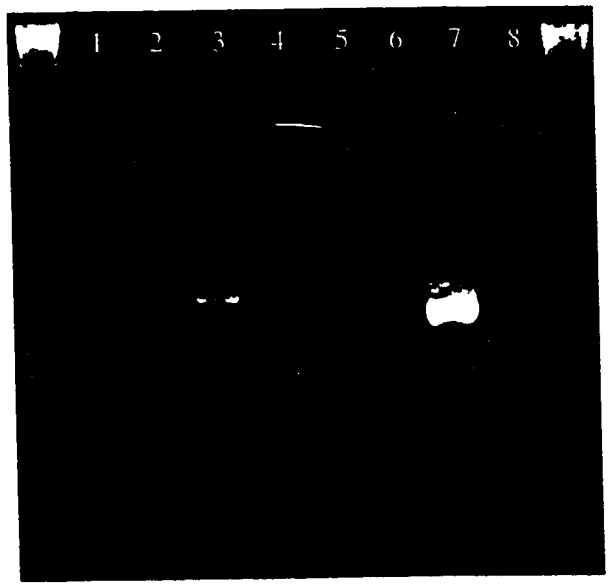
bridization to a 12 kb fragment which could be assigned as pSJF121 by its absence in pSJF17.2. The hybridization to the genomic sample corresponds well with those seen to the fragments resulting from the digestion of the cosmid DNA (figure 3.34).

Three of the PCR products, pEPSY027, pEPSY036, and pEPSY049, do not hybridize to a large *Bam*HI fragment in digested genomic DNA, or from a 12 kb fragment, which would be true if the product had been amplified from the modules found on pSJF121 or pSJF112 (figure 3.35). pEPSY049 shows some differences from known peptide synthetases, so pEPSY027 and pEPSY036 were chosen to screen the Supercos cosmid library of *S. laurentii* genomic DNA. Seven cosmids hybridizing to pEPSY027 were isolated, and six that hybridized to pEPSY036. When the cosmids showing hybridization to pEPSY027 were digested with the restriction endonucleases *Bam*HI and *Not*I, and hybridized with their respective probes, the expected hybridization to a 5 kb *Bam*HI digestion product and a 10 kb *Not*I digestion product was observed (figure 3.36). The cosmids that showed hybridization to pEPSY036, however, did not show a hybridization pattern consistent with that which was observed when the probe was hybridized against endonuclease digested genomic DNA. The strongest hybridizations observed were to cosmid Q-B12, which appears to be from the previously characterized cluster of nonribosomal peptide synthetase modules. The remaining cosmids showed only weak hybridization (figure 3.37). It is unclear why so many weakly hybridizing cosmids were isolated, but no cosmids containing the gene that this fragment was amplified from.

In order to determine if the two PCR products were amplified from a single cluster of nonribosomal peptide synthetases, the probes were cross hybridized to the

cosmids. Hybridization of the cosmids isolated using pEPSY027 to the pEPSY036 probe showed strongest homology to a 13 to 15 kb *Bam*HI fragment, and a 8 kb *Not*I fragment. This is not the hybridization pattern observed when the probe is hybridized against restriction endonuclease digested genomic DNA, nor is it consistent with the hybridizations observed in the cosmids isolated using this probe (figure 3.38). Hybridization of the cosmids isolated using pEPSY036 against the pEPSY027 probe again showed hybridization to the Q-B12 cosmid. with a fairly non-specific hybridization to the large *Bam*HI digestion product observed in the hybridization with the pEPSY036 probe (figure 3.39).

These results strongly indicate that the nonribosomal peptide synthetase modules from which these two PCR products were amplified are not proximally located on the genome. One can reasonably conclude from this information that at least three clusters of nonribosomal peptide synthetase modules can be found in *S. laurentii*.



Lane	1	2	3	4	5	6	7	8
Template	100x	100x	100x	100x	1x	1x	1x	1x
DMSO	10%	15%	10%	15%	10%	15%	10%	15%
Primer	old	old	new	new	old	old	new	new

Figure 3.33: PCR reactions performed in the initial analysis of the new primer set. Two different concentrations of mechanically sheared *S. laurentii* genomic DNA were used as templates, as well as two different DMSO concentrations. The new primers were used under the same conditions as the old primers to determine the relative efficiency of each primer. Amplification was catalyzed with *Taq* DNA polymerase. It is apparent that the new primers amplify the 500 bp product both more specifically and in higher yield than the old primers. The lower concentration of template seems to yield a higher specificity, and higher DMSO concentrations inhibit formation of the product.

Figure 3.34: Southern hybridization experiments conducted using the probe synthesized from pEPSY063, a PCR product believed to have been amplified from the region containing the known and characterized peptide synthetase modules.

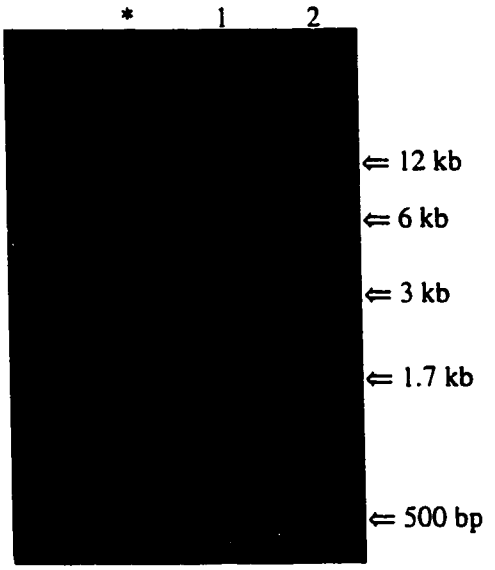
A: The original gel showing hybridization to a large (greater than 15kb) *Bam*HI fragment and a 2.5 kb *Not*I fragment produced by digesting *S. laurentii* genomic DNA.

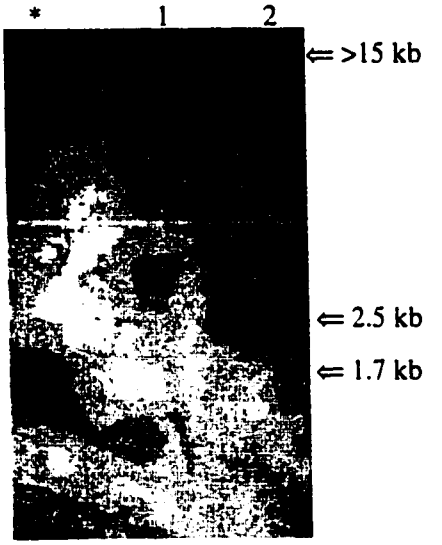
1 *S. laurentii* *Bam*HI **2** *S. laurentii* *Not*I

B: Hybridization against *S. laurentii* genomic DNA and cosmids covering the region surrounding the known cluster of nonribosomal peptide synthetase modules.

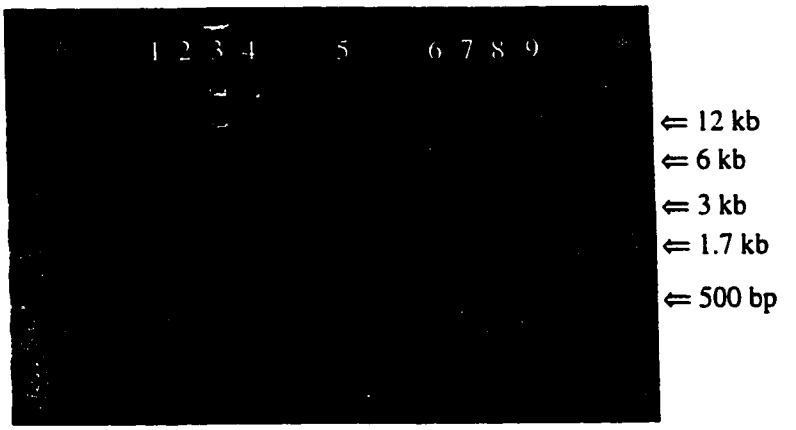
1	*S. laurentii*	*Bam*HI	**6**	*S. laurentii*	*Not*I
2	pSJF 17.2	*Bam*HI	**7**	pSJF 17.2	*Not*I
3	pSJF 18.3	*Bam*HI	**8**	pSJF 18.3	*Not*I
4	pSJF 20.2	*Bam*HI	**9**	pSJF 20.2	*Not*I
5	*S. laurentii*	*Kpn*I			

A


*
1
2
⇐ 12 kb
⇐ 6 kb
⇐ 3 kb
⇐ 1.7 kb
⇐ 500 bp


*
1
2
⇐ >15 kb
⇐ 2.5 kb
⇐ 1.7 kb

B


1 2 3 4 5 6 7 8 9
⇐ 12 kb
⇐ 6 kb
⇐ 3 kb
⇐ 1.7 kb
⇐ 500 bp

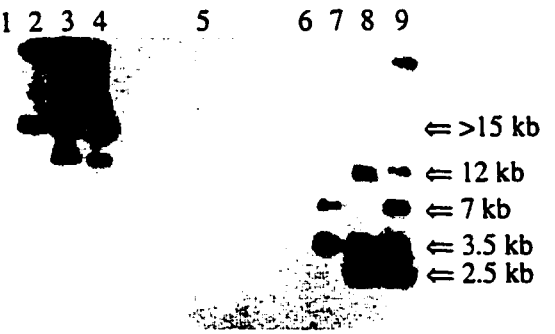
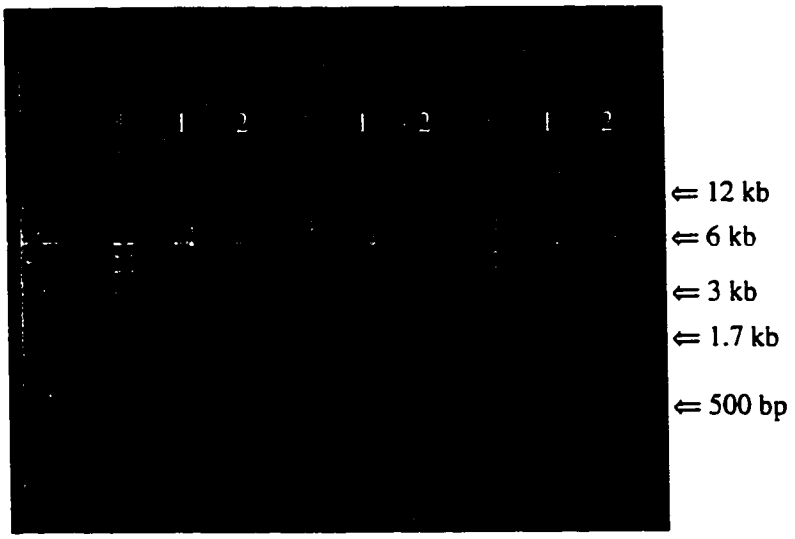

1 2 3 4 5 6 7 8 9
⇐ >15 kb
⇐ 12 kb
⇐ 7 kb
⇐ 3.5 kb
⇐ 2.5 kb

Figure 3.35: Southern hybridization of the pEPSY027, pEPSY036, and pEPSY049 against *Bam*HI and *Not*I restriction endonuclease digested genomic DNA. These fragments have homology to known nonribosomal peptide synthetases and appear not to be derived from the known cluster in *S. laurentii*. Lanes marked 1 are *Bam*HI digested *S. laurentii* genomic DNA, lanes marked 2 are *Not*I digested. Lanes containing molecular weight markers are indicated with an asterix.

Clone	*Bam*HI hybridizing fragments	*Not*I hybridizing fragments
pEPSY027	5 kb	10 kb
pEPSY036	4 kb	7 kb and 1.5 kb
pEPSY049	10 kb and 2.7 kb	1.2 kb

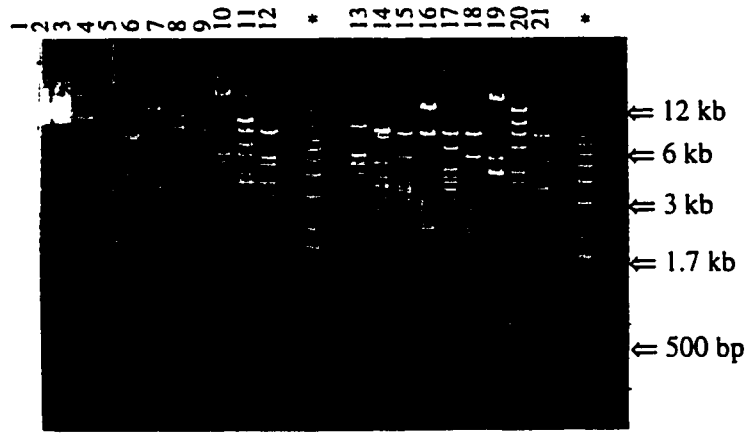


pEPSY027 pEPSY036 pEPSY049

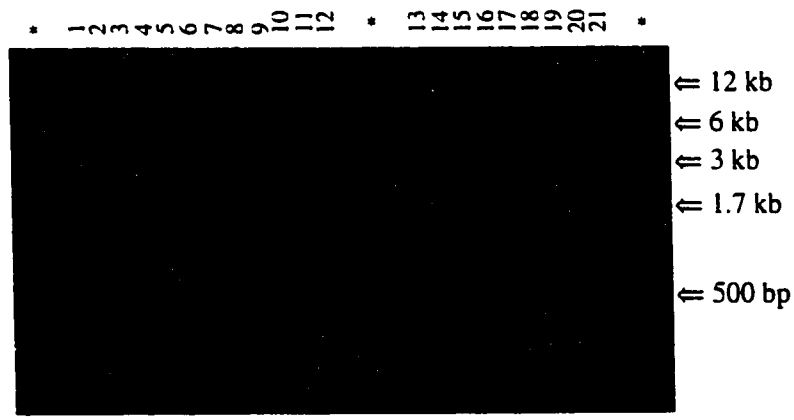
Figure 3.36: Cosmids isolated by screening a cosmid library for hybridization to the pEPSY027 probe. Cosmids were digested with the restriction endonucleases *Bam*HI, *Not*I, and double digested using both (panel A). Panel B shows the autoradiography film of the membrane after hybridization at low stringency (1x SSC, 0.1% SDS, 60°C) Panel C shows the autoradiography film of the membrane after hybridizing at a much higher stringency (0.1x SSC, 0.1% SDS, 65°C). Weaker hybridization was observed to a number of restriction endonuclease digestion products, indicating that this cluster may contain a number of nonribosomal peptide synthetase modules.

