AN ABSTRACT OF THE DISSERTATION OF

Laura L. Grochowski for the degree of Doctor of Philosophy in Pharmacy presented on June 11, 2004.
Title: Molecular Genetics and Enzymology of Secondary Metabolite Biosynthesis: I. Isolation of Natural Product Biosynthesis Gene Clusters from Symbiotic Marine Organisms II. Enzymology of Blasticidin S Biosynthesis.

Abstract approved:

Redacted for Privacy

T. Mark Zabriskie

Molecular genetic and enzymological techniques have been employed to study secondary metabolite biosynthesis. These investigations have focused on two projects: the cloning and heterologous expression of biosynthetic gene clusters from unculturable marine organisms and the characterization of individual enzymes involved in the biosynthesis of the antifungal agent blasticidin S.

The marine environment is proving to be a valuable source of biologically active compounds, but problems associated with sustainable harvest, laboratory culture, and organic synthesis make obtaining sufficient quantities of compounds for drug development both difficult and expensive. A method has been developed for the isolation of biosynthetic gene clusters from complex marine microbe/invertebrate associations. Using this method a mixed polyketide synthase (PKS) and nonribosomal peptide synthetase (NRPS) gene cluster has been cloned from the marine sponge *Jaspis splendens*. The cloned gene cluster was found to code for a PKS with three extension modules and an NRPS with three extension modules. In addition, several open reading frames (ORFs) were identified that may be

involved in the biosynthesis of the PKS starter molecule. Partial characterization of catalytic domains from the NRPS was also completed.

The second project centers on the characterization of enzymes involved in blasticidin S (BS) biosynthesis. Two ORFs were identified in the BS gene cluster encoding gene products predicted to be involved in the early steps of BS biosynthesis. The *bls*G gene product has sequence similarity to lysine 2,3-aminomutase and is believed to be involved in the formation of the β-arginine moiety of BS. A series of heterologous expression studies were undertaken to determine the function of BlsG.

The product of *blsM* exhibits sequence homology with several nucleosidetransferases. *blsM* was cloned from the BS gene cluster, heterologously expressed in *E. coli*, and shown to catalyze the formation of cytosine using cytidine 5'-monophosphate as the preferred substrate. Point mutations were introduced in *blsM* to generate three BlsM mutant enzymes: S92D, E98A, and E98D. All three mutants lost cytidine 5'-monophosphate hydrolysis activity. Surprisingly, the BlsM S92D mutant exhibits cytidine deaminase activity when incubated with cytidine or deoxycytidine, resulting in the formation of uridine and deoxyuridine, respectively.

©Copyright by Laura L. Grochowski
June 11, 2004
All Rights Reserved

Molecular Genetics and Enzymology of Secondary Metabolite Biosynthesis:

I. Isolation of Natural Product Biosynthesis Gene Clusters from Symbiotic Marine Organisms

II. Enzymology of Blasticidin S Biosynthesis

by

Laura L. Grochowski

A DISSERTATION

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented June 11, 2004

Commencement June 2005

Doctor of Philosophy dissertation of Laura L. Grochowski presented on June 11, 2004.

APPROVED:

Redacted for Privacy

Major Professor, representing Pharmacy

Redacted for Privacy

Dean of the College of Pharmacy

Redacted for Privacy

Dean of the Graduate School

I understand that my dissertation will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my dissertation to any reader upon request.

Redacted for Privacy

Laura L. Grochowski, Author

ACKNOWLEDGEMENTS

I would like to express gratitude to my advisor, Dr. Mark Zabriskie, for his willingness to accept me as a student and to provide invaluable mentorship, patient guidance, and support. I would also like to thank Dr. Phil Proteau for his advice and guidance, especially during Dr. Zabriskie's sabbatical. I would like to thank the other members of my graduate committee, Dr. Mark Leid, Dr. Theresa Filtz, and Dr. David Williams for their guidance and assistance throughout my graduate studies as well as for their critical reading of this dissertation.

It is important to recognize the invaluable assistance of Dr. Xihou Yin and Dr. Martha Cone for their mentorship and guidance during the course of my graduate studies and for teaching me the skills necessary for the completion of this work.

I would also like to acknowledge past and present members of the Zabriskie and Proteau groups, especially Dr. Mike Jackson, Dr. Youhhi Woo, Dr. Chanokporn (Shiny) Phaosiri, Dr. Dojung Kim, Roberta Fernandez, David Blanchard, Ann Mintie, and Morgan Parker for their support, friendship, assistance, and encouragement. I would like to thank the faculty and staff of the College of Pharmacy for their continued assistance and support. Special thanks go to Debra Peters and Gary Miller for their kind support and friendship.

Most importantly, I will be forever indebted to my family for their love, support and encouragement. My parents, Dave and Rose Wilhelm, fostered my scientific interests and provided the encouragement and confidence necessary to pursue them. My most sincere gratitude goes to my best friend and beloved husband Paul, for his unconditional love, support and friendship, without which I would not have had the courage, emotional strength or endurance needed to complete this research.

Finally, I would like to acknowledge the American Cancer Society, the Medical Research Foundation of Oregon, National Institutes of Health, and the American Foundation for Pharmaceutical Education for funding myself and this research.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

Page

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF TABLES

LIST OF ABBREVIATIONS

A	Adenylation
ACP	Acyl carrier protein
AdoMet	*S*-adenosylmethionine
AT	Acetyltransferase
AMP	Adenosine monophosphate
AP	Alkaline phosphatase
ATP	Adenosine triphosphate
BAC	Bacterial artificial chromosome
BLAST	Basic logical alignment search tool
BS	Blasticidin S
BSA	Bovine serum albumin
C	Condensation
cDNA	Complementary DNA
CFE	Cell free extract
CGA	Cytosylglucuronic Acid
CMP	Cytidine monophosphate
CoA	Coenzyme A
DBS	Demethylblasticidin S
dCMP	Deoxycytidine monophosphate
DH	Dehydratase
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
EDTA	ethylenedinitrilo-tetraacetic acid
Epi	Epimerase
ER	Enoylreductase
gDNA	Genomic DNA
GST	Glutathione *S*-transferase
g	gravity

h	Hours
HEPPS	4-(2-Hydroxyethyl)piperazine-N'-propanesulfonic acid
HPLC	High pressure liquid chromatography
KR	Ketoreductase
KS	Ketosynthase
LBS	Leucylblasticidin S
LDBS	Leucyldemethylblasticidin S
min	minutes
MT	Methyltransferase
NCBI	National Center for Biotechnology Information
NMR	Nuclear magnetic resonance
NRPS	Nonribosomal peptide synthetase
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PCP	Peptidyl carrier protein
PCR	Polymerase chain reaction
PKS	Polyketide synthase
Ppant	Phosphopantetheine
PP_i	Inorganic pyrophosphate
RAM	Arginine aminomutase
RNA	Ribonucleic Acid
rRNA	Ribosomal RNA
s	seconds
SDS	Sodium dodecyl sulfate
SNAc	Acyl-*S*-*N*-acetylcysteamine
T	Thiolation
TE	Thioesterase
TLC	Thin layer chromatography
tRNA	Transfer RNA

LIST OF SCHEMES

Molecular Genetics and Enzymology of Secondary Metabolite Biosynthesis:

I. Isolation of Natural Product Biosynthesis Gene Clusters from Symbiotic Marine Organisms

II. Enzymology of Blasticidin S Biosynthesis

Forward

Despite the rise in infectious disease, development of multidrug resistant bacteria, and the need to develop new therapies for diseases such as cancer and hypertension, the rate of new drug discovery is decreasing.[1] Over the past twenty years the number of new biologically active compounds has declined and in 2001 the FDA received only sixteen new drug applications.[2] Antibiotics, in particular, are a class of drugs that have not kept pace with the continuing need for new compounds that act through alternate mechanisms of action.[1] These trends demonstrate the need to develop new methodologies and techniques to improve the rate of drug discovery and development.

A common method of drug development is based on the discovery of a bioactive natural product and the subsequent optimization of the compound through a series of structure activity relationship (SAR) studies. These studies are limited by the ability of the medicinal or organic chemist to produce the desired compound analogs in a time and cost efficient manner. With the development of high throughput screening methodologies, it has become difficult for the medicinal chemist, who produces an average of thirty new compounds a year,[3] to keep up with this demand. The field of combinatorial chemistry has largely evolved out of a need to provide novel compounds for high throughput screening applications, however the expectations for drug lead discovery has not been realized.[2] The genomics and proteomics eras are further increasing the demand for novel compounds as new pharmacological targets are rapidly being discovered.[2]

The genomics era has not only provided a plethora of new drug targets, it has also provided a new avenue for drug discovery and development. In 1984 the first cloning and heterologous expression of a bacterial biosynthetic gene cluster was reported,[4] opening the door for molecular biological approaches to drug discovery and development. Dramatic advances in cloning and expression of antibiotic gene clusters

made in the twenty years since then are now being applied to the production and development of biologically active compounds. The work presented in this thesis involves the application of molecular and enzymological techniques in the search and development of new drugs and drug leads.

Two fronts in the development of new, more effective drugs will be discussed. The first project, presented in chapters 1-3, was developed to address a need that is being generated from the large number of drug leads being isolated from marine organisms. Many of the compounds are found to have unique and interesting chemical structures and biological activities; however the important issue of supply must be addressed for the development of new drugs from these compounds. Section 1 of this thesis will discuss one approach to solving this supply problem through the cloning and heterologous expression of the biosynthesis gene clusters from symbiotic marine organisms.

The second section of this thesis involves the investigation of mechanisms of antibiotic biosynthesis. Many natural products that enter clinical and preclinical trials do not achieve drug status. Factors such as solubility, target specificity, efficacy, toxicology and bioavailability limit the usefulness of otherwise promising drug candidates. In the course of drug development, chemical modification is often used to generate compound analogs that demonstrate improved pharmacological properties. This synthetic and semi-synthetic optimization can be both time consuming and expensive. An alternative to traditional synthetic drug lead optimization is the use of molecular genetic techniques to alter the biosynthetic pathways, resulting in the production of novel compounds through combinatorial biosynthesis. Such manipulation, however, requires a complete understanding of the molecular biology and enzymology of secondary metabolism. In order to further our molecular understanding of antibiotic biosynthesis, the enzymology of blasticidin S biosynthesis has been investigated. Blasticidin S is a member of the important class of nucleoside antibiotics and investigation of the biochemistry and enzymology involved in its biosynthesis has many advantages. In addition to potentially providing a means to generate blasticidin S analogs, gaining an understanding of the mechanisms involved in the biosynthesis of blasticidin S will provide insight into the biosynthesis of other related antibiotics.

Isolation of Natural Product Biosynthesis Gene Clusters from Symbiotic Marine Organisms

Chapter 1

General Introduction

Introduction to Marine Natural Products

Historically, soil microorganisms have been the largest producers of antibiotics.[5] Actinomycetes, in particular have been an abundant source of bioactive compounds and produce two thirds of known microbial antibiotics.[6] In recent years, however, the rate of new antibiotic discovery from soil bacteria has diminished, and a 99% rediscovery rate for actinomycete antibiotics has been observed.[5] This decrease in antibiotic discovery is due in part to the inability to culture 99% of soil microbes through current techniques.[5,7] There is a growing effort to gain access to the chemical and biological diversity of these unculturable organisms through the development of new culture techniques and through the cloning of specific genes from the metagenome.[7,8] The metagenome has been described as the genetic composition of a biological community and specifically the genome of microbial communities.[5,8]

With the decreased rate of drug discovery from soil microbes, much attention has been turned to the world's oceans. The marine environment is proving to be equally proliferative with a large number of biological active metabolites and drug candidates being identified from marine sources. Between 1969 and 2002 more than 15,000 compounds have been isolated from marine organisms.[9,10] This number is rapidly expanding and in 2002 alone 677 new compounds were isolated from marine organisms.[11] Marine invertebrates have been shown to produce the highest number of compounds with significant cytotoxicity in a screen by the US National Cancer Institute (NCI).[12] Many of these compounds have been developed as drug candidates and are in preclinical and clinical trials, with a large number being developed for their anticancer activities.[9]

A review of marine compounds that are in clinical trials reveals that a majority was isolated from marine invertebrates.[9] Most of these invertebrates exist in complex assemblages, and often symbiosis, with a variety of microorganisms.[13] These microorganisms may constitute up to 40% of the total biomass of the marine invertebrate.[9] Evidence is accumulating to support the long held belief that, in many cases, the associated microbes are the true source of the active natural product. For example, okadaic acid isolated from a *Halichondria* marine sponge was later found to be produced by dinoflagellates.[9] Additional support, though not proof, of a microbial origin of natural products isolated from marine invertebrates has been provided through a number of celluar localization studies. The use of cell sorting techniques has localized theopalauamide and 13-demethylisodysidenin with proteobacterial and cyanobacterial sponge fractions, respectively.[14,15]

Although the chemical and microbial diversity in the marine environment is proving to be a rich source of bioactive metabolites, problems associated with sustainable harvest, laboratory culture, and organic synthesis makes obtaining sufficient quantities of these compounds for drug development both difficult and expensive. As more marine natural products enter preclinical and clinical trials, the need to provide a consistent supply becomes more pressing. An example of delayed drug development due to inadequate supply is halichondrin B, an anticancer compound from a marine sponge. In order for clinical development of halichondrin B to proceed a method of aquaculture for the producing sponge first needed to be developed.[12]

Harvest and extraction of the bioactive compounds from marine organisms is often not an adequate means of metabolite production.[12] Low concentrations of the desired metabolites as well as the often limited supply of the producing organisms makes this approach both economically and environmentally unsound. For instance, isolation of one gram of the potent anticancer drug ET-743 required the harvest of one metric ton of the tunicate source.[9] Similarly, in order to supply the estimated need of the potential anticancer drug, halichondrin B, would require the estimated harvest of 3,000 to 16,000 metric tons of the producing sponge.[9,10,12] These examples clearly demonstrate that the sustainable harvest of marine organisms for natural product production is unlikely.

An alternative to harvesting invertebrates is to culture the producing organisms either in the laboratory or in aquaculture. Large scale fermentation has proved to be an effective method of producing many microbial natural products such as penicillin and erythromycin.[16] Attempts have been made with varying success to bring the marine symbiotic systems, or just the microorganism, into culture. Harvesting from the wild and/or aquaculture have been used successfully with the bryostatin and ET-743 producers.[12] Unfortunately not all marine organisms are amenable to culture, limiting the applicability of this method. Much like terrestrial microorganisms, many microorganisms found in marine environments are difficult to culture.[17] Even when the organism can be brought into culture, optimization of conditions for growth and metabolite production is often tedious and time consuming.[12] Many issues such as growth rate, presence of the necessary symbionts, regulation of secondary metabolism, and scale must all be addressed for the production of the desired metabolites.

Another means of producing natural products is total synthesis or semi-synthesis from a natural product precursor. Currently, dolastatin 10, hemiasterlins, and ziconotide are produced in this manner.[10,12] This method is limited, however, by the complex structures of many bioactive marine natural products. The synthesis of many of these natural products is often further complicated by the presence of multiple stereocenters. This chemical complexity often makes this approach quite expensive and not economically feasible.[12]

The project described here proposes an alternate approach to the production of bioactive metabolites from marine organisms. Harnessing the biosynthetic genes for these compounds and expressing them in surrogate host cells may circumvent many problems associated with the practical development of drugs from marine sources. The heterologous expression of biosynthetic genes has several advantages. The primary advantage of this approach is the ability to produce large amounts of the desired metabolite through large scale fermentation of the heterologus host.[16] A second advantage is that issues such as identification and culture of the true source organism become irrelevant, as the host organism would have well established culture protocols. Additionally, genetic manipulation of the biosynthetic cluster could lead to metabolite

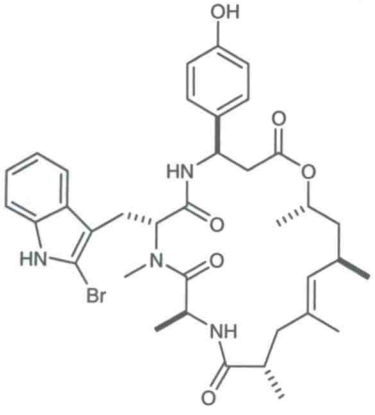
analogs through mutasynthesis and the generation of 'hybrid' gene clusters. This could be quite useful to address issues involved in drug development such as drug activity, absorption, metabolism and resistance.

Cloning biosynthetic gene clusters from the producers of bioactive metabolites presents an additional opportunity to gain access to the chemical diversity of symbiotic systems. The sequence of several microbial genomes has demonstrated the immense potential of these organism to produce natural products.[18,19] Secondary metabolite gene clusters can even out number the known metabolites produced by a given organism. The cloning and heterologous expression of these 'cryptic' gene clusters may lead to the discovery of novel natural products that may have otherwise been undetected. The utility of this approach has recently been demonstrated by Gillespie and coworkers in their studies with unculturable soil microbes. Heterologous expression of a bacterial artificial chromosome (BAC) metagenomic library from soil microorganisms resulted in the production of two antibiotics, turbomycin A and B.[20] Interestingly, the formation of these compounds was found to be a result of introducing just one gene into *E. coli*. This gene product was responsible for the formation of homogentisic acid which reacted with indole produced by *E. coli* to form these antibiotics, demonstrating the formation of a chimeric biosynthetic pathway.[20] These results represent the expression of just a single biosynthetic gene, and not an entire complex gene cluster, however it does demonstrate the potential of this approach.

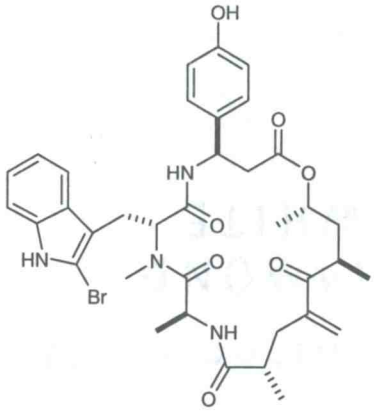
Selection of Jaspamide as an Initial Target

Initial work focused on cloning the biosynthetic genes that produce jaspamide (**1.1**), a bioactive depsipeptide isolated from a marine sponge (Figure 1.1).[21,22] Jaspamide (a.k.a., jasplakinolide) was the first member discovered of a family of a cyclic marine depsipeptides that includes the geodiamolides and neosiphoniamolide A. Jaspamide is formed from the condensation of an unusual tripeptide (L-alanyl-*N*-methyl-2-bromo-D-tryptophanyl-D-β-tyrosine) with a 12-carbon tetraketide that appears to be propionate-derived. Two related jaspamide derivatives, jaspamide B (**1.2**) and C (**1.3**), have also been isolated from *Jaspis splendens* in minor quantities and contain a modified

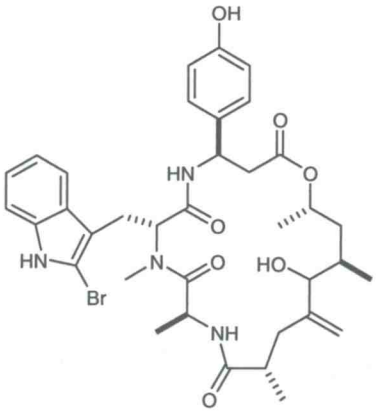
polyketide moiety with an exocyclic methylene and either a ketone or a hydroxyl functionality, respectively.[23]



Jaspamide (**1.1**)



Jaspamide B (**1.2**)



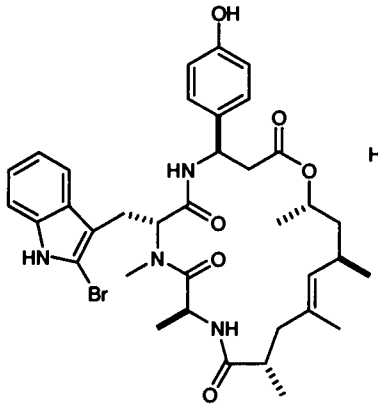
Jaspamide C (**1.3**)

Figure 1.1. Jaspamide producing sponge, *Jaspis splendens,* Photo courtesy M.K. Harper

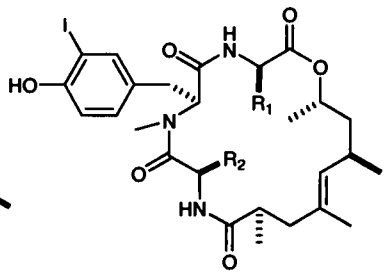
Jaspamide has been shown to have potent insecticidal and antifungal activity as well as toxicity towards renal, CNS, breast and prostate tumor cell lines.[21,24,25] The antiproliferative action of jaspamide occurs through the stabilization of F-actin and the promotion of actin polymerization.[26] Two primary effects are seen in cells treated with jaspamide. One *in vitro* effect of jaspamide is the stabilization of existing actin filaments. This results in the interruption of the cell cycle and has been shown to cause the formation of multinucleated cells.[24,27] The disruption of the cytoskeleton by jaspamide has been shown to induce apoptosis through a caspase-3-like protease-dependent

pathway.[24,28] A second *in vitro* effect of jaspamide is the altered polymerization of actin which results in the formation of perinuclear actin bundles.[29] Jaspamide binds to the interface of actin subunits and reduces the size of polymerization nucleation particles. This results in the formation of amorphous actin masses and the disruption of cellular structure.[26,27,30] This unique mechanism of action has resulted in jaspamide becoming a useful pharmacological tool for studying the role of actin polymerization and depolymerization in various cellular functions such as endocytosis, cell adhesion, locomotion, and vesicle sorting and release.[31-33]

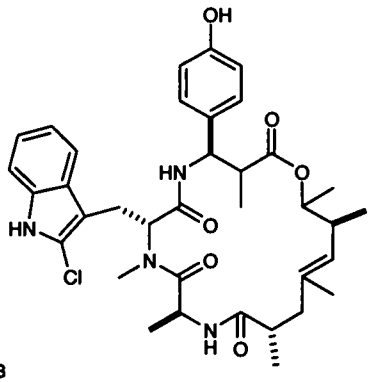
Although largely circumstantial, a large amount of evidence suggests that jaspamide is produced by a microbial symbiont of the *Jaspis* sponge. Support for this theory comes from the large number of sponge species from which jaspamide has been isolated. In addition to *J. splendens*, jaspamide has been isolated from a number of different sponge species representing two taxonomic orders including *Hemisterella minor*,[34] *Cymbastela sp.*,[35] *Dorypleres sp.*,[36]and *Auletta cf. constricta*[37] (Figure 1.2) The geographic distribution of sponges containing jaspamide is also quite varied and include South Pacific collections from Palau, Fiji, Paupa New Guinea, and Vanuatu, as well as South Africa (*H. minor).*[21,22,34,38] Other cyclodepsipeptides that closely resemble jaspamide in structure have been isolated from an even greater geographic range. Geodiamolides (**1.4**) have been isolated from a several sponge genus in the Caribbean (*Geodia sp.*),[39] South Africa (*H. minor)*[34] and the South Pacific (*Cymbastela sp.*, *Pseudaxinyssa sp.*)[34,40,41] while neosiphoniaolide has been isolated from New Caledonia (*Neosiphonia superstes).*[42] The chemical structures of these compounds are quite similar as they all contain an identical polyketide moiety cyclized with a tripepide and suggest a similar biosynthetic mechanism is involved in their production.



Jaspamide (1.1)



Geodiamolide A (1.4): $R_1 = R_2 = CH_3$
Neosiphoniamolide A: R_1 = iPr, R_2 = H

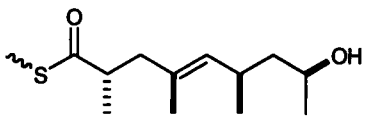


Chondramide D (1.5)

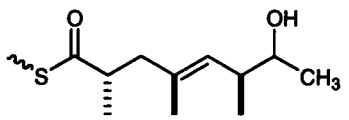
Compound	Source	Taxonomy	Location
Jaspamide	*Jaspis splendens*	order Choristida	Fiji, Palau
	Dorypleres sp.	order Choristida	Papua New Guinea
	Hemisterrella minor	order Hadromerida	South Africa
	Auletta cf. constricta	order Axinellida	Papua New Guinea
	Cymbastela sp.	order Halichondraida	Paupa New Guinea
Chondramide	*Chondromyces crocatus*	Myxobacterium	soil bacterium
Geodiamolide A & B	*Geodia sp.*	order Choristida	Caribbean
Geodiamolide A-F	*Cymbastela sp.*	order Halichondrida	Papua New Guinea
Geodiamolide TA	*Hemiasterella minor*	order Hadromerida	South Africa
Neosiphoniamolide A	*Neosiphonia sp.*	order Lithistida	New Caledonia

Figure 1.2. Taxonomic and geographic distribution of jaspamide and related compounds

It is of particular interest that the compounds most similar to jaspamide, the chondramides (**1.5**), have been isolated from terrestrial myxobacteria such as *Chondromyces crocatus*.[43] Chondramide D differs from jaspamide only in the use of an acetate rather than lactate as a starter unit for the PKS (Figure 1.3), an α-methyl group on the β-tyrosine residue and a chlorine rather than a bromine on the tryptophan residue. In addition, a number of related chondramides have been identified which contain a methoxy group on the α-carbon of β-tyrosine.



Jaspamide linear PKS precursor



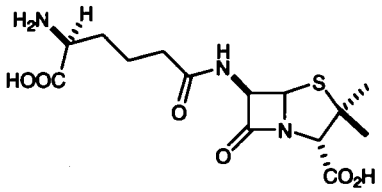
Chondramide linear PKS precursor

Figure 1.3. Comparison of jaspamide and chondramide linear polyketide precursors.

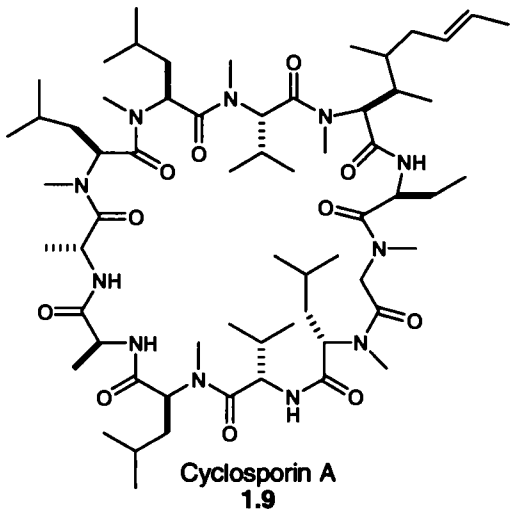
Jaspamide and other depsipeptides are presumably produced through mixed biosynthetic pathways that include nonribosomal peptide synthetases (NRPS) and polyketide synthases (PKS). This mixed biosynthetic origin makes jaspamide an attractive target for entrance into the area of cloning biosynthetic pathways from marine organisms because the gene cluster is expected to contain both PKS and NRPS elements. The modular nature of NRPS and PKS genes and the presence of highly conserved and characteristically spaced signature motifs permits PCR amplification of segments of these genes. These fragments can then be used as probes to screen a genomic library and clone the biosynthetic gene cluster.

Nonribosomal Peptide Synthetases

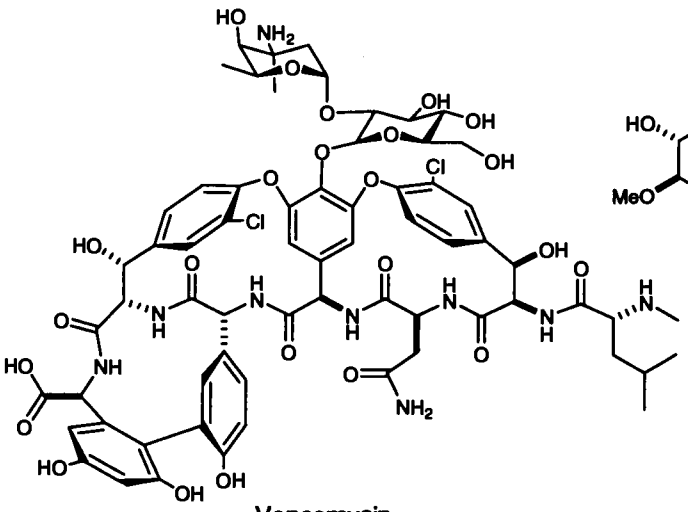
In addition to jaspamide, numerous clinically important compounds are produced by nonribosomal peptide synthesis that exhibit diversity in both structure and biological activity. Several well know antibiotics such as penicillin (**1.6**), vancomycin (**1.7**) and viomycin (**1.8**) as well as immunosuppressive drugs such as cyclosporine (**1.9**), rapamycin (**1.10**) and FK 506 (**1.11**), all have biosynthetic pathways that include nonribosomal peptide synthetases (NRPS).[44] The structural diversity of compounds produced through NRPS biosynthesis is derived largely by the presence of unusual amino acids. Unlike ribosomally synthesized peptides and proteins, nonribosomal peptides are not restricted to the twenty common amino acids and often contain nonproteinogenic amino acids. Over 300 different amino and hydroxyl acid NRPS substrates have been identified including D, β, and *N*-methylated amino acids.[45]



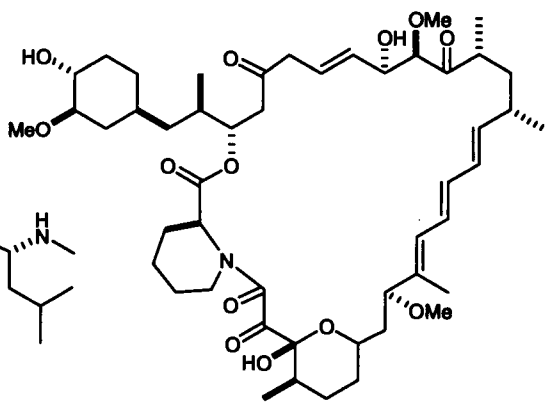
Penicillin N
1.6



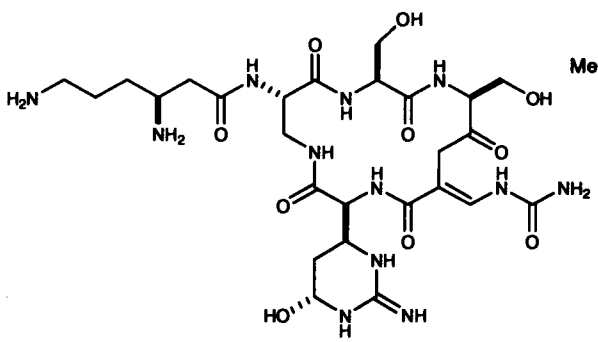
Cyclosporin A
1.9



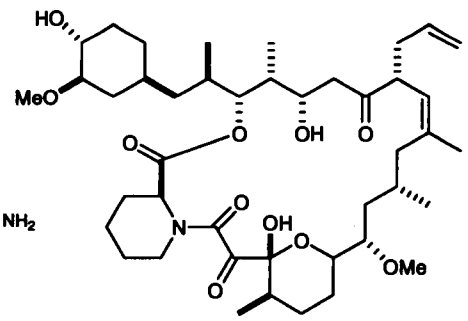
Vancomycin
1.7



Rapamycin
1.10



Viomycin
1.8



FK506
1.11

Nonribosomal peptide synthetases (NRPS) are able to utilize these nonproteinogenic amino acids because they catalyze the biosynthesis of peptides without the use of a nucleic acid template. NRPSs are modular multifunctional proteins or protein complexes that catalyze the formation of peptide bonds through covalently bound enzyme intermediates. The individual modules are organized in an assembly line fashion, with each module incorporating a single amino acid into the growing peptide as shown in Figure 1.4.[46,47] This mechanism of biosynthesis is referred to as a "multiple carrier thiotemplate mechanism" and is also found in the type I polyketide synthases.[44,48]

The NRPS modules may be located on a single multifunctional enzyme or on several, smaller enzymes. In fungal NRPSs the modules are contained on a single multifunctional enzyme that can be quite large. For example, cyclosporin synthetase is a 1.6 MDa multifunctional protein containing 11 amino acid incorporating modules.[49] Bacterial systems usually have multi-enzyme systems with one or more modules located on each enzyme or a large multifunctional enzyme.