	Clone	Enzyme		Clone	Enzyme
1	pEPSY K-G7	*Bam*HI	**13**	pEPSY A-A10	*Bam*HI
2	pEPSY K-G7	*Not*I	**14**	pEPSY A-A10	*Not*I
3	pEPSY K-G7	*Bam*HI/*Not*I	**15**	pEPSY A-A10	*Bam*HI/*Not*I
4	pEPSY E-D5	*Bam*HI	**16**	pEPSY G-F2	*Bam*HI
5	pEPSY E-D5	*Not*I	**17**	pEPSY G-F2	*Not*I
6	pEPSY E-D5	*Bam*HI/*Not*I	**18**	pEPSY G-F2	*Bam*HI/*Not*I
7	pEPSY T-B6	*Bam*HI	**19**	pEPSY L-F8	*Bam*HI
8	pEPSY T-B6	*Not*I	**20**	pEPSY L-F8	*Not*I
9	pEPSY T-B6	*Bam*HI/*Not*I	**21**	pEPSY L-F8	*Bam*HI/*Not*I
10	pEPSY G-D2	*Bam*HI			
11	pEPSY G-D2	*Not*I			
12	pEPSY G-D2	*Bam*HI/*Not*I			

A


1 2 3 4 5 6 7 8 9 10 11 12 * 13 14 15 16 17 18 19 20 21 *
⇐ 12 kb
⇐ 6 kb
⇐ 3 kb
⇐ 1.7 kb
⇐ 500 bp

B


* 1 2 3 4 5 6 7 8 9 10 11 12 * 13 14 15 16 17 18 19 20 21 *
⇐ 12 kb
⇐ 6 kb
⇐ 3 kb
⇐ 1.7 kb
⇐ 500 bp

C

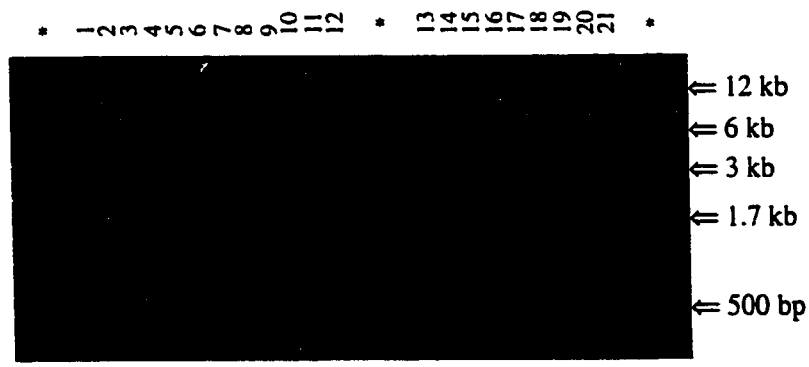
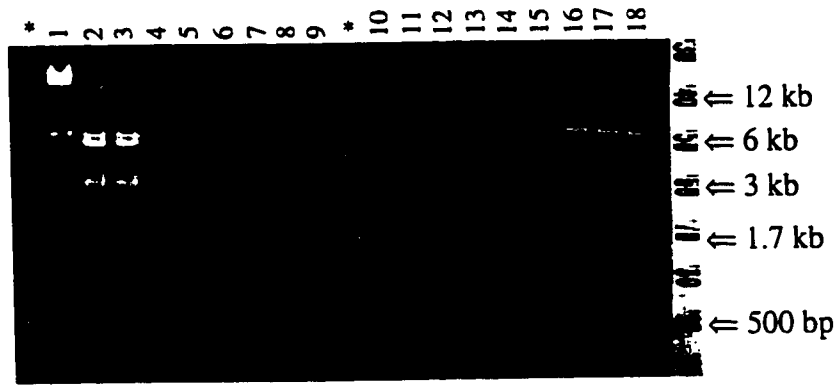

* 1 2 3 4 5 6 7 8 9 10 11 12 * 13 14 15 16 17 18 19 20 21 *
⇐ 12 kb
⇐ 6 kb
⇐ 3 kb
⇐ 1.7 kb
⇐ 500 bp

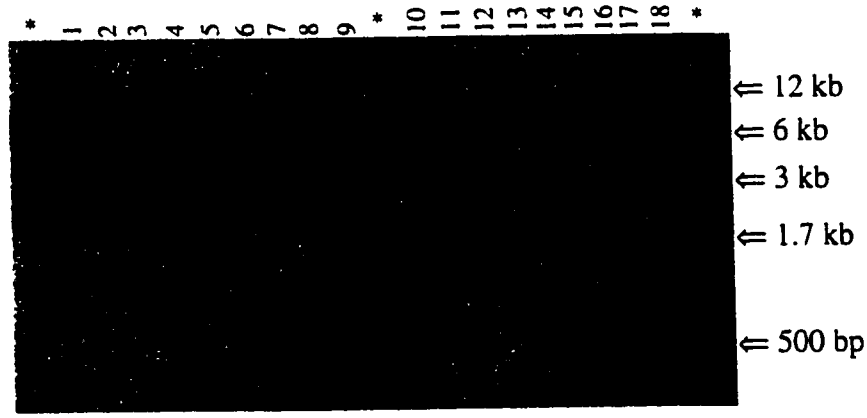
Figure 3.37: Cosmids isolated by screening a cosmid library for hybridization to the pEPSY036 probe. Cosmids were digested with the restriction endonucleases *Bam*HI, *Not*I, and double digested using both. The upper hybridization film shows hybridization at low stringency ($1x$ SSC, 0.1% SDS, 60°C), and the lower was washed at a much higher stringency ($0.1x$ SSC, 0.1% SDS, 65°C). The hybridization observed in the isolated cosmids does not agree with the results obtained when the same probe was hybridized against restriction endonuclease digested genomic DNA, indicating that this region does not contain the region of the genome from which pEPSY036 was amplified, and that cosmids containing that fragment have not yet been isolated.

	Clone	Enzyme		Clone	Enzyme
1	pEPSY Q-B12	*Bam*HI	**10**	pEPSY N-D1	*Bam*HI
2	pEPSY Q-B12	*Not*I	**11**	pEPSY N-D1	*Not*I
3	pEPSY Q-B12	*Bam*HI/*Not*I	**12**	pEPSY N-D1	*Bam*HI/*Not*I
4	pEPSY N-A11	*Bam*HI	**13**	pEPSY P-E3	*Bam*HI
5	pEPSY N-A11	*Not*I	**14**	pEPSY P-E3	*Not*I
6	pEPSY N-A11	*Bam*HI/*Not*I	**15**	pEPSY P-E3	*Bam*HI/*Not*I
7	pEPSY L-F8	*Bam*HI	**16**	pEPSY R-P7	*Bam*HI
8	pEPSY L-F8	*Not*I	**17**	pEPSY R-P7	*Not*I
9	pEPSY L-F8	*Bam*HI/*Not*I	**18**	pEPSY R-P7	*Bam*HI/*Not*I

A


* 1 2 3 4 5 6 7 8 9 * 10 11 12 13 14 15 16 17 18
⇐ 12 kb
⇐ 6 kb
⇐ 3 kb
⇐ 1.7 kb
⇐ 500 bp

B


* 1 2 3 4 5 6 7 8 9 * 10 11 12 13 14 15 16 17 18 *
⇐ 12 kb
⇐ 6 kb
⇐ 3 kb
⇐ 1.7 kb
⇐ 500 bp

C

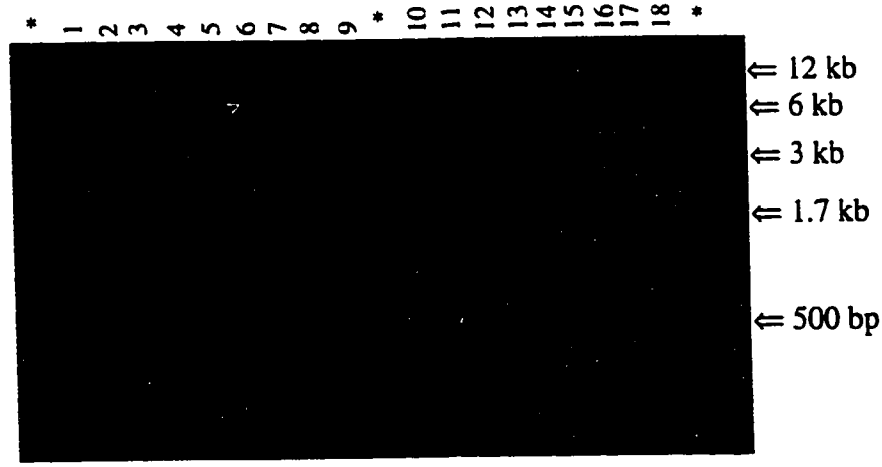
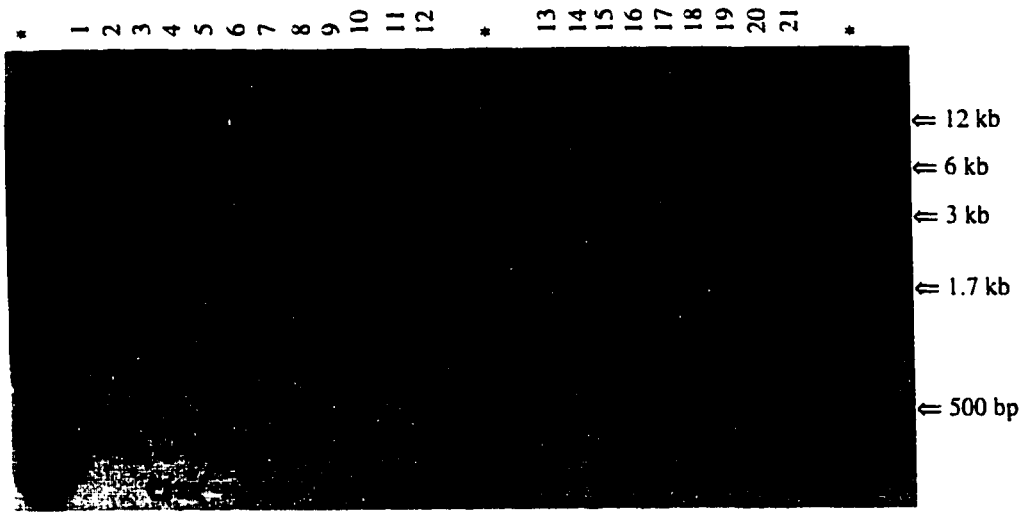

* 1 2 3 4 5 6 7 8 9 * 10 11 12 13 14 15 16 17 18 *
⇐ 12 kb
⇐ 6 kb
⇐ 3 kb
⇐ 1.7 kb
⇐ 500 bp

Figure 3.38: Cross hybridization using the pEPSY036 probe against cosmids isolated using the pEPSY027 probe digested with the restriction endonucleases *Bam*HI, *Not*I, and double digested with both. Panel A shows the autoradiography film of the membrane after washing at low stringency (2x SSC, 0.1% SDS, 60°C). Panel B shows the autoradiography film of the membrane after washing at higher stringency (0.1x SSC, 0.1% SDS, 68°C).

	Clone	Enzyme		Clone	Enzyme
1	pEPSY K-G7	*Bam*HI	**13**	pEPSY A-A10	*Bam*HI
2	pEPSY K-G7	*Not*I	**14**	pEPSY A-A10	*Not*I
3	pEPSY K-G7	*Bam*HI/*Not*I	**15**	pEPSY A-A10	*Bam*HI/*Not*I
4	pEPSY E-D5	*Bam*HI	**16**	pEPSY G-F2	*Bam*HI
5	pEPSY E-D5	*Not*I	**17**	pEPSY G-F2	*Not*I
6	pEPSY E-D5	*Bam*HI/*Not*I	**18**	pEPSY G-F2	*Bam*HI/*Not*I
7	pEPSY T-B6	*Bam*HI	**19**	pEPSY L-F8	*Bam*HI
8	pEPSY T-B6	*Not*I	**20**	pEPSY L-F8	*Not*I
9	pEPSY T-B6	*Bam*HI/*Not*I	**21**	pEPSY L-F8	*Bam*HI/*Not*I
10	pEPSY G-D2	*Bam*HI			
11	pEPSY G-D2	*Not*I			
12	pEPSY G-D2	*Bam*HI/*Not*I			

A


* 1 2 3 4 5 6 7 8 9 10 11 12 * 13 14 15 16 17 18 19 20 21 *
⇐ 12 kb
⇐ 6 kb
⇐ 3 kb
⇐ 1.7 kb
⇐ 500 bp