The catalytic cycle of amino acid incorporation can be broken down into four steps: priming, chain initiation, elongation and termination. These steps are controlled by a series of catalytic domains that are contained within each NRPS module. These domains include a so-called "minimal module" composed of condensation (C), adenylation (A), and peptidyl carrier protein (PCP) domains (Figure 1.5).[50,51] Additional modification of the incorporated amino acid may also occur prior to the next elongation step. These modifications include *N*-methylation, *O*-methylation, epimerization, reduction, oxidation, and heterocyclization and are catalyzed by accessory domains found within the appropriate module.[51,52] The roles of the main catalytic domains in the synthesis of a nonribosomal peptide are described below.

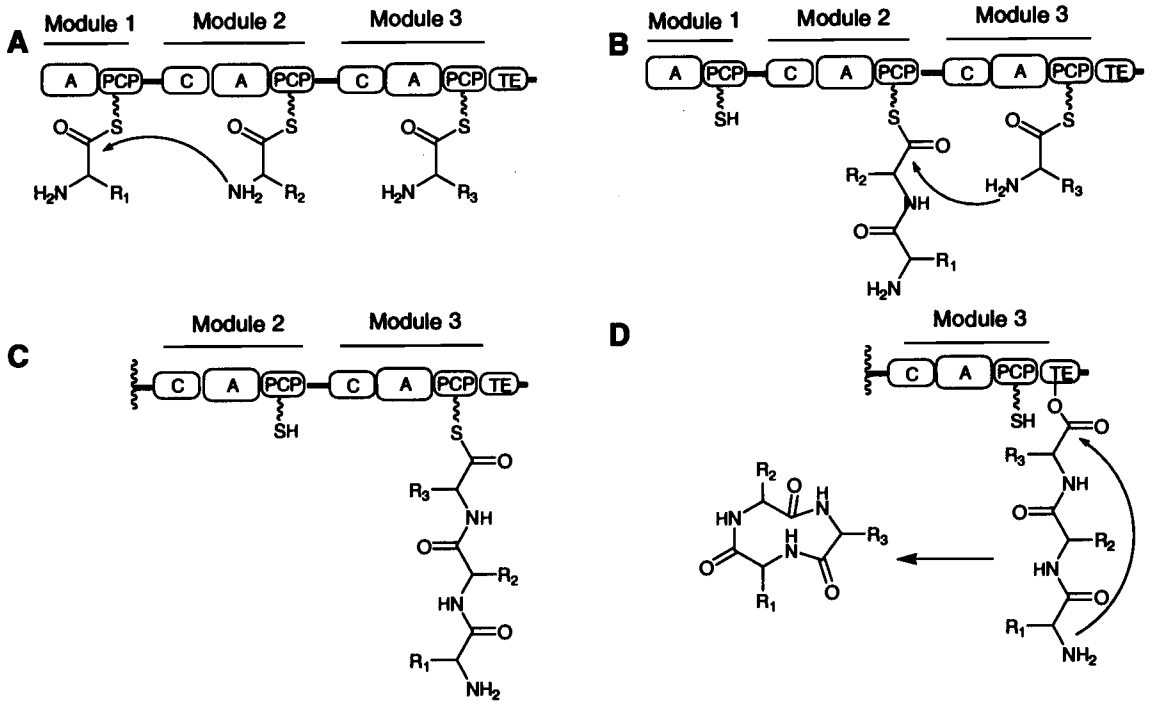


Figure 1.4. Formation of a tripeptide by a NRPS. A. Activation and loading of substrate amino acids by A-domains. B,C Chain initiation and peptide bond formation by C-domains. D. Transfer of peptide to TE domain and release of product.

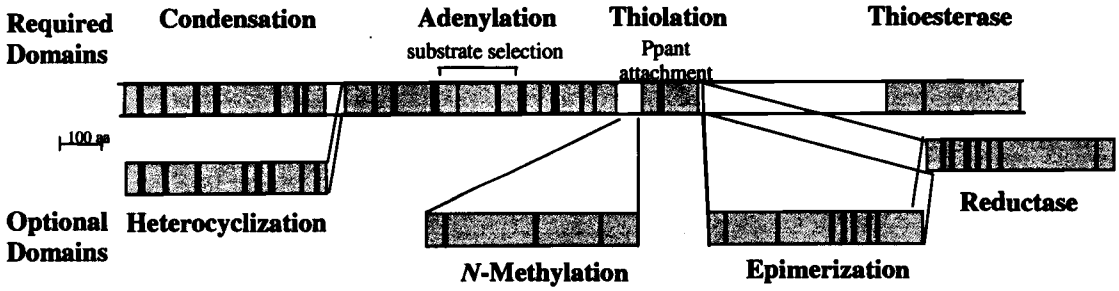


Figure 1.5. Organization of a NRPS module. Dark bands indicate the location of highly conserved amino acid sequence motifs.

Priming of the NRPS for Catalysis

Initial ribosomal synthesis of an NRPS produces an inactive *apo* form of the enzyme. Prior to catalysis, the NRPS is converted to the *holo* form through the addition of a 4'-phosphopantetheine cofactor (Ppant-cofactor) to a conserved serine residue in the PCP domain (also known as the thiolation, or T domain).[53] The Ppant-cofactor performs a similar function in NRPS and PKS enzymes as it does in fatty acid biosynthesis, and acts as a 'swinging arm' that allows for the movement of a covalently bound substrate or peptide intermediate between the catalytic domains of the NRPS modules. A phosphopantetheinyl transferase (PPTase) is often located within the biosynthetic gene cluster and is responsible for the formation of the *holo* enzyme. The PPTase transfers the phosphopantetheine group from coenzyme A (CoA) to the conserved serine residue on the PCP (Figure 1.6).[53]

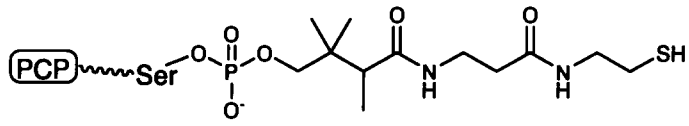


Figure 1.6. 4'-Phosphopantetheine cofactor on a PCP

The PCP domains perform the critical role of directing the flow of the nascent peptide down the biosynthetic assembly line. The mechanism of this direction has been found to be based on differing affinity of the PCP to the various catalytic domains, depending on the state of the substrate bound to the Ppant-cofactor (Figure 1.7).[54] In the absence of a bound substrate, the Ppant-cofactor has highest affinity for the A-domain, facilitating loading of the activated amino acid (Figure 1.7, a). After formation of the aminoacyl-Ppant, binding affinity for the A-domain decreases and the loaded cofactor has highest affinity for the acceptor position of the upstream condensation domain (Figure 1.7, b). Following formation of the peptide bond, the affinity of the Ppant-cofactor changes again and exhibits maximal binding affinity for the donor site of the downstream condensation domain (Figure 1.7, c).

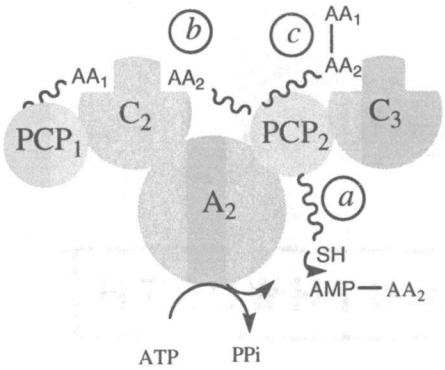


Figure 1.7. The sequence of events involved in one round of peptide elongation by an NRPS module.[54]

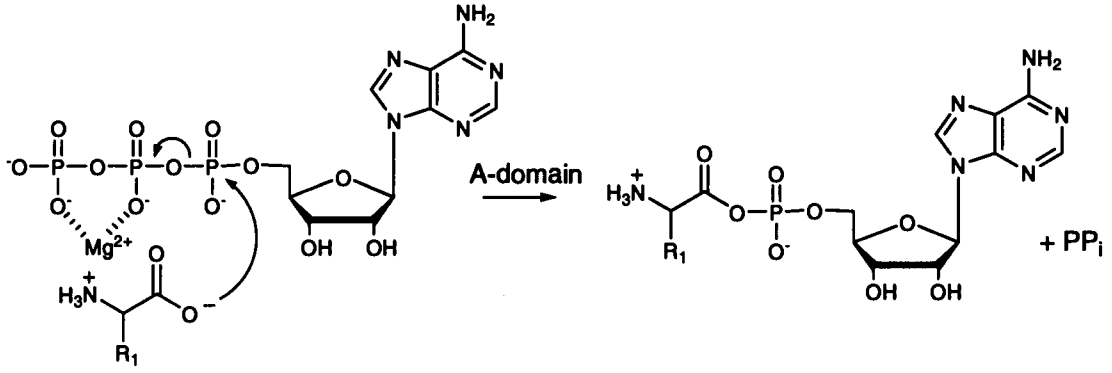
Adenylation Domain

Once formed, the *holo* NRPS modules are loaded with the respective amino acids. Adenylation domains (A-domain) catalyze both the activation of the amino acid substrate as the aminoacyl adenylate and the subsequent loading of the amino acid onto the PCP. A-domains are around 550 amino acids in length and are analogous to tRNA synthetases in activity as they confer amino acid specificity and activate the amino acid for incorporation into the peptide. Despite homology to tRNA synthetases, A-domains are most similar to firefly luciferases in tertiary structure.[55]

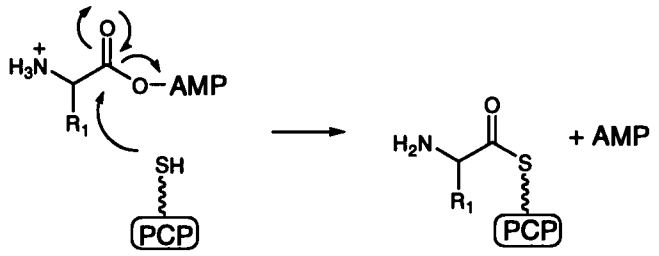
The fidelity of a NRPS is largely reliant on the ability of the A-domain to specifically activate the required amino acid substrate. The crystal structure of the gramicidin S synthetase phenylalanine activating domain (PheA) provided insight into substrate selectivity.[55] Ten amino acid residues were identified that line the amino acid binding pocket of the A-domain and confer substrate specificity. The crystal structure of PheA along with sequence alignments of A-domains of known substrate specificity allowed for the definition of a so called 'nonribosomal code'.[56,57] Key amino acid residues include an aspartate residue that is involved in binding of the α-amine of the amino acid substrate. Similarly, a conserved lysine residue was found to be involved with binding the substrate α-carboxyl group. Several additional residues identified within the PheA A-domain were implicated in binding of the amino acid side chain. The importance of these residues in substrate specificity was demonstrated by Stachelhaus *et al.* by the alteration of A-domain substrate specificity through mutagenesis of these

residues.[56] Application of the 'nonribosomal code' for the prediction of an A-domains substrate specificity will be discussed further in later chapters.

Following substrate selection, the A-domain activates the cognate amino acid as the aminoacyl adenylate at the expense of ATP (Scheme 1.1).[58] The amino acid is then loaded onto the PCP domain through attack of the Ppant thiol on the carboxyl group of the activated amino acid, resulting in formation of a thioester (Scheme 1.2).



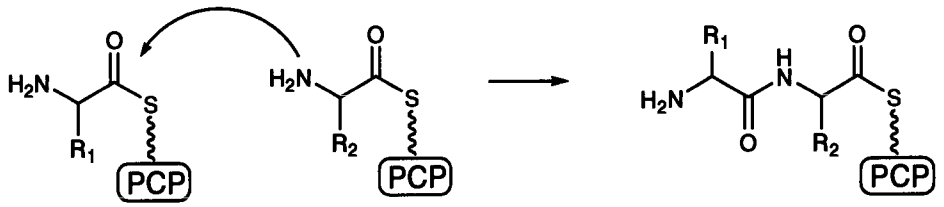
Scheme 1.1. Activation of an amino acid substrate by a NRPS A-domain



Scheme 1.2. Loading of the PCP domain

Chain Initiation and Elongation

Formation of the first peptide bond occurs when the electrophilic aminoacyl-*S*-PCP is attacked by the nucleophilic amino group of a downstream aminoacyl-*S*-PCP (Scheme 1.3). This reaction is catalyzed by the condensation domain (C-domain). The sequential actions of the condensation domains result in the movement of the growing peptide down the NPRS assembly line until it reaches the last module.



Scheme 1.3. Peptide bond formation by a NRPS C-domain

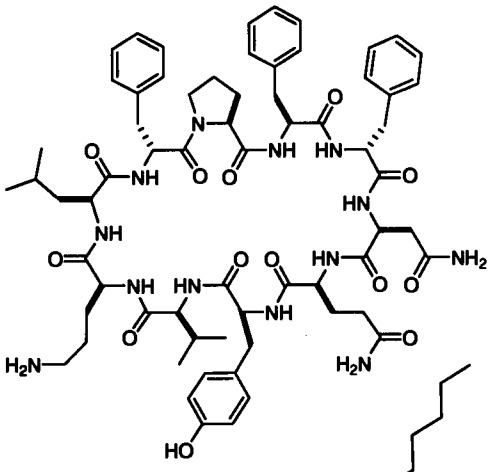
It has been shown that the condensation domain adds an additional level of fidelity to the NRPS systems. Walsh and coworkers have shown that the condensation domains demonstrate a degree of substrate stereoselectivity and do not catalyze peptide bond formation when an inappropriate substrate is loaded onto the PCP.[59] Clugston *et al.* artificially loaded noncognate aminoacyl-CoA derivatives on to the PCP and tested the ability of the downstream condensation domain to form the peptide bond. These experiments determined that the condensation domain exhibits a degree of substrate specificity for the aminoacyl-*S*-PCP in the acceptor position and adds a 'proof reading' functionality to the enzyme.[59,60]

Termination

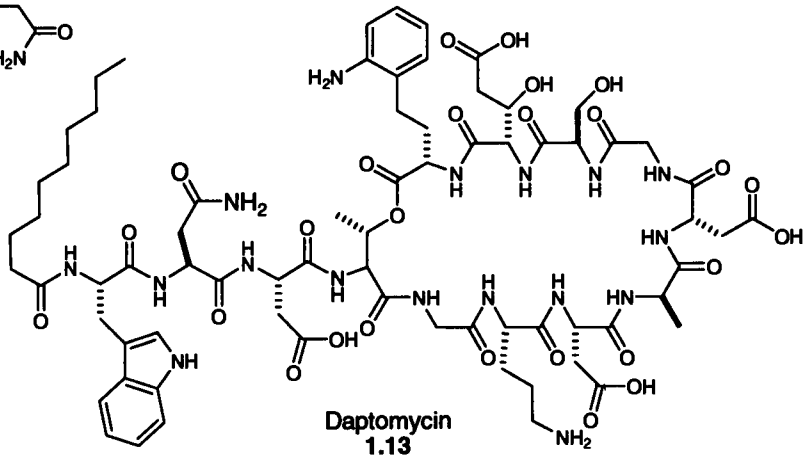
Following formation of the final peptide bond, the nascent linear peptide is cleaved from the enzyme through hydrolysis or cyclization (Figure 1.4). Several mechanisms for this cleavage have been identified, however the most common method of cleavage is through the action of a thioesterase (TE) domain.[61] For cleavage to occur, the peptide is first transferred from the PCP thiol to a serine residue on the TE domain via a transesterification reaction. Release occurs through the nucleophilic attack on the ester, resulting in cleavage of the peptide-enzyme bond. The nature of the nucleophile determines the structure of the released product. Examples of potential nucleophiles include amino-terminus or side chain amines, resulting in the formation of a lactam (e.g. tyrocidine, **1.12**), or a side chain hydroxyl resulting in the formation of a lactone (e.g. daptomycin, **1.13**). The lactone linkage in fengicin (**1.14**) results from the side chain hydroxyl of a tyrosine residue serving as the nucleophile.[51] Alternately, cleavage may occur through a hydrolysis reaction where water serves as the nucleophile, resulting in

the formation of a linear peptide as is the case of ACV (**1.15**), the linear precursor to penicillins.[61,62]

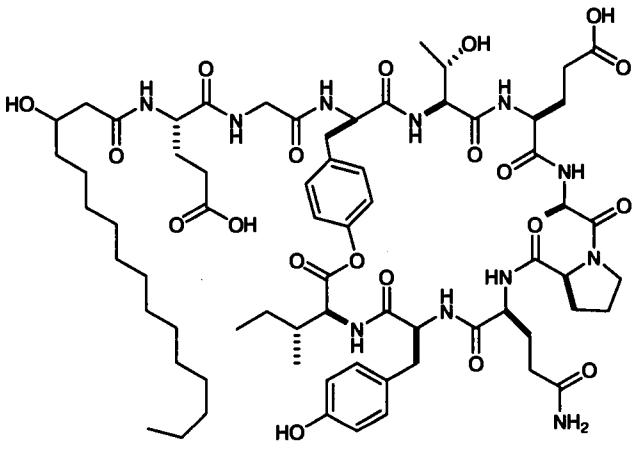
Although TE domains are the most common means of chain termination, alternate mechanisms of cleavage have been identified in several biosynthetic gene clusters. Several NRPS systems utilize a reductase domain to cleave the final thioester bond and release the product in the form of an aldehyde.[61] This is the mechanism used in nostocyclopeptide biosynthesis where the release of the formed peptide is catalyzed through an NAD(P)H-mediated hydride transfer to form the linear peptide aldehyde. Intramolecular attack of the reactive aldehyde by the amino group of the peptide *N*-terminal amino acid results in the formation of an unusual imine bond.[63] Another method of chain termination is through the formation of an amide bond by a *C*-terminal condensation domain. Although there is no direct evidence for this mechanism of cyclization, it is believed to involve the nucleophilic attack of the *N*-terminal amine on the thioester bond.[61] This type of condensation domain is believed to be involved in the biosynthesis of cyclosporin, enniatin and the depsipeptide PF1022A (**1.16**).



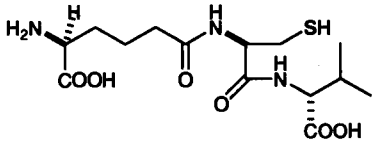
Tyrocidine A
1.12



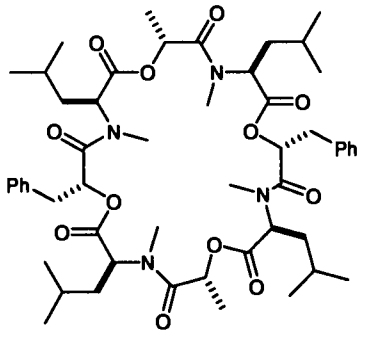
Daptomycin
1.13



Fengicin
1.14



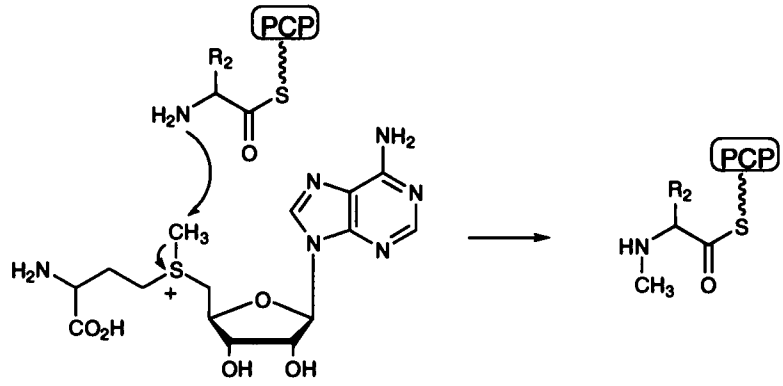
ACV
1.15



PF1022A
1.16

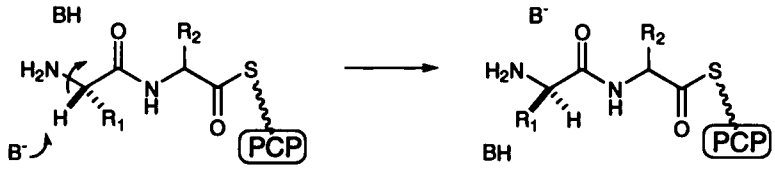
Auxiliary Domains

In addition to the catalytic domains building the central peptide-backbone, each module of a NRPS may contain one or more accessory domains. These accessory domains are responsible for the modification of the substrate amino acid and include epimerization, *N*-methylation, *O*-methylation and heterocyclization domains.[51] The presence of *N*-methylated amino acids is common in nonribosomal peptides. A specialized catalytic domain that is typically inserted into the A-domain is responsible for this transformation.[64] The methyltransferase (MT) domain transfers a methyl group from *S*-adenosylmethionine (AdoMet) to the Ppant-bound aminoacyl substate prior to formation of the peptide bond (Scheme 1.4).[64,65]



Scheme 1.4. Reaction catalyzed by an *N*-methyltransferase domain

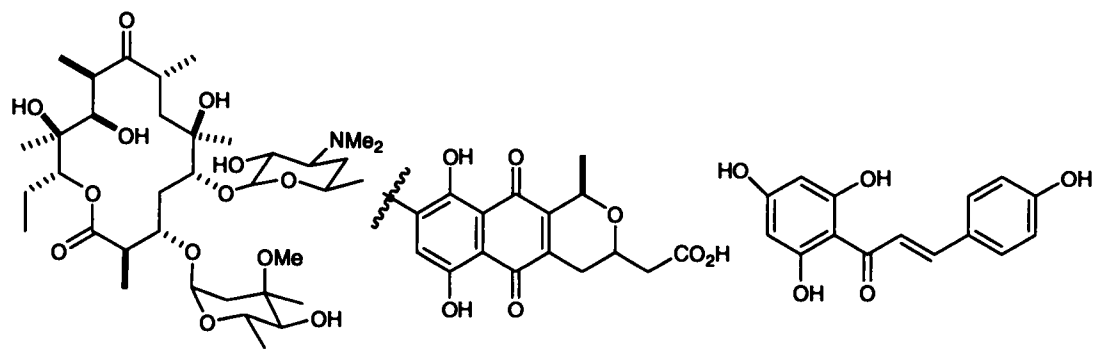
Another common auxiliary domain found in NRPS modules is an epimerization (Epi) domain. Although there are a few rare examples of A-domains that specifically activate D-amino acids, such as the D-alanine incorporating domain of cyclosporin synthetase, it is more common for the A-domain to activate an L-amino acid which is then epimerized. Although the epimerization domain is capable of epimerizing both aminoacyl-*S*-Ppant and peptidyl-*S*-Ppant species, the peptidyl-*S*-Ppant is epimerized much more efficiently (Scheme 1.5).[54] Consequently, it is believed that epimerization typically occurs after peptide bond formation.[51,66]



Scheme 1.5. Epimerization by an E-domain

Polyketide Synthases

Polyketide synthases (PKS) catalyze the formation of polyketides through the successive addition of two- or three-carbon units from either malonyl or methylmalonyl CoA.[67,68] Three types of polyketide synthases have been characterized. The first, type I PKSs, are modular enzyme systems much like the NRPSs. Erythromycin (**1.17**) is perhaps the best studied type I PKS from a mechanistic viewpoint. Type II polyketide synthases are involved in the biosynthesis of aromatic, multicyclic polyketides such as actinorhodin (**1.18**). The type II enzymes are also known as iterative polyketide synthases and are composed of catalytic domains that are repetitively used to generate a β-keto intermediate. Additional enzymes such as ketoreductases, aromatases, and cyclases catalyze the formation of the final multicyclic aromatic product. A third type of PKS is found in higher plants and bacteria and is represented by chalcone synthase, which makes chalcone (**1.19**). These type III enzymes are involved in the biosynthesis of a number of natural products including flavonoids and anthocyanins.[69] The type III PKS are different from the other PKS in that their mechanism does not involve the formation of a covalently bound enzyme intermediate.[68,69]



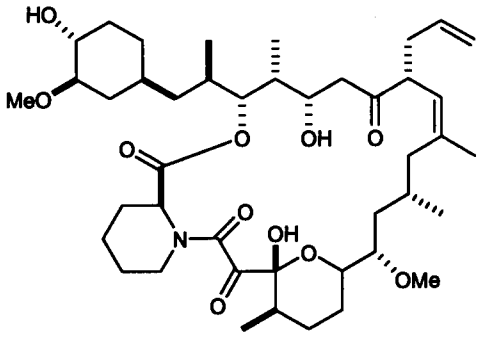
Erythromycin A
Type I PKS
1.17

Actinorhodin
Type II PKS
1.18

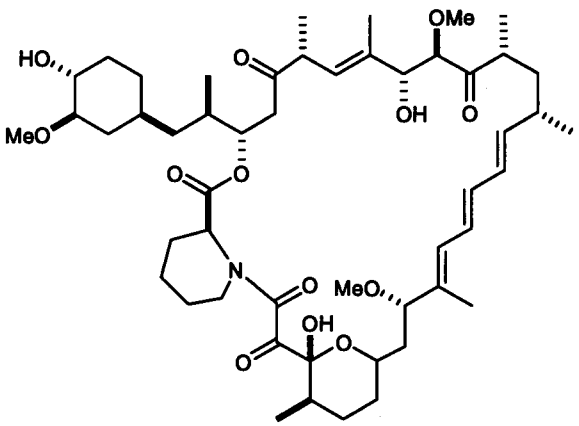
Chalcone
Type III PKS
1.19

The structure of the polyketide portion of jaspamide suggests that it is likely produced by a type I modular PKS. Many natural products produced by type I polyketide synthases have become important drugs including erythromycin (**1.17**), FK 506 (**1.20**), and rapamycin (**1.21**). Type I polyketide synthases are similar to NRPSs in many ways.

Like NRPSs, PKSs are modular enzyme systems that utilize an assembly line type mechanism involving the formation of enzyme bound thioester intermediates.



FK506 **1.20**



Rapamycin **1.21**

A minimal PKS module consists of a ketosynthase (KS)-domain, acyl ltransferase (AT)-domain and an acyl carrier protein (ACP)-domain. Also similar to the NRPS systems, the PKS modules may contain one or more accessory domains that are responsible for the step wise reduction of the ketone to a hydroxyl (ketoreductase (KR)-domain), an enone (dehydratase (DH)-domain) and an alkane (enoylreductase (ER) domain), respectively.[70] The biosynthesis of a representative type I polyketide, 6-deoxyerythronolide B the aglycone of erythromycin, is illustrated in Figure 1.8.[71,72]

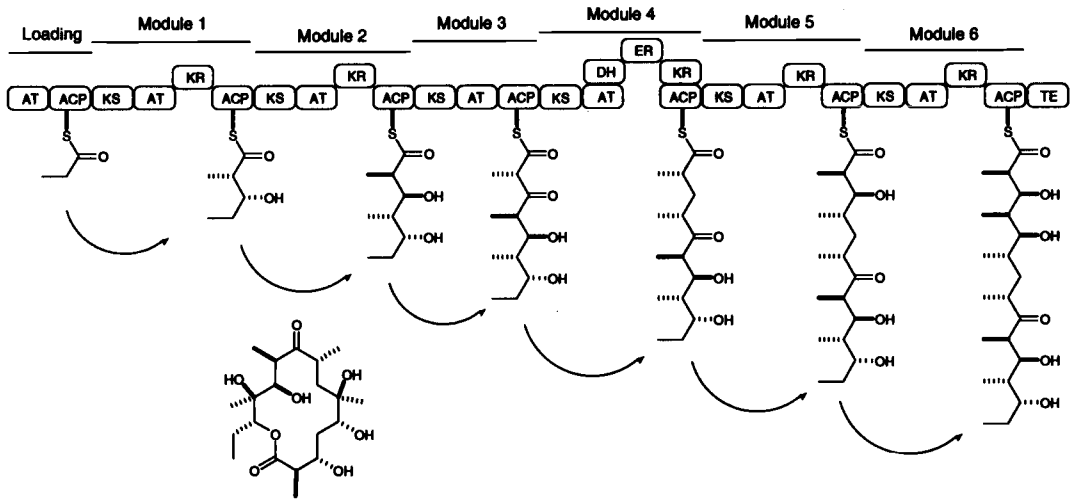


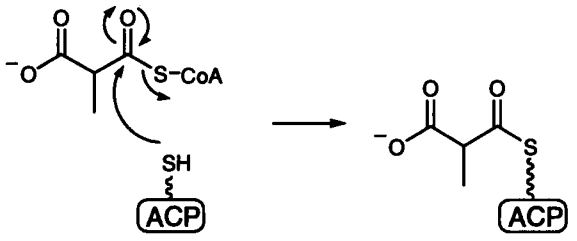
Figure 1.8. Biosynthesis of a type I PKS. Formation of 6-deoxyerythronolide B, macrolide core of erythromycin[71,72]

Although the modular organization of the type I PKS is similar to the NRPS system, there are notable differences between these two classes of enzymes. One difference is found in the quaternary structure. While NRPSs have been shown to be monomers, PKSs are dimers.[73-75] Catalytic activity of PKS is dependent on the formation of a double-helical homodimer that results in the formation of two reaction channels.[73,75] Complementation experiments have shown that the KS and ACP domains of a single active site are located on different monomer subunits of the PKS. These studies also showed that the AT domain is capable of loading either of the two active sites within the dimeric module.[75,76]

A significant source of structural diversity in nonribosomally synthesized peptides is the choice of amino acid extender units. Polyketide syntheses, on the other hand, have a limited choice of extender units and are largely restricted to malonyl and methylmalonyl CoA. Consequently, the structural diversity of polyketides is derived from the choice of starter unit and the presence of auxiliary domains within the activating modules. The role of each of the type I PKS modules in the formation of a polyketide product is discussed below.

Priming and Chain Initiation

The first step in polyketide biosynthesis is the covalent attachment, or loading, of the starter and extender units onto the acyl carrier protein (ACP) (Scheme 1.6). The ACP domain is homologous to the PCP domain of NRPSs and also contains a phosphopantetheine cofactor attached to a highly conserved serine residue within the domain. Also like the PCP, the ACP directs the movement of the substrate through the various catalytic domains.[77]

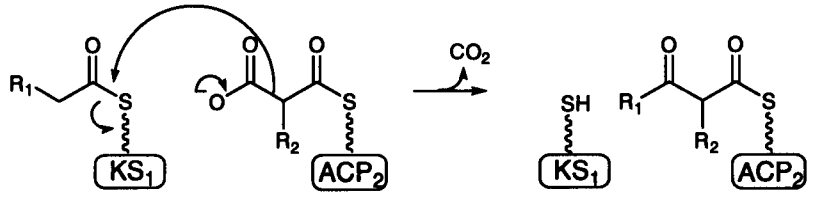


Scheme 1.6. Loading of the ACP domain

The acyl transferase (AT) domain is involved with selection and loading of the extender unit onto the ACP-Ppant cofactor. The substrates for AT domains are CoA-esters of either malonate (malonyl-CoA) or methylmalonate (methylmalonyl-CoA). Unlike the NRPS A-domains, the AT-domain substrates are already activated as CoA-thioesters; it is not necessary for the AT domain to perform this function.

Chain Elongation and Termination

The ketosynthase (KS) domain of a PKS corresponds to the NRPS condensation domain and catalyzes the extension of the growing polyketide chain. The first step in chain elongation is the transfer of starter unit or polyketide chain from the upstream ACP to a cysteine residue located within the KS active site. Chain extension occurs through a Claisen type condensation (Scheme 1.7).[78] The KS domain first catalyzes the decarboxylation of a malonyl or methylmalonyl extender unit tethered to the ACP of the same module. The resulting nucleophile attacks the thioester carbonyl of the polyketide chain bound to the cysteine residue of the KS domain. The net result of the decarboxylative condensation is the extension of the growing polyketide chain by two carbons.[78] If methylmalonyl was used as the extender unit (R=CH_3), a methyl group will be located on the α-carbon.