B

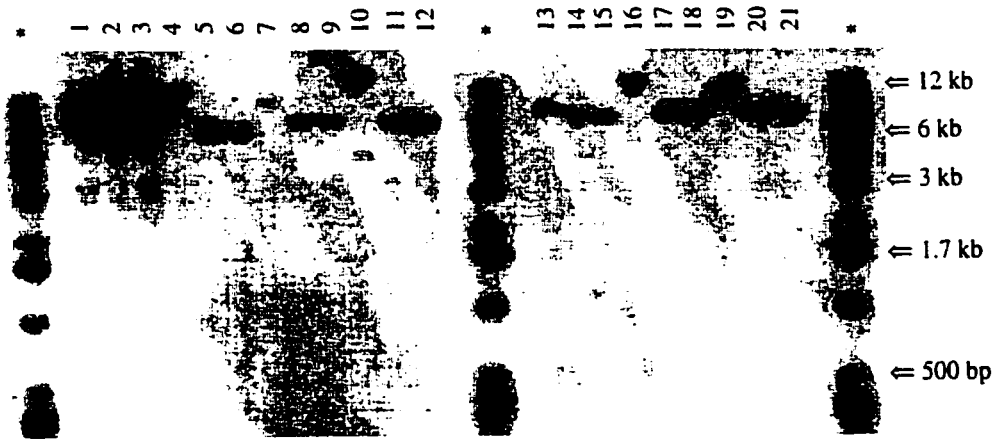
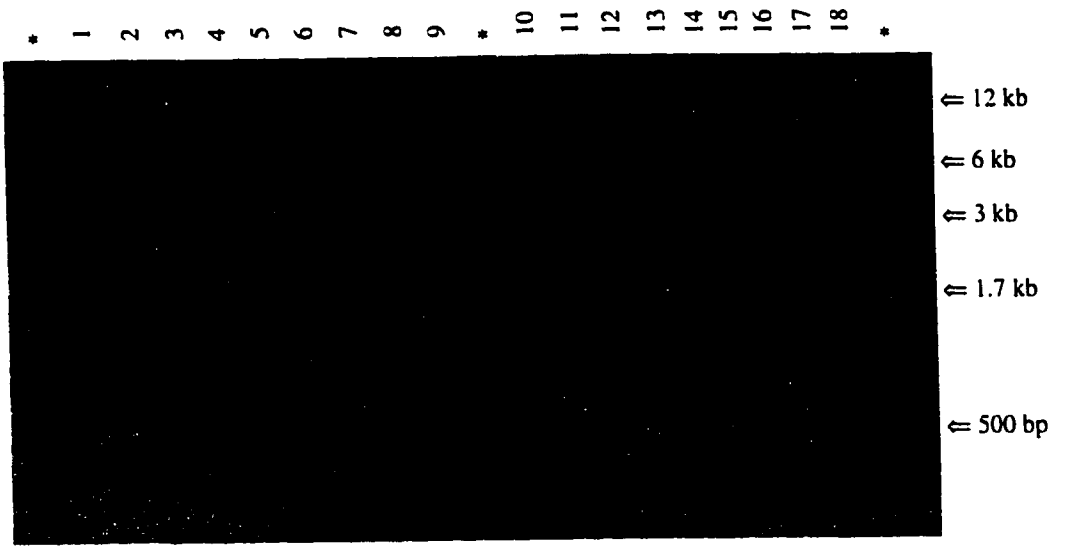

* 1 2 3 4 5 6 7 8 9 10 11 12 * 13 14 15 16 17 18 19 20 21 *
⇐ 12 kb
⇐ 6 kb
⇐ 3 kb
⇐ 1.7 kb
⇐ 500 bp

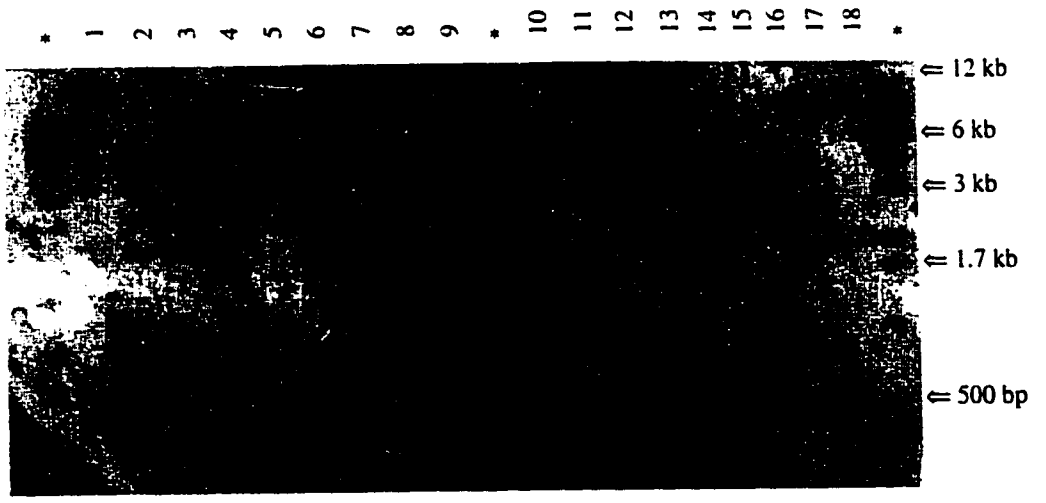
Figure 3.39: Cross hybridization using the pEPSY027 probe against cosmids isolated using the pEPSY036 probe digested with the restriction endonucleases *Bam*HI, *Not*I, and double digested with both. Panel A shows the autoradiography film of the membrane after washing at low stringency (2x SSC, 0.1% SDS, 60°C). Panel B shows the autoradiography film of the membrane after washing at higher stringency (0.1x SSC, 0.1% SDS, 68°C).

	Clone	Enzyme		Clone	Enzyme
1	pEPSY Q-B12	*Bam*HI	**10**	pEPSY N-D1	*Bam*HI
2	pEPSY Q-B12	*Not*I	**11**	pEPSY N-D1	*Not*I
3	pEPSY Q-B12	*Bam*HI/*Not*I	**12**	pEPSY N-D1	*Bam*HI/*Not*I
4	pEPSY N-A11	*Bam*HI	**13**	pEPSY P-E3	*Bam*HI
5	pEPSY N-A11	*Not*I	**14**	pEPSY P-E3	*Not*I
6	pEPSY N-A11	*Bam*HI/*Not*I	**15**	pEPSY P-E3	*Bam*HI/*Not*I
7	pEPSY L-F8	*Bam*HI	**16**	pEPSY R-P7	*Bam*HI
8	pEPSY L-F8	*Not*I	**17**	pEPSY R-P7	*Not*I
9	pEPSY L-F8	*Bam*HI/*Not*I	**18**	pEPSY R-P7	*Bam*HI/*Not*I

A


* 1 2 3 4 5 6 7 8 9 * 10 11 12 13 14 15 16 17 18 *
⇐ 12 kb
⇐ 6 kb
⇐ 3 kb
⇐ 1.7 kb
⇐ 500 bp

B


* 1 2 3 4 5 6 7 8 9 * 10 11 12 13 14 15 16 17 18 *
⇐ 12 kb
⇐ 6 kb
⇐ 3 kb
⇐ 1.7 kb
⇐ 500 bp

Chapter 4

DISCUSSION AND CONCLUSIONS

Thiostrepton is a highly modified sulfur rich peptide antibiotic. It has been used extensively as a selective marker in Streptomycete research, and as a topical antibiotic in veterinary medicine. Previous studies on thiostrepton have elucidated the amino acid precursors of the antibiotic. Research conducted earlier in this laboratory indicated that thiostrepton may be assembled nonribosomally. Comparison with other systems suggests that thiostrepton is produced by the nonribosomal thiotemplate mechanism, in which large enzyme complexes build the peptide by activating an amino acid to its acyl adenylate in a reaction requiring ATP and releasing pyrophosphate, covalently linking it to the enzyme via a nucleophilic attack of the sulfhydryl group of a phosphopantetheine cofactor, and condensation with a similarly activated and attached amino acid or peptide. In addition to the peptide building reactions there is a wide variety of activities required for modifications of the peptide including methylation, dehydration, thiazole and thiazoline ring formation, condensation, hydration, and oxidation.

A recent focus of research in bacterial secondary metabolism has involved the combination of genes involved in the biosynthesis of different structures to produce new, and potentially pharmaceutically interesting structures. The evolution of pathogenic bacteria showing resistance to known antibiotics has made this search for new com-

pounds showing antimicrobial activity a high priority. The modular construction of the genes encoding the nonribosomal peptide synthetases make them an ideal candidate for this approach. With twenty six natural amino acids, and the wide variety of ways in which they are modified in the production of these peptide antibiotics, a near infinite array of potential structures could be produced using these genes.

In an earlier attempt to find the thiostrepton biosynthetic gene cluster it was demonstrated that the region on the genome surrounding the known thiostrepton resistance gene does not contain genes involved in thiostrepton biosynthesis. Previously it has always been found that an antibiotic gene cluster is associated with a resistance mechanism for its product. Examples have been found , however, where the genes are clustered with only one of several resistance-encoding genes. Attempts to find another thiostrepton resistance gene in *S. laurentii* have been unsuccessful.

The focus of this study was to explore other ways to find the thiostrepton biosynthetic gene cluster. Several approaches to finding the cluster were undertaken, including the use of a nonribosomal peptide synthetase gene from another streptomycete as a Southern hybridization probe, purification of a pathway enzyme to use in a reverse genetics approach, and the development of PCR primers based on known conserved core sequences found in nonribosomal peptide synthetases.

Southern hybridization using DNA from genes believed to be similar to a biosynthetic gene encoding the biosynthesis of the antibiotic of interest has been used to find many secondary metabolite clusters. A fragment of a nonribosomal peptide synthetase gene involved in penicillin biosynthesis in *S. clavulerigus* was used to synthesize a probe which was hybridized against restriction endonuclease digested genomic

DNA from three thiopeptide producers. This probe did not show specific hybridization to the genomic DNA from the species tested. While there are core regions of nonribosomal peptide synthetases which show a high degree of amino acid similarity, the overall similarity at the DNA level is fairly low. While nonribosomal peptide synthetase encoding gene clusters have been found using this method, it is not always successful.

Purification of a pathway enzyme for use in a reverse genetics approach has been used to clone several biosynthetic clusters. This method involves purification of a sufficient amount of enzyme to obtain N-terminal sequence information. This sequence information can be used to design an oligonucleotide based on the codon preference of the species being studied. Two enzymes had been partially characterized earlier in this laboratory. The tryptophan methyl transferase methylates the indole two position of tryptophan, an early step in the biosynthesis of the HEQ moiety of thiostrepton. The enzyme that activates HEQ to its acyl adenylate prior to attachment to the antibiotic is an acyl adenylase, catalyzing a reaction requiring ATP that releases pyrophosphate with the formation of the HEQ adenylate. The acyl adenylase was chosen as the target enzyme due to its higher stability and the availability of a rapid assay for enzymatic activity.

Several protein separation techniques were employed while attempting to purify this protein. The protein was shown to have low stability, with a half life at 4°C of about eight hours at pH 6.8, its stability maximum. Stability was found to improve somewhat when glycerol was included in the buffers. The protein was found to not bind to DEAE cellulose or DEAE sephacel anion exchange resins, even at high pH

values, and to precipitate between 35% and 65% saturated ammonium sulfate. Initial purification attempts used cell free extract purified by a negative binding step with DEAE cellulose and ammonium sulfate fractionation, followed by FPLC columns. Hydrophobic interaction chromatography, strong cation chromatography, and strong anion chromatography were employed to purify the protein. After separation by these techniques there was insufficient protein to visualize by SDS-PAGE, followed by Coomassie blue staining.

In order to increase the amount of protein being processed, and to decrease the total amount of time involved in the purification several bulk chromatographic materials were investigated. The phenyl sepharose hydrophobic interaction chromatography resin was shown to bind the adenylase strongly, and in combination with the ammonium sulfate fractionation yielded a 130 fold purification. While the protein will not bind to anion exchange resins, binding was observed to a cation exchange resin at pH 6.0. This is well below the stability maximum of the protein, but this step still afforded a thirty fold purification. Later in the study a stronger cation exchange bulk resin was used which showed excellent binding characteristics at pH 6.8.

A purification scheme was designed using these steps that did not involve any buffer exchange steps, kept the enzyme at its stability maximum, and incorporated glycerol in the buffers where it did not interfere with the chromatography. The entire purification required thirty hours from the harvesting of the cells, and SDS-PAGE analysis following purification showed a clear protein band at the appropriate size.

This protein was isolated and the N-terminal protein sequence determined by Edman degradation. An oligonucleotide was designed based on the protein sequence

obtained, and end labeled for use as a Southern hybridization probe. Probing against restriction endonuclease digested *S. laurentii* genomic DNA did not show any specific hybridization. The lack of specificity observed is probably due to the degeneracy of the coding sequence, and its short length. This experiment also suffers from the inability to have a positive control hybridization, unlike the hybridization using the ACVS probe where a large gene fragment was used to synthesize the probe, which could be hybridized against as a positive control.

These difficulties could be resolved by using a PCR approach with the oligonucleotide based on the N-terminal protein sequence as the forward primer. PCR is considerably more sensitive and specific than a Southern hybridization in that two primers, one forward and one reverse, must anneal to the genomic DNA in close enough proximity to each other that they can produce a single amplified product. For the reverse primers it was necessary to obtain some internal sequence data from the peptide. Two fragments of internal sequence were obtained and used to design reverse primers for use with the original oligonucleotide whose design was based on the N-terminal sequence.

Conditions were found where a clear band was observed using each of the primers, and the resulting fragments were cloned and sequenced. Since both reactions were performed using the same forward primer, it would be expected that the longer sequence would contain the sequence of the shorter. This was not found to be the case, however. This could be due to the weak internal sequence information obtained, leading to the design of incorrect primers, or more likely one of the primers is not specific enough to amplify the correct product. In neither case, however, was homol-

ogy to acyl activating enzymes observed, making it unlikely that the protein isolated for protein sequencing was the acyl adenylase.

Several interesting properties of the acyl adenylase were observed during this attempted purification. The enzyme is very hydrophobic, as demonstrated by its strong binding to the hydrophobic interaction columns, by its lack of binding to anion exchange resins, and its weak interactions with cation exchange resins. This may contribute to the poor stability in buffer solutions, and may explain the stabilizing effect seen when glycerol is present in the buffer. It seems likely that this enzyme is very active, but present in the cell in very low concentration. The lack of observable bands in the expected size range when analyzing active fractions by SDS-PAGE, followed by Coomassie blue staining, supports this hypothesis. Protein sequencing by Edman degradation requires enough protein to be clearly visible by this staining method. The low stability of the enzyme does not allow for combining purified extracts from multiple purification attempts, making it a poor candidate for this approach.

The other main approach undertaken to find the thiostrepton biosynthetic gene cluster involved a direct genetic approach in which PCR primers based on conserved core sequences found in nonribosomal peptide synthetases were employed to amplify fragments of this family of genes from the *S. laurentii* genome. Previous studies in this lab had indicated that thiostrepton is produced nonribosomally, and early work in this study determined that there are large nonribosomal peptide synthetases present in the cell-free extract. This was demonstrated by showing that partially purified cell-free extract was capable of exchanging radiolabeled pyrophosphate into ATP in the presence of natural amino acids, and that radiolabeled amino acids bound to

proteins in the presence of ATP. This activity was also shown to remain above an ultrafiltration membrane with a pore size that does not allow compounds larger than 300 kDa to pass through, showing it to be at least that large, which is consistent with known nonribosomal peptide synthetases. Ribosomal peptide synthesis activates the amino acids using GTP, and would not be observed by this assay. As no other peptide secondary metabolites have been reported to be produced by *S. laurentii*, it seems likely that thiostrepton is produced by the thiotemplate mechanism.

Primers were designed based on conserved amino acid sequences found in nonribosomal peptide synthetases, and used to amplify appropriately sized DNA fragments from *S. laurentii* genomic DNA. These fragments were cloned and sequenced. Three of the clones examined had homology to nonribosomal peptide synthetases, and were later shown to be identical. These were used as probes to screen a cosmid library for hybridizing clones. Cosmids showing hybridization were isolated and their *Bam*HI restriction endonuclease digestion sites mapped. The products obtained by digestion using *Bam*HI were cloned, and the fragment showing strongest hybridization to the probe isolated for further analysis. End sequence analysis of *Not*I subclones of this fragment gave evidence of five nonribosomal peptide synthetase modules present on this fragment. An adjacent *Bam*HI fragment showed weak hybridization to the probe, and appears by Southern hybridization against restriction endonuclease digestion products to contain at least two modules.

These results are promising, but conclusive proof that this is the thiostrepton biosynthetic cluster requires that the gene cluster be disrupted and the loss of production of the antibiotic demonstrated. Several insertional mutants were generated

using the *Not*I restriction endonuclease generated subclones from the region containing the known modules, and the resulting mutants grown under culture conditions in which thiostrepton is produced in the wild type. These mutants were all shown to produce thiostrepton at approximately wild type levels.

It is possible that the thiostrepton biosynthetic cluster is the remnant of an older cluster that contains some inactive nonribosomal peptide synthetase modules. There is reason to believe that this is a possiblity. If one examines the structure of nosiheptide it appears likely that it is the product of a somewhat crippled thiostrepton biosynthetic cluster in which five activating modules are non-functional. In hopes that this is the case end sequence information was obtained for all of the cloned *Bam*HI restriction endonuclease digestion products in the region surrounding the known cluster of nonribosomal peptide synthetase modules. While several interesting homologies were observed, none of them appeared to contain nonribosomal peptide synthetase modules. As seventeen of these modules would be required to construct the peptide chain of thiostrepton this is fairly conclusive proof that the thiostrepton biosynthetic gene cluster is not located in this region.

In search of evidence of another cluster of peptide synthetases new PCR primers were designed on the now much larger database of protein sequences of nonribosomal peptide synthetases. These primers were used in amplification, cloning, and sequencing experiments similar to those performed earlier, and screened to determine if they are from the known cluster. Two modules not originating from the original cluster were found and used to screen a *S. laurentii* cosmid library. Hybridizing clones were isolated, and used in hybridization experiments to determine if these two modules are

proximally located on the genome. These fragments were shown to not be clustered with each other, indicating that there are at least two more clusters of nonribosomal peptide synthetase modules on the *S. laurentii* genome.

Recent results have shown that this is not an unusual result. Moore *et al.* [74] performed PCR amplifications with the primers used in this study on a variety of *Streptomyces* species and found that nonribosomal peptide synthetase modules could be amplified from all of the strains tested, even though many of these are not known to produce peptide secondary metabolites. These primers were also used in a parallel project in our laboratory attempting to find the nosiheptide biosynthetic gene cluster in *S. actuosus*. A cluster containing twelve nonribosomal peptide synthetase modules, the expected number for nosiheptide, was isolated from the genome [68]. A gene deletion experiment was successfully performed that deleted well over half of these modules, and the resulting mutant was shown to still produce nosiheptide at wild type levels [49]. It appears that there can be a multitude of clusters of these nonribosomal peptide synthetases in *Streptomyces* species that are either silent under the conditions tested, or produce secondary metabolites at a level below detection. The state of knowledge about the genetics of these nonribosomal peptide synthetases is reaching a point where predictions on the structure of the metabolite produced can be made based on the gene sequence [66]. These techniques could be used to develop an educated guess as to the products of these silent gene clusters.

Future work on this project would include screening the other nonribosomal peptide synthetase clusters to determine if there are any that are of sufficient size to produce thiostrepton. While there is no conclusive proof that this antibiotic is produced

using the thiotemplate mechanism, it still appears to be the most likely biosynthetic pathway.

Another enzyme from the biosynthetic pathway could be chosen for a reverse genetics approach to finding the cluster as was done in this study. With the wide variety of activities necessary for the biosynthesis of a large and complex structure like thiostrepton there is a large choice of enzymes to purify. One logical choice for purification would be the nonribosomal peptide synthetases that were partially purified and characterized in section 3.3. While we have demonstrated that there are at least three clusters of nonribosomal peptide synthetase modules, thiostrepton is the only peptide secondary metabolite that has been reported to be produced by *S. laurentii*. It is likely that the clusters of nonribosomal peptide synthetase modules that are not involved in thiostrepton biosynthesis are silent, and that any active nonribosomal peptide synthetases are involved in the production of thiostrepton.

If these approaches do not result in finding the thiostrepton biosynthetic gene cluster, it becomes considerably more likely that this antibiotic is produced atypically. It is possible that there is another, yet undiscovered, biosynthetic machinery for producing peptide structures nonribosomally, and that the genes encoding the enzymes responsible for this novel mechanism are not clustered with genes encoding resistance to thiostrepton. It should also not be forgotten that the evidence indicating that thiostrepton is produced nonribosomally is circumstantial, and the possibility still exists that this is the mechanism by which the antibiotic is synthesized in *S. laurentii*.

An interesting study which could be performed based on this research would involve identifying the product of the nonribosomal peptide synthetases clusters found

in *S. laurentii* and demonstrated not to produce thiostrepton. Perhaps by transforming *S. lividans* or *S. coelicolor* with pOJ446 cosmids containing this gene cluster one could observe production of a novel secondary metabolite. With the presence of so many silent clusters in Streptomycete DNA, this could prove a fascinating route to finding new structures.

Thiostrepton is a large and highly modified peptide antibiotic that shows high antimicrobial activity. While its low bioavailability makes it a poor candidate for human medicine, its interesting structure and activity make it an excellent target for directed genetics. This thesis represents a concerted attempt to find the thiostrepton biosynthetic gene cluster, involving three major approaches. While this cluster was not located, much has been learned about nonribosomal peptide biosynthesis in *S. laurentii*. A great deal has also been learned about the enzymology of HEQ activation prior to attachment to thiostrepton. Several clusters of nonribosomal peptide synthetase modules have been found in the *S. laurentii* genome, and one demonstrated not to be involved in thiostrepton biosynthesis. Whether or not the other clusters encode genes involved in the biosynthesis of this molecule remains an open question meriting further study.

Chapter 5

EXPERIMENTAL

5.1 Synthesis

General

Flash column chromatography ("chromatography") was done on 230-400 mesh silica gel. Elemental analyses were carried out by the Canadian Microanalytical Service, Ltd., 207-8116 Alexander Road, RR#7, Delta, B.C. V4G 1G7.

The solvents $CHCl_3$ (CaH_2), CH_2Cl_2 (CaH_2), Et_2O (Na-benzophenone), MeOH ($(MeO)_2Mg$), pyridine (CaH_2) and THF (Na-benzophenone) were distilled from the respective drying agents under argon. All other solvents were HPLC grade and were used without further purification. Argon was dried by passing through a bed of 3 Å molecular sieves. Unless otherwise stated, all reagents were used as supplied by either Sigma or Aldrich.

N-Benzoyl-2-cyano-1,2-dihydroquinoline

Freshly distilled quinoline (46.5 ml, 360 mmol) in CH_2Cl_2 (400 ml) was added to a vigorously stirred solution of KCN (29 g, 445 mmol) in water (160 ml). PhCOCl (58 ml, 360 mmol) was added dropwise and the two-phase system stirred vigorously at room temperature overnight. The two phases were separated and the aqueous phase

was extracted with CH_2Cl_2. The combined organic extracts were washed successively with water, 5% aqueous HCl, 5% aqueous NaOH, and finally water. The organic phase was dried (anhydrous Na_2SO_4), filtered, and the solvent removed by evaporation *in vacuo* to give a white powder. The product was obtained as white needles on recrystallization from 60% aqueous acetone (43.56 g, 167 mmol, 46.5%): mp (from EtOAc) 151 - 153 °C; δ_H (300 MHz, $CDCl_3$) 6.09 (1H, dd, J = 9.1 and 6.4 Hz), 6.19 (1H, d, J = 6.4 Hz), 6.57 (1H, d, J = 8.1 Hz), 6.83 (1H, dd, J = 9.1 and 1.5 Hz), 6.92 (1H, dt, J = 8.0 and 1.6 Hz), 7.09 (1H, dt, J = 7.7 and 1.2 Hz), 7.22 (1H, dd, J = 7.5 and 1.4 Hz), 7.27 (2H, t, J = 7.1 Hz), 7.35 (2H, dd, J = 7.0 and 1.3 Hz), and 7.40 (1H, tt, J = 6.7 and 1.5 Hz); δ_C (75.5 MHz, $CDCl_3$) 42.15, 115.71, 120.2, 125.1, 125.6, 125.9, 127.9, 128.2, 128.3, 129.1, 129.5, 131.4, 133.4, 134.4, and 169.1.

Quinoline-2-carboxylic acid hydrobromide monohydrate

N-Benzoyl-2-cyano-1,2-dihydroquinoline (96.82 g, 372 mmol) was suspended in glacial AcOH (300 ml). Aqueous HBr solution (48% w/v, 290 ml) was added cautiously and then the yellow slurry was heated at reflux for 15 minutes to give a dark red-black solution. The solution was allowed to cool and then left to stand overnight. The crystals which formed were collected by filtration, washed with a little ice-cold glacial AcOH and then with Et_2O. A second crop of crystals was obtained by saturating the filtrate with Et_2O. The combined crops were dried in vacuo to give the product as orange needles. The wet crystals were used directly in the next reaction: mp 224-226°C; δ_H (300 MHz, CD_3OD) 8.05 (1H, dt, J = 7.6 and 0.94 Hz), 8.25 (1H, ddd, J = 8.8, 7.0, and 1.2 Hz), 8.42 (1H, d, J = 8.3 Hz), 8.53 (1H, d, J = 8.6 Hz), 8.57 (1H,

d, J = 8.5 Hz), and 9.33 (1H, d, J = 8.5 Hz); δ_C (75.5 MHz, CD_3OD) 122.7, 123.1, 130.4, 132.2, 132.3, 137.2, 140.5, 144.9, 149.3, and 161.8.

Quinoline-2-carboxylic acid

The hydrobromide prepared above was dissoved with heating in water (250 ml). Concentrated ammonium hydroxide (73 ml) was added until the formation of a white precipitate was observed. Glacial acetic acid (73 ml) was added to the solution, which was warmed until the majority of the solid had disappeared. The solution was filtered while hot, then glacial acetic acid (146 ml) was added, forming a white precipitate that was collected and dried *in vacuo* overnight and used directly in the next reaction.

Methyl quinoline-2-carboxylate

The acid prepared above was dissolved in dry MeOH (700 ml). $SOCl_2$ (70 ml, 960 mmol) was carefully added. The resulting suspension was heated at reflux for 18 h. The suspension was left to cool to room temperature and then carefully poured into aqueous, saturated $NaHCO_3$ solution. The resulting mixture was extracted repeatedly with CH_2Cl_2. The organic extracts were combined, dried over anhydrous Na_2SO_4, filtered, and the solvent removed by evaporation *in vacuo* to give a pale green powder which rapidly turned cream colored on exposure to air (32.6 g, 174 mmol, 46.8% over three steps): mp (from $CHCl_3$) 81 - 82 °C; δ_H (300 MHz, $CDCl_3$) 4.05 (3H, s), 7.63 (1H, dt, J = 7.6 and 0.7 Hz), 7.75 (1H, ddd, J = 8.6, 7.0, and 1.4 Hz), 7.84 (1H, dd, J = 7.7 and 1.1 Hz), 8.16 (1H, dd, J = 8.5 and 1.5 Hz), 8.27 (2H, d, J = 8.5 Hz); δ_C (75.5 MHz, $CDCl_3$) 53.13, 121.0, 127.5, 128.6, 129.3, 130.2, 130.6, 137.2, 147.5, 147.8,

and 165.9.

Methyl 4-acetylquinoline-2-carboxylate

$FeSO_4.7H_2O$ (58 g, 209 mmol) in water (500 ml) and aqueous H_2O_2 solution (500 ml; 209 mmol) were added separately and simultaneously, over a 5 minute period, to a stirred solution of methyl quinoline-2-carboxylate (40.2 g, 215 mmol) and TFA (16 ml, 207 mmol) in acetaldehyde (1,200 ml, 21.6 mol). The addition of the reagents caused the solvent to reflux vigorously. After five minutes the pale orange solution was diluted with water (1,500 ml) and extracted with CH_2Cl_2. The organic extracts were combined, dried over anhydrous Na_2SO_4, filtered, and then the solvent was removed by evaporation *in vacuo* to give a yellow-brown powder. The product was obtained by chromatography on silica gel, eluting with EtOAc-hexanes (1:1), as a pale yellow powder (33.82 g, 148 mmol, 69%): mp 119 - 121 °C (from EtOAc-hexanes); δ_H (200 MHz, $CDCl_3$) 2.66 (3H, s), 3.96 (3H, s), 7.56 (1H, ddd, J = 8.4, 6.9, and 1.5 Hz), 7.67 (1H, ddd, J = 7.7, 7.55, and 1.6 Hz), 8.17 (1H, dd, J = 8.4 and 1.0 Hz), 8.26 (1H, s), 8.35 (1H, dd, J = 8.2 and 1.3 Hz); δ_C (75.5 MHz, $CDCl_3$) 29.70, 53.07, 119.7, 124.5, 125.2, 130.1, 130.3, 130.8, 143.6, 148.3, 165.1, and 200.2.