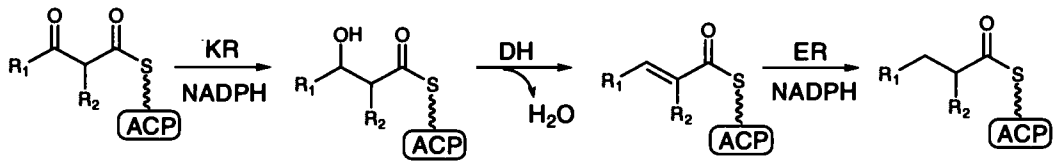
Scheme 1.7. Polyketide chain elongation

After the condensation, but prior to transfer to the KS domain, the ketide unit may undergo up to three consecutive reactions that are catalyzed by auxiliary domains.[79] Up to three auxiliary domains form a reductive loop that may be located within a PKS extension module: a ketoreductase (KR)-domain, a dehydratase (DH)-domain and an enoylreductase (ER)-domain. The oxidative state of each two carbon unit of the polyketide product is dependent on the function of these auxiliary domains within the individual modules. If there are no auxiliary domains within a given module the product

will retain a ketone at that position. Conversely, the presence of all three auxiliary domains in a given module results in the introduction of a saturated two carbon unit into the corresponding position of the product (Figure 1.8).[78]

The reactions catalyzed by these reductive domains are illustrated in Scheme 1.8. The first step in ketide reduction is the NADPH-dependent formation of the secondary alcohol by the KR domain. Dehydration by the DH-domain then results in the formation a double bond between C2 and C3 of the growing polyketide chain. A final reduction catalyzed by the ER-domain is also NADPH-dependent and results in the formation of the saturated bond.[78]

Following any modification of the polyketide, the growing polyketide chain is passed to the cysteine residue of the downstream KS-domain for the next round of chain elongation. After the final elongation step, the polyketide chain is cleaved from the enzyme through the action of a thioesterase domain, as is found with the NRPS systems. The action of the thioesterase domain usually results in the formation of a macrolactone.[44]



Scheme 1.8. Reduction catalyzed by PKS auxiliary domains

Combinatorial Biosynthesis of Polyketides and Nonribosomal Peptides

The need to generate new antibiotics, as well as to generate analogs of well established antibiotics for which resistance has emerged, has helped to fuel the growing interest in combinatorial biosynthesis. Combinatorial biosynthesis can be defined as the manipulation of biosynthetic pathways to generate new antibiotics or antibiotic analogs.[51,80] The manipulation may take several forms and includes the introduction of unnatural starter or extender units, disruption or rearrangement of a module or catalytic domain, or the formation of hybrid gene clusters by combining individual modules from two or more gene clusters.

Opportunities for combinatorial biosynthesis of NRPS and PKS compounds were made possible with the cloning and heterologous expression of the first antibiotic gene

clusters in the 1980's.[4] Heterologous expression of an antibiotic gene cluster was first demonstrated by Malpartida and Hopwood in 1984 with the actinorhodin gene cluster from *Streptomyces coelicolor*.[4] Shortly thereafter, attempts were made to use heterologus expression systems to generate novel "hybrid" natural products.[81] As the number of cloned biosynthetic gene clusters has grown, so have the methods and techniques used to generate novel secondary metabolites through combinatorial biosynthesis.

Studies with a variety of antibiotic producers found that feeding modified precursors can result in the incorporation of the novel compounds into the metabolite.[82-84] This technique has been adapted to the modular PKS systems through the use of acyl-*S*-*N*-acetylcysteamine (SNAc) alternate substrates. The *N*-acetylcysteamine moiety mimics the phosphopantethine cofactor of a *holo* PKS or NRPS (Figure 1.6).[85,86] SNAc precursors have been used with the PKS deoxyerythronolide B (DEBS) gene cluster to generate a variety deoxyerythronolide B analogs.[85,86] A similar method, in conjunction with the Sfp phosphopantetheine transferase, has been used to bypass the adenylation domain and load an alternate amino acid onto a NRPS PCP domain.[60,87] This method has been used to study domain function in the NRPS systems but it may be applied to the generation of novel antibiotics as well.[59,60]

Another method to generate novel antibiotic derivatives is through the genetic alteration of the biosynthetic gene cluster to inactivate or modify the function of a particular enzyme. The inactivation of PKS domains has been used to generate a variety of metabolite analogs. Domain inactivation has been extensively used in the DEBS gene cluster to produce a number of compounds with varying degrees of oxidation on the product and was first used to demonstrate the modular architecture of PKS.[72] Since then, a number of erythromycin analogs have been generated through targeted domain disruption.[71,88] Further, the introduction of multiple mutations into the DEBS gene cluster has allowed for the generation of libraries containing more than 60 6-deoxyerythronolide analogs.[89,90]

The identification of the substrate selectivity "code" in NRPS A-domains has allowed for the alteration of substrate selectivity through site directed mutagenesis and resulted in the incorporation of new amino acids into the antibiotic products.[56,91] Altered

substrate specificity through site directed mutagenesis has also been demonstrated in the DEBS type I PKS system.[92]

The ultimate goal of combinatorial biosynthesis in PKS and NRPS systems is the mixing of enzymes and catalytic domains from different gene clusters to generate hybrid 'engineered' natural products. Although this is a relatively new field of research much progress is being made. This so-called domain swapping has been successfully demonstrated both within a single gene cluster and between multiple gene clusters. Most of the work in this area so far has been performed in PKS systems.[93] One feature that has been exploited is the ability of the TE domain to catalyze the formation of various sized rings.[94,95] Relocation of the TE domain has been used to cleave polyketide intermediates of shorter chain length. Similar relocation of an NRPS TE from surfactin synthetase was shown to result in the formation of truncated linear surfactin analogs.[96] Excised NRPS TE domains have also proved useful in catalyzing the stereospecific cyclization of natural and synthetic-SNAc peptides.[97-100] The broad utility of the TE domain was demonstrated through the use of the tyrocidin synthetase TE in conjunction with solid phase synthesis to generate a library of cyclic peptides.[101]

Generation of new polyketides has also been accomplished by individual domain and module swapping. Kuhstoss *et al.* were able to create a platenolide analog by substituting an acetyl-CoA loading domain with a propionyl-CoA domain.[102] Similar techniques have been used to create novel analogs through exchange of elongation AT domains. This was used in the deoxyerythronolide B synthase (DEBS) gene cluster and resulted in the formation of novel erythromycin analogs.[103,104] The utility to fuse domains from different polyketide gene clusters was demonstrated with the successful fusion of an AT domain from the rapamycin gene cluster into a PKS module from the DEBS cluster.[105] The fusion of entire modules from the DEBS gene cluster, as well as hybrids fusions of DEBS and rapamycin modules, has also been used to generate engineered tri and tetraketides.[106 107]

Domain and module swapping has been met with only limited success in NRPS systems. Although the formation of surfactin derivatives through domain substitution was successful, the yields of the modified products were low.[108] Some success has been

met in various studies aimed at defining domain boundaries and developing methods for combinatorial biosynthesis, several designed di and tripeptides have been generated through NRPS domain fusion.[65,108-110]

Following release from the PKS or NRPS, many polyketides and nonribosomal peptides are further modified by tailoring enzymes. These enzymes present an additional opportunity to generate engineered natural products. The ability of several glycosyl transferases to utilize alternate sugar or aglycone substrates has been demonstrated.[111,112] Cytochrome P450 tailoring enzymes are also commonly found in PKS biosynthetic gene clusters. The P450-hydroxylases involved in pikramycin and oleandomycin biosynthesis, have been shown to catalyze the hydroxylation of alternate substrates.[113,114]

Significance and Objectives

The overall objective for this project was to develop a method for the identification, cloning and expression of biosynthetic gene clusters from marine invertebrates. The specific aim of this project was to clone the jaspamide biosynthetic gene cluster from the sponge which it was isolated from.

The first step necessary for the production of marine natural products in heterologous systems is the isolation of the targeted biosynthetic gene clusters. In marine organisms this can present quite an obstacle as many of these organisms exist in complex assemblages with a variety of microorganisms. This presents a situation where it is difficult to isolate a homogeneous sample of genomic DNA (gDNA) from the organism. Attempts have been made to isolate individual organisms from these complex assemblages through cell sorting, however this has been met with limited success.[115] As it is often not possible to isolate the producing organisms either by laboratory culture or cell sorting, the biosynthetic genes must be identified and isolated from a heterogeneous mixture of organisms. As jaspamide was isolated from a symbiotic marine sponge, the true producer of jaspamide is not known. For that reason, it was necessary to screen a metagenomic library, representing all of the organisms within the *Jaspis splendens* assemblage.

Figure 1.9 shows the experimental design used in the search for the jaspamide mixed biosynthesis gene cluster. First, high quality genomic DNA is isolated from sponge tissue containing all associated microbes. The total genomic DNA is used to generate a metagenomic DNA library as and is used as a template for the PCR amplification of NRPS and PKS gene probes using degenerate primers. The gene probes can then be used to independently screen duplicates of the cosmid library. Cosmids that hybridize with both NRPS and PKS gene probes are subjected to additional screening to identify individual gene clusters. The cosmid(s) that are found to be the best candidates for harboring the jaspamide gene cluster are then subcloned and sequenced. Finally, heterologous expression of the individual catalytic domains or the entire gene cluster can be used to confirm the function of the cloned gene cluster.

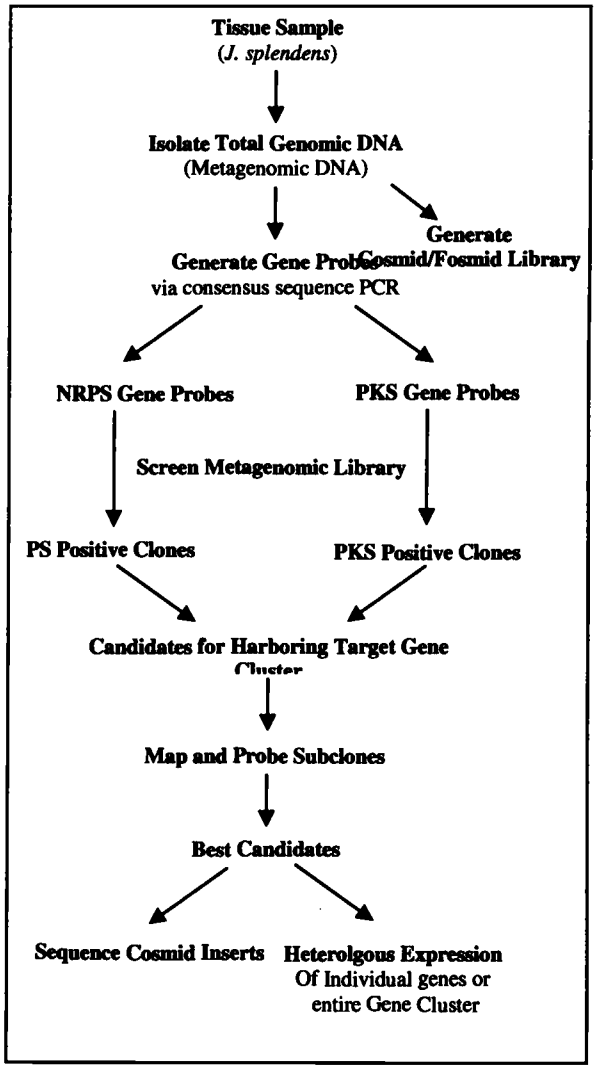


Figure 1.9. Experimental Design for the isolation of mixed NRPS/PKS gene clusters.

This approach was applied to the search for the jaspamide biosynthetic gene cluster and chapter two of this thesis will discuss the identification, cloning and sequencing of a mixed PKS/NRPS biosynthetic gene cluster from the symbiotic marine sponge, *Jaspis splendens*.

Chapter three will discuss the heterologous expression and partial characterization of NRPS adenylation domains coded for by the isolated gene cluster.

References

(1) Cassell, G. H.; Mekalanos, J. Development of antimicrobial agents in the era of new and reemerging infectious diseases and increasing antibiotic resistance. *J.A.M.A.* **2001**, *285*, 601-605.

(2) Newman, D. J.; Cragg, G. M.; Snader, K. M. Natural products as sources of new drugs over the period 1981-2002. *J. Nat. Prod.* **2003**, *66*, 1022-1037.

(3) Maehr, H. Combinatorial chemistry in drug research from a new vantage point. *Bioorg. Med. Chem. Lett.* **1997**, *5*, 473-491.

(4) Malpartida, F.; Hopwood, D. A. Molecular cloning of the whole biosynthetic pathway of a Streptomyces antibiotic and its expression in a heterologous host. *Nature* **1984**, *309*, 462-464.

(5) Handelsman, J.; Rondon, M. R.; Brady, S. F.; Clardy, J.; Goodman, R. M. Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chem. Biol.* **1998**, *5*, R245-249.

(6) Kieser, T. B., M.; Buttner, M.; Chater, K.; and Hopwood, D. *Practical Streptomyces Genetics*; 1*st* ed.; The John Innes Foundation: Nowich, 2000; 613.

(7) Kaeberlein, T.; Lewis, K.; Epstein, S. S. Isolating "uncultivable" microorganisms in pure culture in a simulated natural environment. *Science* **2002**, *296*, 1127-1129.

(8) Rondon, M. R.; August, P. R.; Bettermann, A. D.; Brady, S. F.; Grossman, T. H.; Liles, M. R.; Loiacono, K. A.; Lynch, B. A.; MacNeil, I. A.; Minor, C.; Tiong, C. L.; Gilman, M.; Osburne, M. S.; Clardy, J.; Handelsman, J.; Goodman, R. M. Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. *Appl. Environ. Microbiol.* **2000**, *66*, 2541-2547.

(9) Proksch, P.; Edrada, R. A.; Ebel, R. Drugs from the seas - current status and microbiological implications. *Appl. Microbiol. Biotechnol.* **2002**, *59*, 125-134.

(10) Salomon, C. E.; Magarvey, N. A.; Sherman, D. H. Merging the potential of microbial genetics with biological and chemical diversity: an even brighter future for marine natural product drug discovery. *Nat. Prod. Rep.* **2004**, *21*, 105-121.

(11) Blunt, J. W.; Copp, B. R.; Munro, M. H.; Northcote, P. T.; Prinsep, M. R. Marine natural products. *Nat. Prod. Rep.* **2004**, *21*, 1-49.

(12) Munro, M. H.; Blunt, J. W.; Dumdei, E. J.; Hickford, S. J.; Lill, R. E.; Li, S.; Battershill, C. N.; Duckworth, A. R. The discovery and development of marine compounds with pharmaceutical potential. *J. Biotechnol.* **1999**, *70*, 15-25.

(13) Haygood, M. G.; Schmidt, E. W.; Davidson, S. K.; Faulkner, D. J. Microbial symbionts of marine invertebrates: opportunities for microbial biotechnology. *J. Mol. Microbiol. Biotechnol.* **1999**, *1*, 33-43.

(14) Schmidt, E. W.; Obraztsova, A. Y.; Davidson, S. K.; Faulkner, D. J.; Haygood, M. G. Identification of the antifungal peptide-containing symbiont of the marine sponge *Theonella swinhoei* as a novel δ-proteobacterium, "*Candidatus Entotheonella palauensis*". *Marine Biology* **2000**, *136*, 969-977.

(15) Unson, M. D.; Faulkner, D. J. Cyanobacterial symbiont of chlorinated metabolites from *Dysidea herbacea* (Porifera). *Experientia* **1993**, *49*, 349-352.

(16) Demain, A. L. Small bugs, big business: the economic power of the microbe. *Biotechnol. Adv.* **2000**, *18*, 499-514.

(17) Kelecom, A. Secondary metabolites from marine microorganisms. *An. Acad. Bras. Cienc.* **2002**, *74*, 151-170.

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

(19) Omura, S.; Ikeda, H.; Ishikawa, J.; Hanamoto, A.; Takahashi, C.; Shinose, M.; Takahashi, Y.; Horikawa, H.; Nakazawa, H.; Osonoe, T.; Kikuchi, H.; Shiba, T.; Sakaki, Y.; Hattori, M. Genome sequence of an industrial microorganism *Streptomyces avermitilis*: deducing the ability of producing secondary metabolites. *Proc. Natl. Acad. Sci. U S A* **2001**, *98*, 12215-12220.

(20) Gillespie, D. E.; Brady, S. F.; Bettermann, A. D.; Cianciotto, N. P.; Liles, M. R.; Rondon, M. R.; Clardy, J.; Goodman, R. M.; Handelsman, J. Isolation of antibiotics turbomycin a and B from a metagenomic library of soil microbial DNA. *Appl. Environ. Microbiol.* **2002**, *68*, 4301-4306.

(21) Zabriskie, T. M.; Klocke, J.; Ireland, C. M.; Marcus, A.; Molinski, T.; Faulkner, J.; Xu, C.; Clardy, J. Jaspamide, a Modified Peptide from a Jaspis Sponge, with Insecticidal and Antifungal Activity. *J. Am. Chem. Soc.* **1986**, *108*, 3123-3124.