4-Acetylquinoline-2-carboxylic acid

A 30% (w/v) aqueous NaOH solution (1.33 ml, 10 mmol) was added to a stirred solution of methyl 4-acetylquinoline-2-carboxylate (250 mg, 1.1 mmol) in THF (3.67 ml) at room temperature. The solution immediately turned yellow and a thick yellow precipitate formed within 5 minutes. After 100 minutes the precipitate was dissolved

by the addition of 0.1 N NaOH solution (50 ml). The solution was washed twice with CH_2Cl_2 and then acidified to pH 1 with dilute aqueous HCl. The product was extracted with EtOAc. The organic extracts were pooled, dried over anhydrous Na_2SO_4, filtered, and the solvent removed by evaporation *in vacuo* to give a pale green crystalline solid. The acid was further purified by recrystallization from EtOAc-hexanes (128.3 mg, 54.7%): mp 290°C dec.; ν_{max} (KBr) 3400 (N-H str), 1693 (C=O str), 1610, and 1458 cm^{-1} (deprot acid C=O str); λ_{max} (H_2O) 248 (6000 $M^{-1}cm^{-1}$) and 322 nm (2304); δ_H (300 MHz, $CDCl_3$) 2.80 (3H, s), 7.78 (1H, dt, J =7.7 and 1.3 Hz), 7.88 (1H, dt, J =7.6 and 1.4 Hz), 8.21 (1H, d, J =8.8 Hz), 8.49 (1H, s), 8.56 (1H, d, J = 8.4 Hz).

Methyl 4-(1-hydroxyethyl)quinoline-2-carboxylate

$NaBH_4$ (0.95 g, 25 mmol) was added, as a powder, to a stirred solution of methyl 4-acetylquinoline-2-carboxylate (5.7 g, 25 mmol) in dry MeOH (125 ml) at room temperature. After 20 minutes the excess borohydride was destroyed by the addition of saturated aqueous NH_4Cl solution. The resulting solution was extracted with CH_2Cl_2. The organic extracts were pooled, dried (anhydrous Na_2SO_4), filtered, and the solvent removed *in vacuo* to give a pale yellow oil. The racemic alcohol was then obtained by chromatography on silica gel, eluting with acetone:hexanes (1:1), as a white powder. This reaction was repeated four times, and a fifth time with 4.98 g of reactant (reagents scaled accordingly) yielding an off white powder. (13.04 g, 47.8%): δ_H (200 MHz, $CDCl_3$) 1.51 (3H, d, J = 6.5 Hz), 3.87 (3H, s), 4.06 (1H, br s), 5.54 (1H, q, J = 6.5 Hz), 7.45 (1H, ddd, J = 8.3, 6.9, and 1.3 Hz), 7.58 (1H, ddd, J = 8.4,

6.9, and 1.3 Hz), 7.87 (1H, dd, J = 8.4 and 0.8 Hz), 8.09 (1H, s), 8.11 (1H, d, J = 8.6 Hz); δ_C (75.5 MHz, $CDCl_3$) 24.28, 52.82, 63.91, 116.5, 122.8, 128.2, 129.6, 130.9, 147.2, 147, 153.4, and 165.6.

rac-*4-(1-Hydroxyethyl)quinoline-2-carboxylic acid*

Methyl 4-(1-hydroxyethyl)quinoline-2-carboxylate (562 mg, 2.0 mmol) was dissolved in THF (8 ml). Aqueous 30% (w/v) NaOH solution (1.7 ml) was added and the two phase system stirred at room temperature. After 2 hours the white precipitate was collected by filtration, washed with THF, and recrystallized from EtOH to give the sodium salt of the acid as a white powder (345 mg, 76.6 %): mp 182 °C (evacuated); ν_{max} (nujol) 1610 and 1400 cm^{-1} (deprot acid C=O str); λ_{max} (MeOH) 242 (6430 $M^{-1}cm^{-1}$), 286 (2960), 310 (2200), and 322 nm (1750); δ_H (300 MHz, $CDCl_3$) 1.68 (3H, d, J = 6.6 Hz), 5.70 (1H, q, J = 6.5 Hz), 7.71 (1H, ddd, J = 7.8, 7.6, and 1.3 Hz), 7.82 (1H, ddd, J = 8.4, 7.0, and 1.3 Hz), 8.14 (1H, d, J = 8.7 Hz), 8.18 (1H, d, J = 8.0 Hz), 8.48 (1H, s); δ_C (75.5 MHz, DMSO-d_6) 24.7, 66.8, 117.8, 124.4, 127.1, 128.1, 130.3, 131.2, 148.6, 154.4, 157.0, and 173.4.

Methyl rac-*4-[1-(butanoyloxy)ethyl]quinoline-2-carboxylate*

Dry pyridine (6.3 ml, 45 mmol) and freshly distilled n-PrCOCl (7.9 ml, 45 mmol) were added to a stirred solution of methyl 4-(1-hydroxyethyl)quinoline-2-carboxylate (10.5 g, 45 mmol) in dry CH_2Cl_2 (200 ml). The solution was stirred at room temperature for 3 hours. A white precipitate formed. The suspension was diluted with CH_2Cl_2 (200 ml) and then washed, successively, twice with saturated aqueous $CuSO_4$ solution, once

with water and once with saturated aqueous $NaHCO_3$ solution. The organic phase was dried over anhydrous Na_2SO_4, filtered, and evaporated *in vacuo* to give a brown oil. The butyrate ester was obtained by chromatography on silica gel, eluting with EtOAc-hexanes (1:1), as a white crystalline solid (4.7 g, 34.4%): mp (from EtOAc-hexanes) 68-69 °C; Found: C, 67.7; H, 6.3; N, 4.7%. $C_{17}H_{19}NO_4$ requires: C, 67.8; H, 6.3; N, 4.65%; M+H, 302.1395; requires 302.1392; ν_{max} (CH_2Cl_2 solution) 2962 (C-H str), 1734 (C=O str), 1731 cm^{-1} (C=O str): λ_{max} (MeOH) 246 (9066 $M^{-1}cm^{-1}$) and 294 nm (5420); δ_H (300 MHz, $CDCl_3$) 0.82 (3H, t, J = 7.4 Hz), 1.57 (3H, d, J = 6.6 Hz), 1.57 (2H, sext, J = 7.4 Hz), 2.29 (2H, t, J = 7.4 Hz), 3.95 (3H, s), 6.50 (1H, q, J = 6.6 Hz), 7.54 (1H, ddd, J = 8.4, 6.9, and 1.2 Hz), 7.65 (1H, ddd, J = 8.4, 7.0, and 1.3 Hz), 7.96 (1H, d, J = 8.3 Hz), 8.15 (1H, s), 8.21 (1H, d, J = 8.1 Hz): δ_C (75.5 MHz, $CDCl_3$) 13.31, 18.21, 21.45, 35.99, 52.88, 60.00, 67.52, 116.9, 122.5, 125.9, 128.6, 129.7, 131.4, 147.6, 148.8, 165.6, and 172.3.

Methyl 4-[1-(benzoyloxy)ethyl]quinoline-2-carboxylate

DEAD (34.5 μl, 0.22 mmol) was added dropwise over 2 min to a solution of *rac*-4-(1-hydroxyethyl)quinoline-2-carboxylic acid (25.2 mg, 0.11 mmol), PhCOOH (30.6 mg, 0.25 mmol), and Ph_3P (60.1 mg, 0.23 mmol) in dry CH_2Cl_2 (1.0 ml). The resulting colorless solution was stirred at room temperature for 1.5 hours before being diluted with CH_2Cl_2 (2 ml) and washed with saturated aqueous $NaHCO_3$ solution. The organic phase was dried over anhydrous Na_2SO_4, filtered, and evaporated *in vacuo* to give a colorless oil. The product was obtained by chromatography on silica gel, eluting with EtOAc-hexanes (1:1), as a white crystalline compound (36.2 mg, 99.1 %):

mp 114-118°C; ν_{max} (CH_2Cl_2) 1720 cm^{-1} (C=O str); M+H, 336.1219; $C_{20}H_{17}NO_4$+H requires 336.1236; λ_{max} (MeOH) 244 (18600 $M^{-1}cm^{-1}$) and 294 nm (4500); δ_H (200 MHz, $CDCl_3$) 1.83 (1H, d, J = 6.6 Hz), 4.06 (3H, s), 6.85 (1H, q, J = 6.5 Hz), 7.46 (1H, dt, J = 6.8 and 1.5 Hz), 7.57 (2H, m), 7.70 (1H, dt, J = 6.9 and 1.4 Hz), 7.80 (1H, dt, J = 6.9 and 1.5 Hz), 8.11 (2H, dd, J = 7.8 and 1.6 Hz), 8.18 (1H, d, J = 8.2 Hz), 8.34 (1H, d, J = 8.4 Hz), 8.36 (1H, s); δ_C (75.5 MHz, $CDCl_3$) 21.8, 53.2, 68.7, 100.3, 117.3, 122.8, 126.3, 128.1, 128.4, 128.5, 129.1, 129.68, 129.73, 130.2, 131.6, 133.4, 147.7, 149.1, 165.6, and 165.8.

Methyl rac-4-[1-(acetoxy)ethyl]quinoline-2-carboxylate

Prepared as for methyl *rac*-4-[1-(butanoyloxy)ethyl]quinoline-2-carboxylate from acetyl chloride (0.35 ml, 5 mmol) and methyl 4-(1-hydroxyethyl)quinoline-2-carboxylate (0.46 g, 2.0 mmol). Yield 461.1 mg, 84.5 %: mp (from CH_2Cl_2) 103-104 °C: ν_{max} (CH_2Cl_2 solution) 1740 (C=O) and 1737 cm^{-1} (C=O); λ_{max} (MeOH, nm) 220 (4190 $M^{-1}cm^{-1}$), 248 (6200), 294 nm (4140); δ_H (300 MHz, $CDCl_3$) 1.66 (3H, d, J = 6.6 Hz), 2.14 (3H, s), 4.05 (3H, s), 6.58 (1H, q, J = 6.6 Hz), 7.65 (1H, ddd, J = 8.2, 7.1, and 1.2 Hz), 7.75 (1H, dt, J = 7.7 and 1.3 Hz), 8.05 (1H, d, J = 8.4 Hz), 8.23 (1H, s), 8.30 (1H, d, J = 8.3 Hz); δ_C (75.5 MHz, $CDCl_3$) 21.13, 21.71, 53.20, 67.99, 117.1, 122.7, 126.1, 128.8, 130.0, 131.7, 147.8, 147.8, 148.8, 166.0, and 170.0; m/z (Direct Probe EI-MS) 273 (19%, M^+), 231 (28), 215 (100), 154 (60), 128 (42).

Methyl (R)-4-(1-hydroxyethyl)quinoline-2-carboxylate

Candida cylindracea lipase (Sigma, EC 3.1.1.3, 1.16 mg, 1044 kU) was added to a stirred emulsion of methyl *rac*-4-[1-(butanoyloxy)ethyl]quinoline-2-carboxylate (5.79 g, 19.2 mmol) in 100 mM potassium phosphate buffer, pH 7.0 (100 ml) and DMSO (9 ml). The resulting emulsion was stirred at 50 °C for 3 days with frequent shaking being required to disperse any lumps that formed. The incubation mixture was then diluted with brine and extracted with CH_2Cl_2. The organic phase was dried over anhydrous Na_2SO_4, filtered, and the solvent removed by evaporation *in vacuo* to give a colorless oil. The starting material (4.65 g, 80.3%) and the product (200 mg, 0.82 mmol, 4.3%) were seperated by chromatography on silica gel, eluting with EtOAc-hexanes (1:1). The *R* isomer was identical to the previously prepared racemic compound. The enantiomeric purity of the product was determined by ^{1}H NMR analysis of its (1S)-camphanate derivative (de of 90%).

(1S)-Camphanate ester of methyl (12R)-4-(1-hydroxyethyl)quinoline-2-carboxylate

(1*S*)-(-)-Camphanic acid chloride (500 mg, 2.31 mmol) and dry pyridine (0.185 ml, 2.29 mmol) were added to a stirred solution of methyl (12*R*)-4-(1-hydroxyethyl)-quinoline-2-carboxylate (200 mg, 0.82 mmol) in dry $CHCl_3$ (4.0 ml). The solution was stirred at room temperature for 20 hours and then diluted with $CHCl_3$. The organic phase was washed successively, twice with saturated aqueous $CuSO_4$ solution, twice with 1 N HCl, and twice with saturated aqueous $NaHCO_3$ solution. The organic phase was then dried over anhydrous Na_2SO_4, filtered, and evaporated *in vacuo* to give a pale yellow oil. TLC analysis showed that conversion to the ester

had occurred quantitatively. Purification by chromatography on silica gel, eluting with EtOAc-hexanes (1:1), gave a pale yellow oil (142 mg, 41.9%): Recrystallization from EtOAc:hexanes yielded a white solid that showed none of the undesired diastereomer by ^{1}H NMR (92 mg, 27.1%): mp (from EtOAc: hexanes) 112-113.5 °C; found C, 67.0; H, 6.05; N, 3.45%, $C_{23}H_{25}NO_6$ requires C, 67.1; H, 6.1; N, 3.4%; M+H, 412.1755, requires 412.1760; ν_{max} (neat) 3059 (C-H str), 2971 (C-H str), 1790, 1746, 1724 cm^{-1} (C=O str); λ_{max} (MeOH) 246 (10900 M^{-1}cm^{-1}) and 296 nm (4790); m/z (FAB/$CHCl_3$-NBA) 412 (15%, M+H), 329 (5), 307 (20), 289 (10), 214 (7), 176 (12), δ_H (300 MHz, $CDCl_3$) 0.90 (3H, s), 1.03 (3H, s), 1.10 (3H, s), 1.70 (1H, m), 1.77 (3H d, J = 6.7 Hz), 1.90 (1H, m), 2.10 (1H, m), 2.45 (1H, m), 4.06 (3H, s), 6.72 (1H, q, J = 6.6 Hz), 7.70 (1H, ddd, J 8.4, 7.4, and 1.1 Hz), 7.80 (1H, dt, J = 7.2 and 1.1 Hz), 8.10 (1H, d, J = 8.4 Hz), 8.25 (1H, s), 8.37 (1H, d, J = 8.4 Hz); δ_C (75.5 MHz, $CDCl_3$) 9.63, 16.76, 16.83, 21.71, 28.99, 30.93, 53.21, 54.90, 69.80, 90.83, 117.5, 122.7, 129.1, 130.3, 131.8, 147.9, 165.7, 166.9, and 177.9.

(12R)-4-(1-Hydroxyethyl)quinoline-2-carboxylic acid

The (1S)-camphanate ester of methyl (12R)-4-(1-hydroxyethyl)-quinoline-2-carboxylic acid (4.8 mg, 11.8 mmol, > 98% de) prepared above was dissolved in THF (0.33 ml). Aqueous 30% NaOH solution (0.17 ml) was added and the 2-phase system stirred at 28°C for 2.5 h. The reaction mixture was neutralized and the precipitate dissolved in dilute HCl. The solution was made up to exactly 5.0 ml and then stored frozen prior to use in the enzymatic studies.

Methyl 4-hydroxymethylquinoline-2-carboxylate

TFA (1.8 ml, 24 mmol) was added carefully to a solution of $FeSO_4.7H_2O$ (0.67 g, 2.4 mmol) and methyl 2-quinoline-carboxylate (4.5 g, 24 mmol) in MeOH (150 ml) and the solution brought to reflux. Aqueous H_2O_2 (11.1 ml, 48.4 mmol) was carefully added over a period of 30 minutes to give a pale brown solution which was heated at reflux for a further 15 - 18 hours. The solution was left to cool to room temperature and then concentrated *in vacuo*. The residue (ca. 50 ml) was diluted with water (250 ml) and extracted with $CHCl_3$. The organic extracts were pooled, dried (anhydrous Na_2SO_4), filtered, and the solvent evaporated *in vacuo* to give a dark brown oil. The product was obtained by chromatography on silica gel, eluting with EtOAc-hexanes (4:1), as a pale yellow oil (1.97 g, 52.8% accounting for recovered starting material (1.29 g)): mp (from hexanes: EtOAc) 128-129.5 °C; M+H, 218.0821, $C_{12}H_{11}NO_3$+H requires, 218.0817; ν_{max} (CH_2Cl_2 solution) 3168 (O-H str), 2993 (C-H str), and 1731 cm^{-1} (C=O str); λ_{max} (MeOH) 244 (12050 $M^{-1}cm^{-1}$) and 292 nm (4680); δ_H (300 MHz, $CDCl_3$) 4.07 (3H, s), 5.25 (2H, s), 7.65 (1H, ddd, J = 8.4, 6.9, and 1.3 Hz), 7.77 (1H, ddd, J = 8.6, 6.8, and 1.4 Hz), 8.0 (1H, dd, J = 8.5 and 1.0 Hz), 8.29 (1H, s), and 8.32 (1H, d, J = 8.1 Hz); δ_C (75.5 MHz, $CDCl_3$) 14.2, 21.0, 29.7, 53.2, 60.2, 118.1, 122.8, 126.9, 128.8, 130.1, 165.9, 171.2, and 192.4; m/z (FAB, $CHCl_3$-NBA) 218 (M+H, 100%), 176 (66), and 165 (25).

4-Hydroxymethylquinoline-2-carboxylic acid

Aqueous NaOH solution (30%, 1.33 ml, 10 mmol) was added to a stirred solution of methyl 4-hydroxymethylquinoline-2-carboxylate (0.57 g, 2.63 mmol) in THF (3.67

ml) at room temperature. A yellow precipitate formed immediately. After 10 minutes the product was recovered by filtration and washed with THF (0.52 g, 87.9%): mp 240°C dec.; ν_{max} (KBr) 3421 (N-H str), 1617, and 1458 cm^{-1} (deprot acid C=O str); λ_{max} (H_2O) 242 (5670 $M^{-1}cm^{-1}$) and 292 nm (1620); δ_H (200 MHz, D_2O) 5.00 (2H, s), 7.49 (1H, ddd, J = 8.4, 6.9, and 1.2 Hz), 7.65 (1H, ddd, J = 8.5, 6.9, and 1.4 Hz), 7.79 (1H, s), 7.81 (1H, d, J = 7.0 Hz), and 7.92 (1H, d, J = 8.2 Hz); δ_C (75.5 MHz, D_2O) 61.7, 118.8, 124.3, 127.1, 129.0, 129.5, 130.1, 131.4, 147.3, 149.5, and 155.6.

Methyl 4-formylquinoline-2-carboxylate

Dry DMSO (0.118 ml, 1.66 mmol) was added to a stirred solution of $(COCl)_2$ (0.07 ml, 0.8 mmol) in CH_2Cl_2 (5.1 ml) at -78°C under an argon atmosphere. After 5 minutes a solution of methyl 4-hydroxymethylquinoline-2-carboxylate (150 mg, 0.69 mmol) in DMSO:CH_2Cl_2 (1:2, 1.4 ml) was added dropwise. After another fifteen minutes triethylamine (0.484 ml, 3.5 mmol) was added. The mixture was stirred for another 15 minutes at -78°C, then the flask was allowed to warm to room temperature for 90 minutes with continued stirring. The reaction mixture was diluted with H_2O and extracted with CH_2Cl_2. The organic extracts were pooled, dried (anhydrous Na_2SO_4), filtered, and the solvent evaporated *in vacuo* to yield a brown solid which was recrystallized using EtOAc/hexanes (90 mg, 0.42 mmol, 61%): mp 120-121°C; ν_{max} (KBr) 2848 (aldehyde C-H str), 1725 (ester C=O str), and 1695 cm^{-1} (aldehyde C=O str); λ_{max} (CH_3OH) 246 (6075 $M^{-1}cm^{-1}$) and 298 nm (4732); δ_H (300 MHz, $CDCl_3$) 4.12 (3H, s), 7.86 (2H, m), 8.47 (1H, d, J = 8.0 Hz), 8.59 (1H, s), 9.08 (1H, d, J = 7.8 Hz), 10.53 (1H, s); δ_C (300 MHz, DMSO-d_6) 53.1, 124.4, 124.5, 15.0, 130.7,

131.2, 131.5, 137.9, 148.1, 137.9, 148.1, 148.2, 164.8, 194.3.

5.2 Microbiology

Growth of S. laurentii

S. laurentii ATCC 31255 was stored as a spore suspension in 20% glycerol at -78°C. Fermentations were carried out in a two step process. Vegetative cultures (100 ml, 15 g/l Trypticase Soy Broth, 15 g/l soluble starch, 50 g/l sucrose in distilled water) were inoculated with 100 μl of spore supsension and grown for 48 hours at 28°C with shaking at 200 rpm (ISF-4V shaker cabinet, A. Kühner) in 500 ml baffled Erlenmeyer flasks. Production cultures (100 ml, 15 g/l trypticase soy broth, 15 g/l $CaSO_4$, 11 g/l yeast extract, 50 g/l glucose, 1 ml/l $2x$ trace elements solution [23]) were inoculated with 10 ml of vegetative culture and grown for 72 hours in 500 ml baffled Erlenmeyer flasks at 28°C with shaking at 200 rpm. For growth of *S. laurentii* to be used for genomic DNA isolation YMB (YEME without sucrose [23]) medium was used.

Spores were harvested by the Hopwood protocol [23] from petri dishes containing R6 medium (R2YE, alternative method, without sucrose [23]) that had been inoculated with 100 μl of spore suspension and grown for 120 hours at 30°C. Selective growth was performed with the following antibiotic concentrations: Apramycin sulfate (100 μg/ml), thiostrepton (50 μg/ml), and nalidixic acid (100 μg/ml).

Assay for thiostrepton production

Thiostrepton production was assayed by extracting 2 ml production culture with 2 ml chloroform. The layers were separated, and one ml of the chloroform diluted with one

ml of glacial acetic acid. 15 μl was analyzed by reverse phase HPLC (Hamilton PRP-1), run isocratically with 60% aqueous acetonitrile, and compared with thiostrepton standards. The peak believed to be due to thiostrepton was confirmed by collecting the peak onto a circular piece of filter paper (Whatcom), drying the disk, and using it to test growth inhibition. The disks were placed onto LB medium, and a soft agar overlay containing *Staph. aureus*, *S. lividans*, or *S. laurentii* applied to the plate [59]. These cultures were then incubated for 48h at 30°C, and examined for clear zones indicating antimicrobial activity.

Growth of E. coli

E. coli cultures were grown in or on LB medium [59] at 37°C for 12-16 hours. Liquid cultures were shaken at 300 rpm (Model G25 incubator/shaker, New Brunswick, Inc.) in unbaffled Erlenmeyer flasks or test tubes. For selective growth antibiotics were used in the following concentrations: Carbenicillin (100 μg/ml), apramycin sulfate (100 μg/ml), chloramphenicol (50 μg/ml), nalidixic acid (100 μg/ml) and tetracycline (10 μg/ml).

Conjugation

E. coli S17-1 was transformed with pKC1132 containing the fragment of DNA. *E. coli* cells that had been transformed were grown overnight in LB liquid medium. Fresh *S. laurentii* spore suspension (100 μl) was washed once with water and resuspended in 500 μl 0.05 M TES pH 7.2 and 500 μl $2x$ pregermination medium [23]. These spores were then incubated for 4 hours at 28 °C. Overnight S17-1 culture (100 μl) containing

plasmid was mixed with 100 μl of the pregerminated spores, and the mixture spread onto AS-1 solid medium (1 g/l yeast extract, 0.2 g/l L-alanine, 0.2 g/l L-arginine, 0.5 g/l L-asparagine, 5 g/l soluble potato starch, 2.5 g/l NaCl, 10 g/l Na_2SO_4, 22 g/l agar, pH 7.5). These cultures were incubated overnight at 28°C. The *E. coli* cells were partially removed by washing with sterile distilled water, then the plates were overlayed with 3 ml AS-1 soft agar containing 1.5 mg nalidixic acid to kill all *E. coli* cells, and 1.5 mg apramycin sulfate to select for *S. laurentii* cells containing inserted DNA, and incubated at 28°C. Colonies were evident after 2-3 days. Observed colonies were transferred to R6 medium containing 100 μg/ml apramycin sulfate after one week.

5.3 Protein purification

General

Amino acids and ATP were obtained from Sigma. Buffer salts and glycerol were obtained from Baker, Inc. All buffers and column materials were degassed *in vacuo* prior to use. Buffer A: 50 mM KH_2PO_4, pH 6.8, 1 mM DTT; Buffer B: 100 mM TRIS, pH 8.0, 1 mM DTT; Buffer C: 50 mM KH_2PO_4, pH 5.8, 10% glycerol, 1 mM DTT. Centrifugation was performed using either a Sorvall RC-5C centrifuge or a benchtop microcentrifuge (Sorvall 12S). Ammonium sulfate fractionation, column equilibration, SDS-PAGE analysis, and transfer of proteins to PVDF membranes were performed as described by Doonan *et al.* [15]. Protein purification resins were obtained and used as described in table 5.2.

Table 5.1: Bacterial strains used

Strain	Relevant features	Source or reference
E. coli SURE	Cosmid cloning host	Stratagene
E. coli XL-1 Blue	Cloning host	Stratagene
E. coli XL10-Gold Kan	Cloning host	Stratagene
E. coli S17-1	Conjugation host	Simon *et al.* [62]
Staphylococcus aureus	Thiostrepton sensitive species	ATCC 35556
Streptomyces actuosus	Nosiheptide producer	ATCC 25421
S. lividans TK24	Thiostrepton sensitive Streptomycete	Hopwood *et al.* [23]
S. laurentii	Thiostrepton producer	ATCC 31255
S. laurentii TS203	pKC1132 insertion mutant (Am^R)	Smith [64]
S. laurentii TS207	pKC1132 insertion mutant (Am^R)	Smith [64]
S. laurentii TS208	pKC1132 insertion mutant (Am^R)	Smith [64]
S. laurentii TS209	pKC1132 insertion mutant (Am^R)	Smith [64]
S. laurentii TS210	pKC1132 insertion mutant (Am^R)	Smith [64]
S. laurentii TS211	pKC1132 insertion mutant (Am^R)	Smith [64]
S. laurentii TS212	pKC1132 insertion mutant (Am^R)	Smith [64]
S. laurentii TS213	pKC1132 insertion mutant (Am^R)	Smith [64]
S. laurentii TS214	pKC1132 insertion mutant (Am^R)	Smith [64]
S. laurentii TS215	pKC1132 insertion mutant (Am^R)	Smith [64]
S. laurentii TS216	pKC1132 insertion mutant (Am^R)	Smith [64]
S. laurentii TS217	pKC1132 insertion mutant (Am^R)	Smith [64]
S. laurentii TS218	pKC1132 insertion mutant (Am^R)	Smith [64]
S. laurentii TS219	pKC1132 insertion mutant (Am^R)	Smith [64]
S. laurentii pSJF202'	pKC1132 insertion mutant (Am^R)	page 94
S. laurentii pSJF203'	pKC1132 insertion mutant (Am^R)	page 94
S. laurentii pSJF204'	pKC1132 insertion mutant (Am^R)	page 94
S. laurentii pSJF205'	pKC1132 insertion mutant (Am^R)	page 94
S. laurentii pSJF206'	pKC1132 insertion mutant (Am^R)	page 94
S. laurentii pSJF207'	pKC1132 insertion mutant (Am^R)	page 94
S. laurentii pSJF209'	pKC1132 insertion mutant (Am^R)	page 94

Preparation of cell-free extract

Two liters of production medium were harvested by centrifugation, and washed twice with buffer A containing 20% glycerol. PMSF was added to 1 mM, and the cells lysed mechanically using a French pressure system (AMINCO, $2x$ 14,000-16,000 psi). The cell debris was removed by centrifugation. PEI was added to 0.3% to precipitate the DNA from the sample, and the suspension separated by centrifugation, yielding the cell-free extract.

5.4 *Molecular biology*

General

Restriction endonucleases were purchased from Gibco BRL, Boehringer Mannheim, Stratagene, Inc., or Promega and used according to the manufacturers' guidelines. Ligation of sticky ended DNA fragments was achieved using shrimp alkaline phosphatase (Boehringer Mannheim) to dephosphorylate the vector, and T4 ligase (Gibco BRL) to ligate the insert, according to the manufacturers' recommendations. The pPCR Script cloning kit (Stratagene, Inc.) was used to clone PCR-generated blunt ended DNA fragments.

DNA isolation and purification

Plasmid DNA was isolated from *E. coli* cell cultures for sequencing using either Invitrogen or Qiagen DNA purification columns according to manufacturers' recommendations. The "standard" miniprep procedure [59] was used for other procedures.