(22) Crews, P.; Manes, L.; Boehler, M. Jasplakinolide, A Cyclopepsipeptide from the Marine Sponge, *Jaspis* sp. *Tetrahedron Lett.* **1986**, *27*, 2797-2800.

(23) Zampella, A.; Giannini, C.; Debitus, C.; Roussakis, C.; D'Auria, M. V. New jaspamide derivatives from the marine sponge *Jaspis splendans* collected in Vanuatu. *J. Nat. Prod.* **1999**, *62*, 332-334.

(24) Senderowicz, A. M.; Kaur, G.; Sainz, E.; Laing, C.; Inman, W. D.; Rodriguez, J.; Crews, P.; Malspeis, L.; Grever, M. R.; Sausville, E. A.; et al. Jasplakinolide's inhibition of the growth of prostate carcinoma cells in vitro with disruption of the actin cytoskeleton. *J. Natl. Cancer Inst.* **1995**, *87*, 46-51.

(25) Takeuchi, H.; Ara, G.; Sausville, E. A.; Teicher, B. Jasplakinolide: interaction with radiation and hyperthermia in human prostate carcinoma and Lewis lung carcinoma. *Cancer Chemother. Pharmacol.* **1998**, *42*, 491-496.

(26) Bubb, M. R.; Senderowicz, A. M.; Sausville, E. A.; Duncan, K. L.; Korn, E. D. Jasplakinolide, a cytotoxic natural product, induces actin polymerization and competitively inhibits the binding of phalloidin to F-actin. *J. Biol. Chem.* **1994**, *269*, 14869-14871.

(27) Holzinger, A.; Meindl, U. Jasplakinolide, a novel actin targeting peptide, inhibits cell growth and induces actin filament polymerization in the green alga Micrasterias. *Cell. Motil. Cytoskeleton* **1997**, *38*, 365-372.

(28) Odaka, C.; Sanders, M. L.; Crews, P. Jasplakinolide induces apoptosis in various transformed cell lines by a caspase-3-like protease-dependent pathway. *Clin. Diagn. Lab. Immunol.* **2000**, *7*, 947-952.

(29) Spector, I.; Braet, F.; Shochet, N. R.; Bubb, M. R. New anti-actin drugs in the study of the organization and function of the actin cytoskeleton. *Microsc. Res. Tech.* **1999**, *47*, 18-37.

(30) Bubb, M. R.; Spector, I.; Beyer, B. B.; Fosen, K. M. Effects of jasplakinolide on the kinetics of actin polymerization. An explanation for certain *in vivo* observations. *J. Biol. Chem.* **2000**, *275*, 5163-5170.

(31) Rotsch, C.; Radmacher, M. Drug-induced changes of cytoskeletal structure and mechanics in fibroblasts: an atomic force microscopy study. *Biophys. J.* **2000**, *78*, 520-535.

(32) Lee, E.; Shelden, E. A.; Knecht, D. A. Formation of F-actin aggregates in cells treated with actin stabilizing drugs. *Cell. Motil. Cytoskeleton* **1998**, *39*, 122-133.

(33) Holzinger, A. Jasplakinolide. An actin-specific reagent that promotes actin polymerization. *Methods Mol. Biol.* **2001**, *161*, 109-120.

(34) Talpir, R.; Benayahu, Y.; Kashman, Y.; Pannell, L.; Schleyer, M. Hemiasterlin and Geodiamolide TA; Two New Cytotoxic Peptides from the Marine Sponge *Hemiasterella minor*. *Tetrahedron Lett.* **1994**, *35*, 4453-4456.

(35) Chevallier, C.; Richardson, A. D.; Edler, M. C.; Hamel, E.; Harper, M. K.; Ireland, C. M. A new cytotoxic and tubulin-interactive milnamide derivative from a marine sponge *Cymbastela sp. Org. Lett.* **2003**, *5*, 3737-3739.

(36) Akee, R.; Cartner, L.; McCloud, T.; Colin, P.; Newman, D. Distribution and Preparative Isolation of Jaspamide, a Marine Anticancer Cyclic Peptide. *American Society of Pharmacognosy- 38th Annual Meeting-July 26-30*: Iowa City, IA, 1997.

(37) Crews, P.; Farias, J.; Emrich, R.; Keifer, P. Milnamide A, an Unusual Cytotoxic Tripeptide from the Marine Sponge *Auletta cf. constricta*. *J. Org. Chem.* **1994**, *59*, 2932-2934.

(38) Braekman, J. C.; Daloze, D.; Moussiaux, B.; Riccio, R. Jaspamide from the Marine Sponge *Jaspis johnstoni*. *J. Natural Products* **1987**, *50*, 994-995.

(39) Chan, W.; Tinto, W. F.; Manchand, P. S.; Todaro, L. J. Stereostructrues of Geodiamolides A and B, Novel Cyclodepsipeptides from the Marine Sponge *Geodia sp. J. Org. Chem.* **1987**, *52*, 3091-3097.

(40) Dilip de Silva, E.; Andersen, R. J.; Allen, T. Geodiamolides C to F, New Cytotoxic Cyclodepsipeptides from the Marine Sponge *Pseudaxinyssa sp.* *Tetrahedron Lett.* **1990**, *31*, 489-492.

(41) Coleman, J. E.; Van Soest, R.; Andersen, R. J. New Geodiamolides from the Sponge *Cymbastela sp.* Collected in Papau New Guinea. *J. Nat. Prod.* **1999**, *62*, 1137-1141.

(42) D'Auria, M. V.; Gomez Paloma, L.; Minale, L.; Zampella, A.; Debitus, C.; Perez, J. Neosiphoniamolide A, a novel cyclodepsipeptide, with antifungal activity from the marine sponge *Neosiphonia superstes*. *J. Nat. Prod.* **1995**, *58*, 121-123.

(43) Jansen, R.; Kunze, B.; Reichenbach, H.; Hofle, G. Chondramides A-D, New Cytostatic and Antifungal Cyclodepsipeptides from *Chondromyces crocatus*

(Myxobacteria): Isolation and Structure Elucidation. *Liebigs. Ann.* **1996**, 285-290.

(44) Schwarzer, D.; Marahiel, M. A. Multimodular biocatalysts for natural product assembly. *Naturwissenschaften* **2001**, *88*, 93-101.

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

(46) Stachelhaus, T.; Marahiel, M. A. Modular structure of genes encoding multifunctional peptide synthetases required for non-ribosomal peptide synthesis. *FEMS Microbiol. Lett.* **1995**, *125*, 3-14.

(47) Stachelhaus, T.; Marahiel, M. A. Modular structure of peptide synthetases revealed by dissection of the multifunctional enzyme GrsA. *J. Biol. Chem.* **1995**, *270*, 6163-6169.

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

(49) Weber, G.; Schorgendorfer, K.; Schneider-Scherzer, E.; Leitner, E. The peptide synthetase catalyzing cyclosporine production in *Tolypocladium niveum* is encoded by a giant 45.8-kilobase open reading frame. *Curr. Genet.* **1994**, *26*, 120-125.

(50) Cane, D. E.; Walsh, C. T.; Khosla, C. Harnessing the biosynthetic code: combinations, permutations, and mutations. *Science* **1998**, *282*, 63-68.

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

(52) Walsh, C. T. Polyketide and nonribosomal peptide antibiotics: modularity and versatility. *Science* **2004**, *303*, 1805-1810.

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

(54) Linne, U.; Marahiel, M. A. Control of directionality in nonribosomal peptide synthesis: role of the condensation domain in preventing misinitiation and timing of epimerization. *Biochemistry* **2000**, *39*, 10439-10447.

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

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

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

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

(59) Clugston, S. L.; Sieber, S. A.; Marahiel, M. A.; Walsh, C. T. Chirality of peptide bond-forming condensation domains in nonribosomal peptide synthetases: the C5 domain of tyrocidine synthetase is a (D)C(L) catalyst. *Biochemistry* **2003**, *42*, 12095-12104.

(60) Belshaw, P. J.; Walsh, C. T.; Stachelhaus, T. Aminoacyl-CoAs as probes of condensation domain selectivity in nonribosomal peptide synthesis. *Science* **1999**, *284*, 486-489.

(61) Keating, T. A.; Ehmann, D. E.; Kohli, R. M.; Marshall, C. G.; Trauger, J. W.; Walsh, C. T. Chain termination steps in nonribosomal peptide synthetase assembly lines: directed acyl-S-enzyme breakdown in antibiotic and siderophore biosynthesis. *Chembiochem* **2001**, *2*, 99-107.

(62) Miller, D. A.; Luo, L.; Hillson, N.; Keating, T. A.; Walsh, C. T. Yersiniabactin synthetase: a four-protein assembly line producing the nonribosomal peptide/polyketide hybrid siderophore of *Yersinia pestis*. *Chem. Biol.* **2002**, *9*, 333-344.

(63) Becker, J. E.; Moore, R. E.; Moore, B. S. Cloning, sequencing, and biochemical characterization of the nostocyclopeptide biosynthetic gene cluster: molecular basis for imine macrocyclization. *Gene* **2004**, *325*, 35-42.

(64) Haese, A.; Schubert, M.; Herrmann, M.; Zocher, R. Molecular characterization of the enniatin synthetase gene encoding a multifunctional enzyme catalysing *N*-methyldepsipeptide formation in *Fusarium scirpi*. *Mol. Microbiol.* **1993**, *7*, 905-914.

(65) Schauwecker, F.; Pfennig, F.; Grammel, N.; Keller, U. Construction and in vitro analysis of a new bi-modular polypeptide synthetase for synthesis of *N*-methylated acyl peptides. *Chem. Biol.* **2000**, *7*, 287-297.

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

(67) Cane, D. E.; Walsh, C. T. The parallel and convergent universes of polyketide synthases and nonribosomal peptide synthetases. *Chem. Biol.* **1999**, *6*, R319-325.

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

(69) Jez, J. M.; Noel, J. P. Mechanism of chalcone synthase. pKa of the catalytic cysteine and the role of the conserved histidine in a plant polyketide synthase. *J. Biol. Chem.* **2000**, *275*, 39640-39646.

(70) Khosla, C. Harnessing the Biosynthetic Potential of Modular Polyketide Synthases. *Chem. Rev.* **1997**, *97*, 2577-2590.

(71) Donadio, S.; McAlpine, J. B.; Sheldon, P. J.; Jackson, M.; Katz, L. An erythromycin analog produced by reprogramming of polyketide synthesis. *Proc. Natl. Acad. Sci. U S A* **1993**, *90*, 7119-7123.

(72) Donadio, S.; Staver, M. J.; McAlpine, J. B.; Swanson, S. J.; Katz, L. Modular organization of genes required for complex polyketide biosynthesis. *Science* **1991**, *252*, 675-679.

(73) Staunton, J.; Caffrey, P.; Aparicio, J. F.; Roberts, G. A.; Bethell, S. S.; Leadlay, P. F. Evidence for a double-helical structure for modular polyketide synthases. *Nat. Struct. Biol.* **1996**, *3*, 188-192.

(74) Sieber, S. A.; Linne, U.; Hillson, N. J.; Roche, E.; Walsh, C. T.; Marahiel, M. A. Evidence for a monomeric structure of nonribosomal Peptide synthetases. *Chem. Biol.* **2002**, *9*, 997-1008.

(75) Kao, C. M.; Pieper, R.; Cane, D. E.; Khosla, C. Evidence for two catalytically independent clusters of active sites in a functional modular polyketide synthase. *Biochemistry* **1996**, *35*, 12363-12368.

(76) Gokhale, R. S.; Lau, J.; Cane, D. E.; Khosla, C. Functional orientation of the acyltransferase domain in a module of the erythromycin polyketide synthase. *Biochemistry* **1998**, *37*, 2524-2528.

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

(78) Katz, L. Manipulation of Modular Polyketide Synthases. *Chem. Rev.* **1997**, *97*, 2557-2576.

(79) Donadio, S.; Staver, M. J.; McAlpine, J. B.; Swanson, S. J.; Katz, L. Biosynthesis of the erythromycin macrolactone and a rational approach for producing hybrid macrolides. *Gene* **1992**, *115*, 97-103.

(80) Reynolds, K. A. Combinatorial biosynthesis: lesson learned from nature. *Proc. Natl. Acad. Sci. U S A* **1998**, *95*, 12744-12746.

(81) Hopwood, D. A.; Malpartida, F.; Kieser, H. M.; Ikeda, H.; Duncan, J.; Fujii, I.; Rudd, B. A.; Floss, H. G.; Omura, S. Production of 'hybrid' antibiotics by genetic engineering. *Nature* **1985**, *314*, 642-644.

(82) Dutton, C. J.; Gibson, S. P.; Goudie, A. C.; Holdom, K. S.; Pacey, M. S.; Ruddock, J. C.; Bu'Lock, J. D.; Richards, M. K. Novel avermectins produced by mutational biosynthesis. *J. Antibiot. (Tokyo)* **1991**, *44*, 357-365.

(83) Kawashima, A.; Seto, H.; Ishiyama, T.; Kato, M.; Uchida, K.; Otake, N. Fluorinated blasticidin S. *Agric. Biol. Chem.* **1987**, *51*, 1183-1184.

(84) Denoya, C. D.; Fedechko, R. W.; Hafner, E. W.; McArthur, H. A.; Morgenstern, M. R.; Skinner, D. D.; Stutzman-Engwall, K.; Wax, R. G.; Wernau, W. C. A second branched-chain alpha-keto acid dehydrogenase gene cluster (bkdFGH) from *Streptomyces avermitilis*: its relationship to avermectin biosynthesis and the construction of a bkdF mutant suitable for the production of novel antiparasitic avermectins. *J. Bacteriol.* **1995**, *177*, 3504-3511.

(85) Jacobsen, J. R.; Hutchinson, C. R.; Cane, D. E.; Khosla, C. Precursor-directed biosynthesis of erythromycin analogs by an engineered polyketide synthase. *Science* **1997**, *277*, 367-369.

(86) Jacobsen, J. R.; Keatinge-Clay, A. T.; Cane, D. E.; Khosla, C. Precursor-directed biosynthesis of 12-ethyl erythromycin. *Bioorg. Med. Chem. Lett.* **1998**, *6*, 1171-1177.

(87) Admiraal, S. J.; Khosla, C.; Walsh, C. T. The loading and initial elongation modules of rifamycin synthetase collaborate to produce mixed aryl ketide products. *Biochemistry* **2002**, *41*, 5313-5324.

(88) Katz, L.; Donadio, S. Polyketide synthesis: prospects for hybrid antibiotics. *Annu. Rev. Microbiol.* **1993**, *47*, 875-912.

(89) Xue, Q.; Ashley, G.; Hutchinson, C. R.; Santi, D. V. A multiplasmid approach to preparing large libraries of polyketides. *Proc. Natl. Acad. Sci. U S A* **1999**, *96*, 11740-11745.

(90) McDaniel, R.; Thamchaipenet, A.; Gustafsson, C.; Fu, H.; Betlach, M.; Ashley, G. Multiple genetic modifications of the erythromycin polyketide synthase to produce a library of novel "unnatural" natural products. *Proc. Natl. Acad. Sci. U S A* **1999**, *96*, 1846-1851.

(91) Eppelmann, K.; Stachelhaus, T.; Marahiel, M. A. Exploitation of the selectivity-conferring code of nonribosomal peptide synthetases for the rational design of novel peptide antibiotics. *Biochemistry* **2002**, *41*, 9718-9726.

(92) Del Vecchio, F.; Petkovic, H.; Kendrew, S. G.; Low, L.; Wilkinson, B.; Lill, R.; Cortes, J.; Rudd, B. A.; Staunton, J.; Leadlay, P. F. Active-site residue, domain and module swaps in modular polyketide synthases. *J. Ind. Microbiol. Biotechnol.* **2003**, *30*, 489-494.

(93) Cropp, T. A.; Kim, B. S.; Beck, B. J.; Yoon, Y. J.; Sherman, D. H.; Reynolds, K. A. Recent developments in the production of novel polyketides by combinatorial biosynthesis. *Biotechnol. Genet. Eng. Rev.* **2002**, *19*, 159-172.

(94) Cortes, J.; Wiesmann, K. E.; Roberts, G. A.; Brown, M. J.; Staunton, J.; Leadlay, P. F. Repositioning of a domain in a modular polyketide synthase to promote specific chain cleavage. *Science* **1995**, *268*, 1487-1489.