Table 5.2: Protein purification resins

Resin	Characteristics	Buffer	Volume	Source
Mono S	Methyl sulphonate strong cation exchanger	A	1 ml	AP Biotech
Mono Q	Quaternary ammonium strong anion exchanger	B	1 ml	AP Biotech
Phenyl Superose	Hydrophobic interaction chromatographic resin	A + 30% ammonium sulfate	1 ml	AP Biotech
DEAE Cellulose	Anion exchange resin	A + 20% glycerol	35 ml	Sigma
DEAE Sephacel	Anion exchange resin	A + 20% glycerol	20 ml	Sigma
CM Sephadex	Cation exchange resin	C	5 ml	Sigma
Phenyl Sepharose	High substitution fast flow hydrophobic interaction chromatography resin	A + 1 M ammonium sulfate	10 ml	AP Biotech
SP Sepharose	Sulphopropyl strong cation exchange resin	A + 10% ammonium sulfate	5 ml	AP Biotech
ATP affinity C-8	ATP affinity resin with 9 atom spacer attached via C-8	A	1ml	Sigma
ATP affinity N-6	ATP affinity resin with 11 atom spacer attached via N-6	A	200 μl	Sigma
ATP affinity ribose	ATP affinity resin with 9 atom spacer attached via the ribose hydroxyls	A	200 μl	Sigma
Affi-Gel Blue	Cibachrome Blue dye affinity resin	A	200 μl	Bio-Rad

Genomic DNA was isolated from *Streptomyces* species using the rapid small scale preparation described by Hopwood *et al.* [23], and collected by spooling on a pasteur pipette.

Southern hybridization

Electrophoresis, transfer to Nytran membranes, and Southern hybridization were performed as described by Maniatis *et al.* [59]. Probes were synthesized either by generating randomly primed and labeled fragments using the RadPrime labeling kit (Gibco BRL) and [α^{32}P]dCTP (NEN-Dupont), or end labeled using the RTS T4 kinase end labeling kit (Gibco BRL) and [γ^{32}P]dATP.

PCR conditions

DNA polymerases used in PCR amplification studies were obtained from Stratagene, Inc., Gibco BRL, or Boehringer Mannheim, as noted in the text, and used according to manufacturers' guidelines with the noted modifications (table 5.3). PCR reactions were performed using an MJ Research model PTC-200 thermocycler.

DNA sequencing

Manual sequencing reactions were performed using the Sequenase kit (US Biochemical) and [$\alpha-^{35}$S]dCTP (NEN-Dupont) according to manufacturer's recommendations. Sequencing reactions were analyzed on gels of 6% polyacrylamide, 7M urea, TBE [59] in a Poker Face gel electrophoresis apparatus (VWR).

PCR cycle sequencing was performed by BigDye (Applied Biosystems, Inc.) chem-

istry using the PCR cycles recommended by the manufacturer, and analyzed with ABI DNA sequencers. Sequence data were read, edited, and analyzed using Sequencher DNA analysis software (Gene Codes, Inc.), and compared to database sequences using the NCBI BLAST server [1].

Table 5.3: PCR conditions for Taq DNA polymerases and Pfu DNA polymerases. T_A is the annealing temperature, which was varied from 50°C to 68°C. Denaturing temperature was generally higher than manufacturers' recommendations due to the high GC content of templates and primers

Step	*Taq* DNA Polymerases	*Pfu* DNA Polymerases
1	96°C 3'	98°C 45"
2	T_A 30"	97°C 45"
3	72°C 1.5'	T_A 45"
4	96°C 45"	72°C 2'
5	Repeat 2-4 29 times	
6	55°C 30"	72°C 10'
7	72°C 30'	4°C $\rightarrow \infty$
8	4°C $\rightarrow \infty$	

Table 5.4: Vectors and clones

Vector or clone	Relevant characteristics	Source or reference
pBluescript II KS$^-$	*E. coli* cloning vector for plasmid libraries and sequencing, ampicillin resistance gene, shortened to pKS$^-$ for brevity	Stratagene
pBC KS$^-$	*E. coli* cloning vector for plasmid libraries and sequencing, chloramphenicol resistance gene	Stratagene
pPCR-Script Amp	Plasmid used to clone blunt end PCR products	Stratagene
SuperCos	*E. coli* cosmid vector	Stratagene
pOJ446	*E. coli*-Streptomycete shuttle cosmid vector	Bierman *et al.* [5]
pKC1132	*E. coli* vector, containing *ori*T, used for integrational mutations	Bierman *et al.* [5]
pSJF 14.3	cosmid clone in pOJ446	page 91
pSJF 17.2	cosmid clone in pOJ446	page 91
pSJF 18.1	cosmid clone in pOJ446	page 91
pSJF 18.3	cosmid clone in pOJ446	page 91
pSJF 20.2	cosmid clone in pOJ446	page 91
Supercos 5.2	cosmid clone in SuperCos	page 91
Supercos 9.1	cosmid clone in SuperCos	page 91
Supercos 9.7	cosmid clone in SuperCos	page 91
Supercos 9.4	cosmid clone in SuperCos	page 91
Supercos 3.1	cosmid clone in SuperCos	page 91
Supercos 5.3	cosmid clone in SuperCos	page 91
Supercos 4.1	cosmid clone in SuperCos	page 91
pEPSY K-G7	cosmid clone in SuperCos	page 109
pEPSY E-05	cosmid clone in SuperCos	page 109
pEPSY T-B6	cosmid clone in SuperCos	page 109
pEPSY G-D2	cosmid clone in SuperCos	page 109
pEPSY A-A10	cosmid clone in SuperCos	page 109
pEPSY G-F2	cosmid clone in SuperCos	page 109
pEPSY L-F8	cosmid clone in SuperCos	page 109
pEPSY Q-B12	cosmid clone in SuperCos	page 110
pEPSY N-A11	cosmid clone in SuperCos	page 110
pEPSY L-F8	cosmid clone in SuperCos	page 110
pEPSY N-D1	cosmid clone in SuperCos	page 110
pEPSY P-E3	cosmid clone in SuperCos	page 110
pEPSY R-D7	cosmid clone in SuperCos	page 110

Table 5.5: Subclones

Plasmid	Parent	Size (kb)	Enzyme(s)	Vector	Notes
pSJF 103	pSJF 17.2	5.5	*Bam*HI	pKS$^-$	
pSJF 105	pSJF 17.2	2.0	*Bam*HI	pKS$^-$	
pSJF 107	pSJF 17.2	0.9	*Bam*HI	pKS$^-$	
pSJF 108	pSJF 17.2	0.7	*Bam*HI	pKS$^-$	
pSJF 111	pSJF 20.2	3.1	*Bam*HI	pKS$^-$	
pSJF 112	pSJF 20.2	1.4	*Bam*HI	pKS$^-$	
pSJF 113	pSJF 20.2	0.8	*Bam*HI	pKS$^-$	
pSJF 114	pSJF 18.1	23	*Bam*HI	pKS$^-$	
pSJF 116	pSJF 18.1	6.5	*Bam*HI	pKS$^-$	
pSJF 117	pSJF 18.1	2.5	*Bam*HI	pKS$^-$	
pSJF 118	pSJF 18.1	2.5	*Bam*HI	pKS$^-$	
pSJF 119	pSJF 18.1	1.0	*Bam*HI	pKS$^-$	
pSJF 121	pSJF 18.3	12	*Bam*HI	pKS$^-$	
pSJF 202	pSJF 114	6.0	*Not*I	pBC KS$^-$	
pSJF 202'	pSJF 114	6.0	*Not*I	pKC1132	
pSJF 203	pSJF 114	3.2	*Not*I	pBC KS$^-$	
pSJF 203'	pSJF 114	3.2	*Not*I	pKC1132	
pSJF 204	pSJF 114	3.1	*Not*I	pBC KS$^-$	
pSJF 204'	pSJF 114	3.1	*Not*I	pKC1132	
pSJF 205	pSJF 114	2.2	*Not*I	pBC KS$^-$	
pSJF 205'	pSJF 114	2.2	*Not*I	pKC1132	
pSJF 206	pSJF 114	1.8	*Not*I	pBC KS$^-$	
pSJF 206'	pSJF 114	1.8	*Not*I	pKC1132	
pSJF 207	pSJF 114	1.2	*Not*I	pBC KS$^-$	
pSJF 207'	pSJF 114	1.2	*Not*I	pKC1132	
pSJF 208	pSJF 114	7.0	*Sac*I	pKS$^-$	Internal deletion
pSJF 209	pSJF 114	6.0	*Not*I	pKS$^-$	Internal deletion
pSJF 209'	pSJF 114	6.0	*Not*I/*Bam*HI	pKC1132	
pSJF 211	pSJF 114	0.5	*Kpn*I	pKS$^-$	Internal deletion
pSJF 301	pSJF 203	0.5	*Sma*I	pBC KS$^-$	Internal deletion
pSJF 302	pSJF 203	1.5	*Kpn*I	pBC KS$^-$	Internal deletion
pSJF 305	pSJF 204	1.2	*Sma*I	pKS$^-$	
pSJF 306	pSJF 204	1.0	*Sac*I	pKS$^-$	
pSJF 307	pSJF 204	1.8	*Sac*I	pKS$^-$	
pSJF 308	pSJF 204	1.0	*Sac*I	pBC KS$^-$	Internal deletion

Table 5.6: Manufacturers and suppliers

Aldrich	Milwaukee, WI 53201
Alltech (Hamilton)	Deerfield, IL 60015
AMINCO	Silver Spring, MD 20910
AP Biotech	Piscataway, NJ 08855
Applied Biosystems, Inc. (Perkin-Elmer)	Norwalk, CT 06859
ATCC	Rockville, MD 20852
Bio-Rad, Inc.	Hercules, California 94547
Boehringer Mannheim Biochemicals	Indianapolis, IN 46250
Gene Codes, Inc.	Ann Arbor, MI 48108
Gibco BRL	Rockville, MD 20849
Invitrogen	Carlsbad CA 92008
J.T. Baker	Phillipsburg, NJ 08865
Millipore	Bedford Massachusetts, 01730
MJ Research	Waltham, MA 02451
New Brunswick, Inc.	Edison, NJ
New England Biochemicals	Beverly, MA 01915
New England Nuclear (NEN) - Dupont	Boston, MA 02118
Promega	Madison, WI 53711
Qiagen	Valencia, CA 91355
Sigma	St. Louis, MO 63103
Sorvall (Dupont)	Newtown, CT 06470
Stratagene, Inc.	La Jolla, CA 92037
US Biochemicals	Piscataway, NJ 08855
VWR Scientific	West Chester, PA 19380
Whatman	Clifton, NJ 07014

BIBLIOGRAPHY

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

[2] B. Anderson, D. C. Hodgkin, and M. W. Viswamitra. The structure of thiostrepton. *Nature*, 225:233–255, 1970.

[3] M. Aoki, T. Ohtsuka, M. Yamada, Y. Ohba, H. Yoshizaki, H. Yasuno, T. Sano, J. Watanabe, K. Yokose, and H. Seto. Cyclothiazomycin, a novel polythiazole-containing peptide with renin inhibitory activity. Taxonomy, fermentation, isolation and physico-chemical characterization. *J. Antibiot.*, 44:582–588, 1991.

[4] F. Benazet, M. Cartier, J. Florent, C. Godard, G. Jung, J. Lunel, D. Mancy, C. Pascal, J. Renaut, P. Tarridec, J. Theilleux, R. Tissier, M. Dubost, and L. Ninet. Nosiheptide, a sulfur-containing peptide antibiotic isolated from *Streptomyces actuosus* 40037. *Experientia*, 36:414–416, 1980.

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

[6] M. Bodanszky, J. Alicino, C.A. Birkhimer, and N.J. Williams. Degradation of

thiostrepton. The structure of thiostreptine. *J. Am. Chem. Soc.*, 84:2003–2004, 1962.

[7] M. Bodanszky, J. Fried, J.T. Sheehan, N.J. Williams, J. Alicino, A.I. Cohen, B.T. Keeler, and C.A. Birkhimer. Thiostrepton. Degradation products and structural features. *J. Am. Chem. Soc.*, 86:2478–2484, 1964.

[8] M. Bodanszky, J.T. Sheehan, J. Fried, N. Williams, and C. Birkhimer. Degradation of thiostrepton. Thiostreptoic acid. *J. Am. Chem. Soc.*, 82:4747–4748, 1960.

[9] M.L. Chiu, M. Folcher, P. Griffin, T. Hold, T. Klatt, and C.J. Thompson. Characterization of the covalent binding of thiostrepton to a thiostrepton-induced protein from *Streptomyces lividans*. *Biochemistry*, 35:2332–2341, 1996.

[10] M.J. Cron, D.F. Whitehead, I.R. Hooper, B. Heineman, and J. Lein. Braymycin, a new antibiotic. *Antibiot. Chemother.*, 6:63–67, 1953.

[11] E. Cundliffe. How antiobiotic-producing microorganisms prevent suicide. *Annu. Rev. Microbiol.*, 43:207–233, 1989.

[12] E. Cundliffe and J. Thompson. Ribose methylation and resistance to thiostrepton. *Nature*, 278:859–861, 1979.

[13] Jr. J.W. Davis. Studies with quinolines. I. Synthesis of quinaldic acid and some of its amide derivatives. *J. Org. Chem.*, 24:1691, 1959.

[14] M. Debono, R.M. Molloy, J.L. Occolowitz, J.W. Paschal, A.H. Hunt, K.H. Michel, and J.W. Martin. The structures of A10255B, -G, and -J. New thiopeptide antibiotics produced by *Streptomyces gardneri*. *J. Org. Chem.*, 57:5200–5208, 1992.

[15] S. Doonan, editor. *Protein Purification Protocols*, volume 59 of *Methods in Molecular Biology*. Human Press, 1996.

[16] C.N.C. Drey, G.W. Kenner, H.D. Law, R.C. Sheppard, M. Bodanszky, J. Fried, N.J. Williams, and J.T. Sheehan. Degradation of thiostrepton. Derivatives of 8-hydroxyquinoline. *J. Am. Chem. Soc.*, 83:3906–3908, 1961.

[17] K.D. Entian and W.M. de Vos. Genetics of subtilin and nisin biosyntheses. *Antonie van Leeuwenhoek*. 69:109–138, 1996.

[18] T. Frenzel, P. Zhou, and H.G. Floss. Formation of 2-methyltryptophan in the biosynthesis of thiostrepton: Isolation of S-adenosyl-methionine:tryptophan 2-methyltransferase. *Arch. Biochem. Biophys.*, 278:35–40, 1990.