(95) Kao, C. M.; Luo, G.; Katz, L.; Cane, D. E.; Khosla, C. Manipulation of Macrolide Ring Size by Directed Mutagenesis of a Modular Polyketide Synthase. *J. Am. Chem. Soc.* **1995**, *117*, 9105-9106.

(96) de Ferra, F.; Rodriguez, F.; Tortora, O.; Tosi, C.; Grandi, G. Engineering of peptide synthetases. Key role of the thioesterase-like domain for efficient production of recombinant peptides. *J. Biol. Chem.* **1997**, *272*, 25304-25309.

(97) Kohli, R. M.; Trauger, J. W.; Schwarzer, D.; Marahiel, M. A.; Walsh, C. T. Generality of peptide cyclization catalyzed by isolated thioesterase domains of nonribosomal peptide synthetases. *Biochemistry* **2001**, *40*, 7099-7108.

(98) Trauger, J. W.; Kohli, R. M.; Mootz, H. D.; Marahiel, M. A.; Walsh, C. T. Peptide cyclization catalysed by the thioesterase domain of tyrocidine synthetase. *Nature* **2000**, *407*, 215-218.

(99) Trauger, J. W.; Kohli, R. M.; Walsh, C. T. Cyclization of backbone-substituted peptides catalyzed by the thioesterase domain from the tyrocidine nonribosomal peptide synthetase. *Biochemistry* **2001**, *40*, 7092-7098.

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

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

(102) Kuhstoss, S.; Huber, M.; Turner, J. R.; Paschal, J. W.; Rao, R. N. Production of a novel polyketide through the construction of a hybrid polyketide synthase. *Gene* **1996**, *183*, 231-236.

(103) Katz, L.; McDaniel, R. Novel macrolides through genetic engineering. *Med. Res. Rev.* **1999**, *19*, 543-558.

(104) Ruan, X.; Pereda, A.; Stassi, D. L.; Zeidner, D.; Summers, R. G.; Jackson, M.; Shivakumar, A.; Kakavas, S.; Staver, M. J.; Donadio, S.; Katz, L. Acyltransferase domain substitutions in erythromycin polyketide synthase yield novel erythromycin derivatives. *J. Bacteriol.* **1997**, *179*, 6416-6425.

(105) Oliynyk, M.; Brown, M. J.; Cortes, J.; Staunton, J.; Leadlay, P. F. A hybrid modular polyketide synthase obtained by domain swapping. *Chem. Biol.* **1996**, *3*, 833-839.

(106) McDaniel, R.; Kao, C. M.; Hwang, S. J.; Khosla, C. Engineered intermodular and intramodular polyketide synthase fusions. *Chem. Biol.* **1997**, *4*, 667-674.

(107) Holzbaur, I. E.; Ranganathan, A.; Thomas, I. P.; Kearney, D. J.; Reather, J. A.; Rudd, B. A.; Staunton, J.; Leadlay, P. F. Molecular basis of Celmer's rules: role of the ketosynthase domain in epimerisation and demonstration that ketoreductase domains can have altered product specificity with unnatural substrates. *Chem. Biol.* **2001**, *8*, 329-340.

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

(109) Duerfahrt, T.; Doekel, S.; Sonke, T.; Quaedflieg, P. J.; Marahiel, M. A. Construction of hybrid peptide synthetases for the production of alpha-l-aspartyl-l-phenylalanine, a precursor for the high-intensity sweetener aspartame. *Eur. J. Biochem.* **2003**, *270*, 4555-4563.

(110) Doekel, S.; Marahiel, M. A. Dipeptide formation on engineered hybrid peptide synthetases. *Chem. Biol.* **2000**, *7*, 373-384.

(111) Doumith, M.; Legrand, R.; Lang, C.; Salas, J. A.; Raynal, M. C. Interspecies complementation in *Saccharopolyspora erythraea* : elucidation of the function of oleP1, oleG1 and oleG2 from the oleandomycin biosynthetic gene cluster of *Streptomyces antibioticus* and generation of new erythromycin derivatives. *Mol. Microbiol.* **1999**, *34*, 1039-1048.

(112) Gaisser, S.; Reather, J.; Wirtz, G.; Kellenberger, L.; Staunton, J.; Leadlay, P. F. A defined system for hybrid macrolide biosynthesis in *Saccharopolyspora erythraea*. *Mol. Microbiol.* **2000**, *36*, 391-401.

(113) Betlach, M. C.; Kealey, J. T.; Ashley, G. W.; McDaniel, R. Characterization of the macrolide P-450 hydroxylase from *Streptomyces venezuelae* which converts narbomycin to picromycin. *Biochemistry* **1998**, *37*, 14937-14942.

(114) Gaisser, S.; Lill, R.; Staunton, J.; Mendez, C.; Salas, J.; Leadlay, P. F. Parallel pathways for oxidation of 14-membered polyketide macrolactones in *Saccharopolyspora erythraea*. *Mol. Microbiol.* **2002**, *44*, 771-781.

(115) Hildebrand, M.; Waggoner, L. E.; Lim, G. E.; Sharp, K. H.; Ridley, C. P.; Haygood, M. G. Approaches to identify, clone, and express symbiont bioactive metabolite genes. *Nat. Prod. Rep.* **2004**, *21*, 122-142.

Isolation of Natural Product Biosynthesis Genes from Symbiotic Marine Organisms

Chapter 2

Cloning a Mixed NRPS/PKS Gene Cluster from *J. splendens*

Introduction

In the broadest sense, the overall goal of this project was to develop a method for the identification, cloning, and expression of biosynthetic gene clusters from symbiotic marine organisms. In order to develop such a method, our initial aim was to identify the jaspamide biosynthetic gene cluster. Several specific aims required for the success of this objective include: isolation of high quality genomic DNA from the complex sponge/microbe assemblage, design of degenerate primers for the amplification of NRPS and PKS gene probes, generation and screening of a heterogenomic library, identification and sequencing of potential jaspamide gene clusters, and the expression of all or part of the identified gene clusters in order to provide proof of function. The first four objectives outlined above are the focus of this chapter.

One of the features that made jaspamide a good initial target compound is its suspected microbial origin. This is because several features of eukaryotic genes make them less suitable for the cloning of biosynthetic gene clusters. First, unlike bacterial secondary metabolism, biosynthetic genes are often not clustered in eukaryotic organisms. The lack of gene clustering may require the individual isolation of each biosynthetic gene involved in the pathway. This process can not only be quite tedious, but also requires accurate prediction of the types of enzymes likely to be involved in the pathway if a degenerate PCR method is to be applied. In addition, the sequence of sufficient gene homologs would need to be available to allow for the design of degenerate PCR primers. This approach makes it more difficult to identify novel biosynthetic genes that are not represented in public databases.

An alternate approach is to use random mutagenesis to generate libraries of the producing organism with single gene disruptions. The mutants can then be characterized

based on their ability to produce the compound of interest or related analogs.[1] Sequencing of the mutation can then be used to identify genes that are involved in the biosynthetic pathway. One disadvantage of this approach is the likelihood that a variety of non-biosynthetic genes could also disrupt antibiotic biosynthesis. Genes involved in functions such as primary metabolism, cell signaling and regulation of secondary metabolism could all result in a similar phenotype.[2,3] Consequently, it would be necessary to analyze a very large number of mutants in order to identify potential antibiotic biosynthetic genes.

Another hurdle that must be overcome in the isolation of biosynthetic genes in eukaryotes is the presence of introns. The inability to predict the presence, number and size of introns in the target gene makes it difficult to amplify unknown genes using PCR based approaches because the size of the desired product can not be accurately predicted. Consequently, isolation of eukaryotic biosynthetic genes requires generation of cDNA libraries from mRNA. Although methods for the production of cDNA libraries are well established, the success of this approach relies on the expression of the target gene at the time that the mRNA is harvested. This represents an additional problem when working with unculturable organisms where it is difficult to study the timing of metabolite production. The presence of the desired metabolite may not always correlate to expression of the metabolic genes and may simply reflect the stability and storage potential of the desired compound. In these situations, it would be easy to miss the expression of the target biosynthetic genes.

The process of identifying antibiotic biosynthetic genes from prokaryotic organisms does not share many of these limitations. The lack of introns in bacterial genes allows for the generation of gene probes by PCR, as the product size can be readily predicted based on the sequence of similar, known genes. In addition, it is common for related bacterial biosynthetic genes to be localized in the same region of the chromosome forming gene clusters.[4] This feature is often taken advantage of in the search for bacterial biosynthetic genes.[5,6] Oligonucleotide probes can be designed to identity specific genes that are believed to be involved in biosynthesis of the target antibiotic, such as NRPS or PKS genes. Analysis of the chromosomal region around these genes can then be used to

identify other genes in the cluster. In this way, it is possible to identify novel genes that code for enzymes with unique activities and that code for unusual chemical transformations. The first successful use of sequence homology to identify new biosynthetic genes and pathways was reported by Malpartida in the successful use of actinorhodin genes to hybridize to genomic DNA of other polyketide antibiotic producing *Streptomyces*.[7]

Based on the suspected microbial origin of jaspamide, a genomic DNA based approach was developed for the isolation of the biosynthetic gene cluster. This approach required isolation of metagenomic DNA, amplification of specific NRPS and PKS gene probes, screening a metagenomic library and sequencing of a candid gene cluster. Using the methods described below, a mixed PKS/NRPS gene cluster was cloned and sequenced from an unculturable *Jaspis* sponge-microbe assemblage.

Results

Isolation of Metagenomic DNA from Jaspis splendens

The first step in the cloning of marine biosynthetic gene clusters is the purification of high quality genomic DNA (gDNA). Several methods of gDNA preparation were tried in order to find optimal conditions for the isolation of high molecular weight gDNA from marine organisms. In addition to *J. splendens*, gDNA was isolated from a variety of other marine organisms, including *Geodia* and *Phorbas* sponges, cyanobacteria, and tunicates in order to determine if the developed method could be broadly applied.

One of the primary issues to be addressed was the identification of optimal storage conditions for collected marine tissue samples. The best method to store the tissue samples at the time of collection in order to preserve the integrity and quality of the gDNA needed to be determined. gDNA was prepared from *J. splendens* samples, shown to contain jaspamide, that had been stored under a variety of conditions. Two samples were large chunks of sponge (>3 cm^3) that had been frozen at -20 °C or -80 °C for three years. A second set of tissue samples were freshly collected and had been stored in RNAlater™ at -20 °C as either large chunks, thin slices, or minced tissue. RNAlater™ is

a solution marketed by Ambion that is used to preserve mRNA in tissue samples. The main reason for the use of this solution was to preserve the mRNA in case a cDNA approach to identify genes was needed. The use of this solution may aid in the preservation of DNA, as well as the mRNA, in the sample.

Initial studies indicated that the highest quality gDNA was isolated from sponge tissue that had been minced prior to storage in RNAlater. After determining the optimal storage conditions for the marine tissue samples, the next objective was to determine the best method of isolating high molecular weight gDNA.

The first method used to isolate gDNA from marine samples was based on a protocol to isolate gDNA from terrestrial *Streptomyces*.[8] Although gDNA could be isolated by this method, the yield was quite low. Another disadvantage of this method is the use of phenol/chloroform/isoamyl alcohol washes. Care needed to be taken in the decanting of the gDNA solution from these washes as contaminating phenol in the gDNA prep can interfere with restriction enzymes and DNA polymerase used in subsequent experiments. A survey of several commercially available gDNA isolation kits was conducted to determine if they could be used to acquire adequate amounts of quality gDNA from these samples.

Three genomic DNA purification kits were tried in parallel to determine which protocol produces the best yields and quality gDNA. The kits that evaluated were: Promega's Wizard Genomic DNA Purification Kit, BioRad's Quantum Prep AquaPure Genomic DNA Isolation Kit, and Gentra Systems' Genomic DNA Purification Kit for yeast and Gram Positive Bacteria. Initially two isolation protocols were tried for each kit, according to manufacturer's directions. In all cases sponge tissue was frozen in liquid nitrogen and ground by hand with a microfuge tube pestle until the tissue was the consistency of a thick paste. The first protocol used for each was that described for the isolation of genomic DNA from Gram positive bacteria. The second protocol evaluated for each of the three kits was for the isolation of genomic DNA from mouse tail. The second protocol was chosen because it was the only protocol available for the isolation of DNA from animal tissue. It was thought that the overnight digestion of the ground tissue in proteinase K used in this method may help lyse any endosymbiotic cells that may be

imbedded within the protective sponge tissue. The methods were modified from the manufacturer's protocols only in the nature of the sample from which the DNA was isolated (ground sponge tissue instead of pelleted bacterial cells or mouse tail).

Initial results indicated that the Promega kit, following the protocol for isolation of DNA from mouse tail, produced the greatest quantity and highest quality of DNA from tissue samples stored at -80 °C in the absence of RNAlater™. However, for samples that were preserved in RNAlater™, the PureGene kit produced the greatest quantity, best quality, and highest molecular weight of DNA from sponge tissue (Figure 2.1).

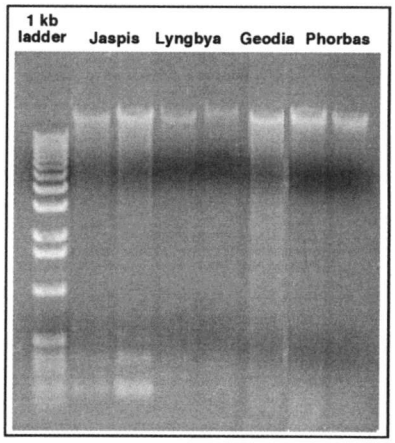


Figure 2.1. Metagenomic DNA isolated from marine organisms

Other factors were also found to have an impact on the quality and quantity of gDNA isolated. First, it was found that scaling up the protocol often resulted in the degradation of the genomic DNA. The highest yield (ng DNA/mg tissue sample) was found when gDNA was purified from 30-40 mg tissue samples. Use of larger tissue samples was found to result not only in a lower yield, but also in a higher degree of degradation. It was also observed that the highest yields of gDNA were obtained when the tissue sample was thoroughly homogenized prior to DNA abstraction. The difficulty of homogenizing the tissue samples varied between the different marine organisms tested. The *J. splendens* tissue was relatively easy to homogenize and could be accomplished by first mincing the tissue with microscissors and then homogenizing with a microfuge tube pestle. Other sponge samples were not as easily homogenized. This included the *Phorbas* sponge samples, which due to a large amount of silica present in the tissue, could not be minced with the microscissors and required extensive grinding with the microfuge tube pestle. Another factor that appeared to have an impact on the quality of DNA was the source organism(s). It was observed that the gDNA from some organisms, such as *Geodia* sponge samples, were more susceptible to degradation during purification.

Based on these studies, it was determined that the PureGene gDNA purification kit using the protocol for gram positive bacteria gave us the highest yield of consistently high quality DNA. Additionally, we found that storage of marine tissue samples in RNAlater for several years did not have an adverse effect on the quality of gDNA isolated using the PureGene protocol.

Amplification of NRPS and PKS Gene Probes by PCR

Design of PCR Primers.

A second issue that must be considered is the design and choice of degenerate PCR primers for the amplification of gene probes form the metagenomic DNA. A common approach to cloning biosynthetic gene clusters is the use of degenerate PCR primers designed to amplify regions within specific types of genes. This approach takes advantage of the high degree of sequence conservation often found within enzyme families. Both NRPSs and PKSs contain multiple highly conserved sequence motifs that can be used for the design of degenerate PCR primers to amplify gene probes. These gene probes can be used to screen metagenomic libraries constructed from the producing organisms.

A variety of NRPS and PKS PCR primers were used to optimize the amplification of gene probes from the metagenomic DNA. Primers used in this work are shown in Table 2.1. The sequence of several primers were taken from the literature and had been used successfully to amplify NRPS gene clusters from other organisms.[9-11] Additional primers were designed to take advantage of other conserved sequence motifs located within the NRPS A-domain. Several new primers were developed in order to increase the likelihood of hybridizing to GC rich gene sequences. We wanted to bias one set of primers toward high GC organisms due to the suspected identity of the jaspamide producer. As mentioned previously, the structure of jaspamide is very similar to the myxobacterial metabolite chondramide. One hypothesis is that marine bacteria related to the *Myxobacteria* may be the true source of jaspamide. Like *Streptomyces*, *Myxobacteria* have DNA with a high percentage of GC.[12] We wanted to use degenerate primers that would have a higher probability of hybridizing to genes from this type of organism.