[19] S.K. Ghosh, N.K. Mukhopadhyay, S. Majumcer, and S.K. Bose. Purification of the constituent enzyme fractions of mycobacillin synthetase. *Biochem. J.*, 235:81–85, 1986.

[20] O. D. Hensens and G Albers-Schönberg. ^{13}C NMR study of thiostrepton and thiopeptin components. *J. Antibiot.*, 36:832–845, 1983.

[21] O.D. Hensens and G. Albers-Schönberg. Total structure of the highly modified peptide antibiotic components of thiopeptin. *J. Antibiot.*, 36:814–831, 1983.

[22] D.A. Hopwood, M.J. Bibb, K.F. Chater, and T. Kieser. Plasmid and phage vectors for gene cloning and analysis in Streptomyces. *Methods Enzymol.*, 153:116–166, 1987.

[23] D.A. Hopwood, M.J. Bibb, K.F. Chater, T. Kieser, C. Bruton, H.M. Kieser, D.J. Lydiate, S.P. Smith, J.M. Ward, and H. Schremph. *Genetic manipulation of Streptomyces, a laboratory manual.* The John Innes Foundation, Norwich, UK, 1985.

[24] S.M. Jeong, H. Yoshikawa, and H. Takahashi. Isolation and characterization of the *secE* homologue gene of *Bacillus subtilis*. *Mol. Microbiol.*, 10:133–142, 1993.

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

[26] U. Keller, H. Kleinkauf, and R. Zocher. 4-Methyl-3-hydroxyanthranilic acid activating enzyme from actinomycin-producing *Streptomyces chrysomallus*. *Biochemistry*, 23:1479–1484, 1984.

[27] J. Kettenring, L. Colombo, P. Ferrari, P. Taveccia, M. Nebuloni, K. V'ekey, G.G. Gallo, and E. Selva. Antibiotic GE2270 A: a novel inhibitor of bacterial protein synthesis. II. Structure elucidation. *J. Antibiot.*, 44:702–715, 1991.

[28] H. Kleinkauf and H. von Döhren. A nonribosomal system of peptide biosynthesis. *Eur. J. Biochem*, 236:335–351, 1996.

[29] R.H. Lambalot, A.M. Gehring, R.S. Flugel, P. Zuber, M. LaCelle, M.A. Marahiel, R. Reid, Khosla. C, and C.T. Walsh. A new enzyme superfamily - the phosphopantetheinyl transferases. *Chemistry and Biology*, 3:923–936, 1996.

[30] Y. Li, D.C. Dosch, R.H. Woodman, H.G. Floss, and W.R Strohl. Transcriptional organization and regulation of the nosiheptide resistance gene in *Streptomyces actuosus*. *Ind. Microbiol.*, 8:1, 1991.

[31] H. Malina and M. Robert-Gero. Characterization of an 8.7-kilobase thiostrepton resistance-encoding plasmid (pGIF3) of *Streptomyces incarnatus*. *App. Env. Microbiol.*, 58:895–9, 1992.

[32] A.J. Manusco, S.-L. Huang, and D. Swern. Oxidation of long-chain and related alcohols to carbonyls by dimethyl sulfoxide "activated" by oxalyl chloride. *J. Org. Chem.*, 43:2480, 1978.

[33] M.A. Marahiel, T. Stachelhaus, and H.D. Mootz. Modular peptide synthetases involved in nonribosomal peptide synthesis. *Chem. Rev.*, 97:2651–2673, 1997.

[34] P. Mazodier, R. Petter, and C. Thompson. Intergeneric conjugation between *Escherichia coli* and *Streptomyces* species. *J. Bacteriol.*, 171:3583–3585, 1989.

[35] G.A. McConkey, M.J. Rogers, and T.F. McCutchan. Inhibition of *Plasmodium*

falciparum protein synthesis: Targeting the plastid-like organelle with thiostrepton. *J. Biol. Chem.*, 272:2046–2049, 1997.

[36] R. McDaniel, A. Thamchaipenet, C. Gustafsson, H. Fu, M. Betlach, M. Betlach, and G. Ashley. Multiple genetic modifications of the erythromycin polyketide synthase to produce a library of novel "unnatural" natural products. *Proc. Natl. Acad. Sci. USA*, 96:1846–1851, 1999.

[37] F. Minisci, E. Vismara, and S. Levi. A general, selective and convenient procedure of homolytic formylation of heteroaromatic bases. *J. Org. Chem.*, 51:536, 1986.

[38] F. Minisci, F. Vismara, F. Fontana, and D. Redaelli. Polar effects in free-radical reactions. New catalytic processes in the substitution of heteroaromatic bases by nucleophilic free radicals. *Gazz. Chim. Ital.*, 117:363, 1987.

[39] U. Mocek, J.M. Beale, and H.G. Floss. Reexamination of the ^{1}H and ^{13}C NMR spectral assignments of thiostrepton. *J. Antibiot.*, 42:1649-1652, 1989.

[40] U. Mocek, Z. Zeng, D. O'Hagan, P. Zhou, L.D. Fan, J. Beale, and H.G. Floss. Biosynthesis of the modified peptide antibiotic thiostrepton in *Streptomyces azureus* and *Streptomyces laurentii*. *J. Am. Chem. Soc.*, 115:7992–8001, 1993.

[41] T. Nakamura, Y.and Gojobori and T. Ikemura. Codon usage tabulated from the international DNA sequence databases; its status 1999. *Nucl. Acids Res.*, 27:292, 1999.

[42] G.H. Nesbitt and P.R. Fox. Clinical evaluation of Panolog Cream used to treat canine and feline dermatoses. *Vet. Med. Small Anim. Clin.*, 76:535–538, 1981.

[43] H.C. Neu. The crisis in antibiotic resistance. *Science*, 257:1064–1073, 1992.

[44] P.T. Northcote, M. Siegel, D.B. Borders, and M.D. Lee. Glycothiohexide alpha, a novel antibiotic produced by *"Sebekia" sp.*, LL-14E605. III. Structural elucidation. *J. Antibiot.*, 29:901–908. 1976.

[45] K. Ochi. The tsr gene-coding plasmid pIJ702 prevents thiopeptin from inhibiting ppGpp synthesis in *Streptomyces lividans*. *FEMS Microb. Biol.*, 52(1-2):219–23. 1989.

[46] K. Ochi. A relaxed (rel) mutant of *Streptomyces coelicolor* A3(2) with a missing ribosomal protein lacks the ability to accumulate ppGpp, A-factor and prodigiosin. *J. Gen. Microbiol.*, 136(12):2405–12, 1990.

[47] K. Ochi. Streptomyces relC mutants with an altered ribosomal protein ST-L11 and genetic analysis of a *Streptomyces griseus* relC mutant. *J. Bacteriol.*, 172(7):4008–16, 1990.

[48] P. O'Farrell. Hydrophobic interaction chromatography. *Methods in Molecular Biology*, 59:151–155, 1996.

[49] S. O'Hogan. Characterization of a peptide synthetase gene cluster in the nosiheptide producer *Streptomyces actuosus*. Master's thesis, University of Washington, 1997.

[50] J.F. Pagano, M.J. Weinstein, H.A. Stout, and R. Donovick. Thiostrepton, a new antibiotic. I. *In vitro* studies. *Antibiot. Ann.*, pages 554–559, 1955/1956.

[51] F.D. Popp, W. Bount, and P.J. Melvin. Reissert compound studies. II. Nature of the quinoline. *J. Org. Chem.*, 26:4930, 1961.

[52] B.T. Porse, I. Leviev, A.S. Mankin, and R.A. Garrett. The antibiotic thiostrepton inhibits a functional transition within protein L11 at the ribosomal GTPase centre. *J. Mol. Biol.* 276:391–404, 1998.

[53] N.D. Priestley, T.M. Smith, P.R. Shipley., and H.G. Floss. Studies on the biosynthesis of thiostrepton: 4-(1-hydroxyethyl)quinoline-2-carboxylate as a free intermediate on the pathway to the quinaldic acid moiety. *Bioorg. Med. Chem.*, 4:1135–47, 1996.

[54] M.S. Puar, T.M. Chan, V. Hegde, M. Patel, P. Bartner, K.J. Ng, and B.N. Pramanik. Sch 40832: A novel thiostrepton from *Micromomspora carbonacea. J. Antibiot.*, 51:221–224, 1998.

[55] A. Reissert. Ueber die einführung der benzoyl gruppe in tertiäre cyclische basen. *Ber. Dtch. Chem. Ges.*, 38:1603, 1905.

[56] G. Rosendahl and S. Douthwaite. Ribosomal proteins L11 and L10.$(L12)_4$ and the antibiotic thiostrepton interact with overlapping regions of the rRNA backbone in the ribosomal GTPase centre. *J. Mol. Biol.*, 234:1013–1020, 1993.

[57] G. Rosendahl and S. Douthwaite. The antibiotics micrococcin and thiostrepton interact directly with 23S rRNA nucleotides 1067A and 1095A. *Nucl. Acids Res.*, 22:357–363, 1994.

[58] R.S. Roy, A.M. Gehring, J.C. Milne, P.J. Belshaw, and C.T. Walsh. Thiazole and oxazole peptides: biosynthesis and molecular machinery. *Nat. Prod. Rep.*, 16:249–263, 1999.

[59] J. Sambrook, E.F. Fritsch, and T. Maniatis. *Molecular cloning, a laboratory manual.* Cold Spring Harbor Laboratory Press, 1989.

[60] F.J. Schmidt, J. Thompson, K. Lee, J. Dijk, and E. Cundliffe. The binding site for ribosomal protein L11 within 23S ribosomal RNA of *Escherichia coli*. *J. Biol. Chem.*, 256(23):12301–5, 1981.

[61] J. Shoji, H. Hinoo, Y. Wakisaka, K. Koizumi, M. Mayama, S. Matsuura, and K. Matsumoto. Isolation of three new antibiotics, thiocillins I, II, and III, related to micrococcin P. *J. Antibiot.*, 29:366–374, 1976.

[62] R. Simon, U. Priefer, and A. Pühler. A broad host range mobilization for *in vivo* genetic engineering: Transposon mutagenesis in Gram positive bacteria. *Bio/Technology*, 1:784–791, 1983.

[63] T.M. Smith. Studies on thiostrepton biosynthesis. Unpublished work, 1992.

[64] T.M. Smith. *Genetic and Biochemical Studies of Thiostrepton Biosynthesis in* Streptomyces laurentii. PhD thesis, University of Washington, 1993.

[65] T.M. Smith, J.-F. Jiang, P. Shipley, and H.G. Floss. The thiostrepton resistance gene in *Streptomyces laurentii* is located within a cluster of ribosomal protein operons. *Gene*, 164:137–142, 1995.

[66] T. Stachelhaus, H.D. Mootz, and M. Marahiel. The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. *Chemistry and Biology*. 6:493–505, 1999.

[67] B.A. Steinberg, W.P. Jambor, and L.O. Snydam. Thiostrepton, a new antibiotic. III. *In vivo* studies. *Antibiot. Ann.*, pages 562–565, 1955/1956.

[68] S. Storkova. *Search for the nosiheptide biosynthetic gene cluster in* Streptomyces actuosus. PhD thesis, Academy of Sciences of the Czech Republic and University of Washington, 1997.

[69] J. Thompson and E. Cundliffe. Resistance to thiostrepton, siomycin, and sporangiomycin in actinomycetes that produce them. *J. Bacteriol.*, 142:281–291, 1980.

[70] K. Tori, K. Tokura, K. Okabe, M. Ebata, H. Otsuka, and G. Lukacs. Carbon-13 NMR studies of peptide antibiotics, thiostrepton and siomycin A: the structure relationship. *Tet. Lett.*, 3:185–188, 1976.

[71] W.H. Trejo, L.D. Dean, J. Plussic, E. Meyers, and W.E.J. Brown. *Strepomyces laurentii*, a new species producing thiostrepton. *J. Antibiotics (Tokyo)*, 30:369, 1977.

[72] P. Trieu-Cuot, C. Carlier, P. Martin, and P. Courvalin. Plasmid transfer by conjugation from *Escherichia coli* to Gram-positive bacteria. *FEMS Microb. Lett.*, 48:289–294, 1987.

[73] K. Turgay and M. Marahiel. A general approach for identifying and cloning peptide synthetase genes. *Peptide Research*, 7:238–242, 1994.

[74] C.G. Uy, P.R. August, T. Liu, and B.S. Moore. Genetic screen for marine actinomycete natural products. *Pacific Northwest Streptomyces Meeting*, 1997.

[75] E. E. van Tamelen and V. B. Haarsted. The direct oxidative conversion of 2-methyltryptophan to 4-acetylquinoline. *Tet. Lett.*, page 390, 1961.

[76] J. Vandeputte and J.D. Dutcher. Thiostrepton, a new antibiotic. II. Isolation and chemical characterization. *Antibiot. Ann.*, pages 560–561, 1955-1956.

[77] J. Walker. Total structure of the polythiazole-containing antibiotic micrococcin P. A ^{13}C nuclear magnetic resonance study. *J. C. S. Chem. Comm.*, pages 706–708, 1977.

[78] J.M. Walker. The Dansyl-Edman method for peptide sequencing. *Methods in Molecular Biology*, 32:329–334, 1994.

[79] R.H. Woodman. *Molecular and Biochemical Studies on Thiopeptide Antibiotics.* PhD thesis, The Ohio State University, 1991.

[80] P. Zhou. *Studies on antibiotic biosynthesis.* PhD thesis, Shanghai Medical University, 1993.

[81] P. Zhou, D. O'Hagan, U. Mocek, A. Zeng, L.D. Yuen, T. Frenzel, C. Unkefer, J. Beale, and H.G. Floss. Biosynthesis of the antibiotic thiostrepton. Methylation of tryptophan in the formation of the quinaldic acid moiety by transfer of the methionine methyl group with net retention of configuration. *J. Am. Chem. Soc.*, 111:7274–7276, 1989.

[82] G. Zubay. *Biochemistry.* Addison-Wesley, Reading, MA, 1993.

VITA

Paul Richard Shipley was born on August 13, 1969 in Bremerton, Washington to Sue Ann and Richard Shipley. He obtained his Bachelor of Science in Chemistry at the University of Washington in March, 1992, and his Master of Science in Chemistry, also at the University of Washington, in June, 1993.