Streptomyces NRPS genes were used for primer design because they typically have high GC content and more *Streptomyces* NRPS gene clusters are available in the public databases than myxobacterial NRPS clusters. Sequence alignments of several *Streptomyces* NRPS genes were used to identify a consensus sequence in order to design degenerate primers from these highly conserved regions.

A second set of gene probes was needed to identify type I PKS genes. Degenerate primers for type I PKS used in this work were published by Beyer *et al.* and had been used for the amplification of PKS from myxobacteria.[13]

Table 2.1. Degenerate PCR Primers used in this work. Degeneracy is indicated as follows: M is A or C; I is inosine; R is G or A; H is A,C, or T; W is A or T; Y is T or C; N is A,G,C or T; K is G or T; and S is G or C.

ORGANISM	PRIMER NAME	SEQUENCE	REFERENCE
A2 (CORE I)		**L K A G G A Y V P I D P**	
Streptomyces	PSI	CTS AAG GCS GGC GGI KMC TAC GTS CCS IAC CC	M.Cone,unpub.
Bacillis subtilis	N/A	TTI AAR GCR GGY GGI GCI TAT GTG CCG ATY GAY CC	Borchert,1992
P. fluorescens	N/A	CTC AAG GCG GGC GGI GCC TAC GTG CCC IAC CC	Rajendran,1999
Cyanobacteria sp.	LGA2F1	GCN GGY GGY GCN TAY GTN CC	Neilan, 1999
A3 (Core II)		**I Y T S G T T G K P K G**	
Bacillis subtilis	N/A	ATI TAY ACI TCH GGI ACI ACA GGI AAG CCA AAA GG	Borchert,1992
Streptomyces	PSII-IVNT	ATC TAC ACS TCS GGC ACS ACS GGC AAG CCS AAG GG	
Streptomyces	PSIIR	CTT SGG CIK SCC GGI SGI SCC CGA GGT GTA GA	M.Cone,unpub.
P. fluorescens	N/A	CTT GGG CIG GCC GGI CGI CCC GGA GGT GTA GA	Rajendran,1999
A7 (Core IV)		**M/L Y R (ST) G D**	
Streptomyces	LGA7F1	AGT TAC GCG ASC GGC GAC	This work
Streptomyces	LGA7F2	CTG TAC CGC AGC GGC GAC	This work
A8 (Core V)		**Q V K I R G**	
Cyanobacteria sp.	LGA8R1	CCN CGD ATY TTN ACY TG	Neilan, 1999
Streptomyces	PS5R3	SCG GWA SCC GCG SAY CTT SAC CTG	M.Cone, unpub.
T (Core VI)		**X S X G G L X F F**	
Streptomyces	LGTR1	ATS GAG TCG CCG CCS AGC	This work
Streptomyces	LGTR2	ATC GAG KCG CCG CCC AGS	This work
Bacillis subtilis	LGTR3	CGA GTG KCC GCC GAT GGA	This work
Streptomyces	PSII-VICT	AW IGA CKS ICC ICC SRR SIM GAA GAA	
PKS KS primers	KS1Up	MIGGARGCIHWISMIAGTGAYCCICARCAI	Beyer,S. 1999
	KSD1	GGRTCICCIARISWIGTICCIGTICCRTG	

Amplification of NRPS and PKS Gene Probes by PCR

After a sufficient amount of high quality gDNA had been isolated from *J. splendens*, it was used to amplify gene probes by PCR. Prior to being used in PCR, the isolated genomic DNA was partially digested with *Sau*3A or mechanically sheared by passing through a 20 gauge syringe needle in order to generate smaller fragments and

facilitate improved hybridization. The PCR conditions used for the amplification of gene probes needed to be optimized to obtain good yields while maintaining specificity. Polyketide synthase (PKS) probes specific to the ketosynthase domain were amplified easily with little optimization of reaction conditions required (Figure 2.2 A). Upon cloning and sequencing, these probes were found to have high sequence similarity to known PKS genes. The closest matches in the public database to the translated sequences were MtaE (*Stigmatella aurantiaca*) PksD (*Mycobacterium leprae*) and McyD (*Microcystis aeruginosa*).[14]

Initial experiments using the NRPS primers revealed that the primer pairs LGA7F1/LGTR2, LGA2F1/LGA8R1 and LGA2F1/LGTR2 gave the most consistent amplification of NRPS genes from the metagenomic DNA. These primers amplify the region between the A7-PCP, A2-A8, and A2-PCP conserved sequence motifs of NRPS adenylation domains, respectively (Figure 2.2, B, C).

Following purification, the PCR products were cloned into pGEMT-*easy* and unique clones were identified and sequenced. In the initial round of screening, two unique clones were identified that had translated sequences with high similarity to known NRPS genes. The translated sequence of the first clone, JaspF2, had the highest similarity to cyanobacterial and myxobacterial NRPSs such as McyA (*Microcystis aeruginosa*), NosA (*Nostoc sp.*) and Mxa1 (*Myxococcus xanthus*). Likewise, the translation of the insert from the second clone, JaspF13, showed high similarity with NosC (*Nostoc sp.*), MycC (*Bacillus subtilis*) and SafA (*M. xanthus*). These results, specifically identification of products similar to myxobacterial genes, were encouraging due to the similarity of jaspamide and chondramide.

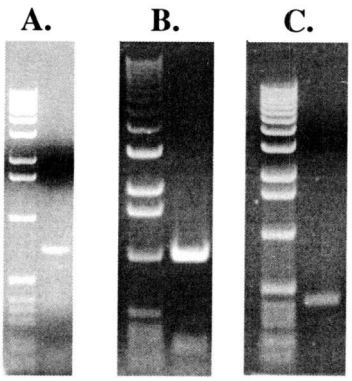


Figure 2.2. NRPS and PKS gene probes amplified from *J. splendens* metagenomic DNA. A. 0.7 kb PCR products amplified with PKS degenerate primers for conserved β-ketoacylsynthase motifs. B. 1 kb PCR products amplified with degenerate NRPS primers for A2 and A8 conserved sequence motifs. C. 0.4 kb PCR product amplified with degenerate NRPS primers for A7 and PCP conserved sequences.

In order to gain insight into the types of gene clusters that were being amplified, the substrate specificity for several of the NRPS probes was predicted using the method developed independently by Challis *et al.* and Stachelhaus *et al.*[15,16] The key selectivity conferring amino acids for the translated amplicons were identified through sequence alignment and compared to those in the binding pocket of domains with known specificity. The first unique clone, JasF2, was found to have the strictly conserved aspartate in the FDxS motif shifted to FxDS. The movement of this aspartate residue has only been seen in adenylation domains that activate β-amino acids.[15] This was of particular interest because jaspamide contains a β-tyrosine residue. The second unique clone, JasF13, most closely matched the sequence of serine activating adenylation domains.

Successful amplification of NRPS gene fragments could also be accomplished with the A7-PCP primer pair (Figure 2.2, C). The cloned PCR products from these reactions also had high sequence similarity to known NRPS genes. For example, the predicted gene product of one clone, JaspB, had significant similarity to cyanobacterial McyA (*M. aeruginosa*) and the myxobacterial SafA (*M. xanthus*). The substrate specificity for the A7-PCP amplicons could not be predicted because the amplified region does not contain the substrate selectivity region between motifs A4 and A5. For that reason, A2-A8 or A2-PCP amplicons were used as gene probes.

Despite the ability of the A2-PCP primer pairs to consistently amplify NRPS genes, several unexpected results were observed. One frequently encountered problem was the amplification of multiple PCR products when the A2/PSII-IVCT primer pair was used. These primers were expected to amplify a 1.4-1.5 kb fragment, however the reactions frequently yielded four bands. The first and strongest band corresponded to 1.6 kb, a second to 1.5 kb, a third to 1.2 kb and the last band was 1.0 kb. The amplification of multiple fragments was found to be caused by the PSII-VICT reverse primer. Other researchers in our laboratory observed similar results with this primer. The amplification of additional products is believed to be due to hybridization of the primer to a second region in the A-domain, resulting in a smaller PCR product (Xihou Yin, personal communication). This would account for the two pairs of PCR products. One possible

explanation for the extra 1.6 kb product is the possible variation in the size of the A2 and T region. This may vary is size due to the presence of 'optional' NRPS domains such as an *N*-methylation domain, however *N*-methylation domains are typically ~400 bp in size. Thus, both the 1.6 kb and 1.5 kb PCR products were purified and cloned into pGEM-T *easy* for sequencing. No auxiliary domains were detected in any of the products.

An additional problem of primer specificity was encountered during the construction of a genomic library from a Fijian *J. splendens* sample. While eight unique clones were identified from A2-PCP amplicons, only two were found to have sequence similarity to known NRPS genes. One of the clones was predicted to activate a glycine residue while the other was predicted to activate an aspartate or asparagine residue. The remaining clones showed weak sequence similarity to adenylate cyclases and NADH dependent dehydrogenases.

A second round of PCR was attempted using degenerate primers LGPSII-VICT and LGPSII-VINT, for conserved sequence motifs A3 and PCP, in an attempt to improve the specificity. These primers were used to amplify an expected 1.2 kb product; however the five unique clones all had sequence similarity to ABC transporters and not NRPS modules. A consistent feature in PCR amplification from the Fijian sponge sample is that the products were quite 'dirty' and contained many non-specific bands smaller than 1.0 kb. Optimization of PCR conditions did not improve the specificity of the reactions and resulted either in the characteristic "dirty" products or no products.

In an attempt to increase the specificity of the PCR and reduce the amount of non-specific hybridization, an additional amplification step was added to the protocol. It was found that the percentage NRPS gene products could be increased by performing two rounds of PCR using "nested" PCR primers. A first round of PCR was conducted using the LGA2F1/LGTR2 or LGA2F1/LGPSII-VICT primer pairs. The purified PCR products from this first round of PCR were then used for a second round of PCR under more stringent conditions with LGA2F1 and LGA8R1 primers. The PCR products from the second round of PCR were then cloned and sequenced. The two step nested PCR method resulted in a dramatically higher percentage of NRPS positive clones.

Generation and Screening of a J. splendens Metagenomic Library

The general approach to the identification of mixed biosynthetic gene clusters from marine organisms is based on the common phenomenon of gene clustering. Many NRPS and PKS natural product gene clusters are contained on greater than 30 kb of DNA, making it is highly unlikely that a 35-45 kb cosmid insert of gDNA would contain elements from unrelated NRPS and PKS genes. In order to take advantage of this gene clustering, a metagenomic library was constructed using cosmids with 30-40 kb inserts. Hybridization of both NRPS and PKS gene probes to a single cosmid is a good indication that that cosmid contains a mixed NRPS/PKS gene cluster. Further screening of NRPS/PKS positive clones could then be used to select a candidate that is most likely to harbor the jaspamide gene cluster.

Once NRPS and PKS gene probes were amplified from *Jaspis* DNA, they were labeled for chemiluminescent detection and used to screen duplicates of a *Jaspis* metagenomic cosmid library. Two cosmid libraries were screened according to the methods of Berger and Kimmel.[17] The first library consisted of approximately 9000 clones and was generated from a *Jaspis* sample collected in Pohnpei in 1998. The second library consisted of approximately 10,000 clones and was generated from a *Jaspis* sample collected in Fiji in 2001.

Many colonies that appeared to hybridize to both PKS and NRPS probes were identified. These colonies were cultured and their cosmids were purified and subjected to additional screening by Southern analysis. Of the fifty-three original positive clones screened, four were found to hybridize to both PKS and NRPS probes. In all cases, the hybridization signal of the PKS probe was significantly weaker than that of the NRPS probe. The four positive cosmids, identified as JC1, JC2, JC3, and JC7, were digested with *Bam*HI and subjected to a final round of Southern hybridization to identify bands that hybridized to the PKS and NRPS probes, respectively.

The PKS probe was found to hybridize to the 7 kb *Bam*HI fragment found in all four clones, as well as to the 5.2 kb *Bam*HI fragment of JC2. The NRPS probe shared hybridization to the 7 kb *Bam*HI fragment and had an additional, stronger hybridization with the 9 kb *Bam*HI fragment of JC1, JC3, and JC7, as well as the 5.2 kb band of JC2.

Restriction mapping and initial sequencing of the four cosmids showed that they contained fragments of the same gene cluster. In order to identify additional cosmids that may overlap the cloned gene cluster, as well as to identify additional candidates that may harbor the jaspamide biosynthetic gene cluster, a second genomic library was generated from a freshly harvested sample of *J. splendens* collected at Beagle Pass, Fiji in 2001. When constructing the second genomic library, a fosmid vector was chosen to try and increase library size.

The construction of the fosmid library was conducted in the same manner as the cosmid library with one exception. One concern with using a single copy fosmid system was the potential to miss a positive clone due to inadequate sensitivity of the system. With the cosmid system, the presence of multiple copies of the cosmid within each bacterial colony increases the sensitivity for successful probe hybridization. With the fosmid system, however, there is only a single copy of the plasmid in each cell. This could result in decreased assay sensitivity and result in missing a potentially positive signal. For this reason, the fosmids were maintained in a single copy state during preparation of the fosmid library and induced to a high copy number just prior to cell lysis and fixing of the cosmid DNA to the nylon membrane. This was accomplished by regenerating colony lifts on LB agar plates supplemented with appropriate antibiotics and arabinose for the induction of high copy numbers. It is believed that exposure of the *E. coli* cells to arabinose that had diffused through the membrane would result in the induction of high fosmid copy number and provide adequate template for probe hybridization. Although no controls were run to determine if such an induction were occurring, no significant differences in detection sensitivity was observed between the cosmid and fosmid libraries.

The Fijian *Jaspis* fosmid library was larger that the previously screened Pohnpei library (~10,000 colonies). Although a number of clones were identified that hybridized to either the NRPS or the PKS probes, no mixed gene clusters were identified in the initial screening of the Fijian *Jaspis* fosmid library or in subsequent Southern analysis of the PKS and NRPS positive clones.

Sequencing the Mixed PKS/NRPS Gene Cluster

Cosmid Mapping

The *Bam*HI fragments for each of the four cosmids from the Pohnpei genomic library were cloned into pBluescript KS+ and used for subsequent sequencing and mapping of the cosmid inserts. A preliminary map for each of the four cosmid inserts was constructed using restriction digestion analysis. Initial digestion of the four cosmids and cloned *Bam*HI fragments suggested that clones JC1, JC3, and JC7 were overlapping cosmids with three common *Bam*HI fragments (5.7 kb, 7 kb and 9 kb). Clones JC1 and JC3 appeared to be nearly identical with the only difference being the size of the largest *Bam*HI fragment (~11 kb in JC1 and ~13kb in JC3). Similarly, JC2 was found to have the 12, 7 and 5.7 kb restriction fragments in common with the other three clones, however it also had an unique 5.2 kb *Bam*HI fragment. Cosmid JC7 was found to have the smallest insert and contained the common 5.2, 7 and 9 kb fragments in addition to 3.5 and 4.5 kb fragments.

Digestion of each of the subcloned 5.7 kb and 7 kb *Bam*HI fragments from all four clones (as well as the 9 kb *Bam*HI fragment from clones JC1, JC3, and JC7) with *Ava*I, *Bgl*II, *Eco*RI, *Eco*RV, *Hind*III, *Not*I, *Nco*I, *Sma*I, *Sal*I, and *Xba*I confirmed that they had the same restriction sites and were thus likely to be overlapping fragments. The restriction patterns for each of these digests, coupled with the restriction patterns of the parental cosmids and other subcloned fragments, were used to create a preliminary map for each of the cosmids. Several fragments, including the JC7 3.5 and 4.5 kb *Bam*HI fragments, could not be definitively mapped based on restriction analysis alone.

The initial maps constructed through restriction analysis were refined by sequencing the ends of each fragment. Sequence alignment determined that the JC7 3.5 kb fragment overlapped with the 11-13 kb *Bam*HI fragments in JC1, JC2 and JC3 locating it just upstream of the 5.7 kb *Bam*HI fragment (Figure 2.3).

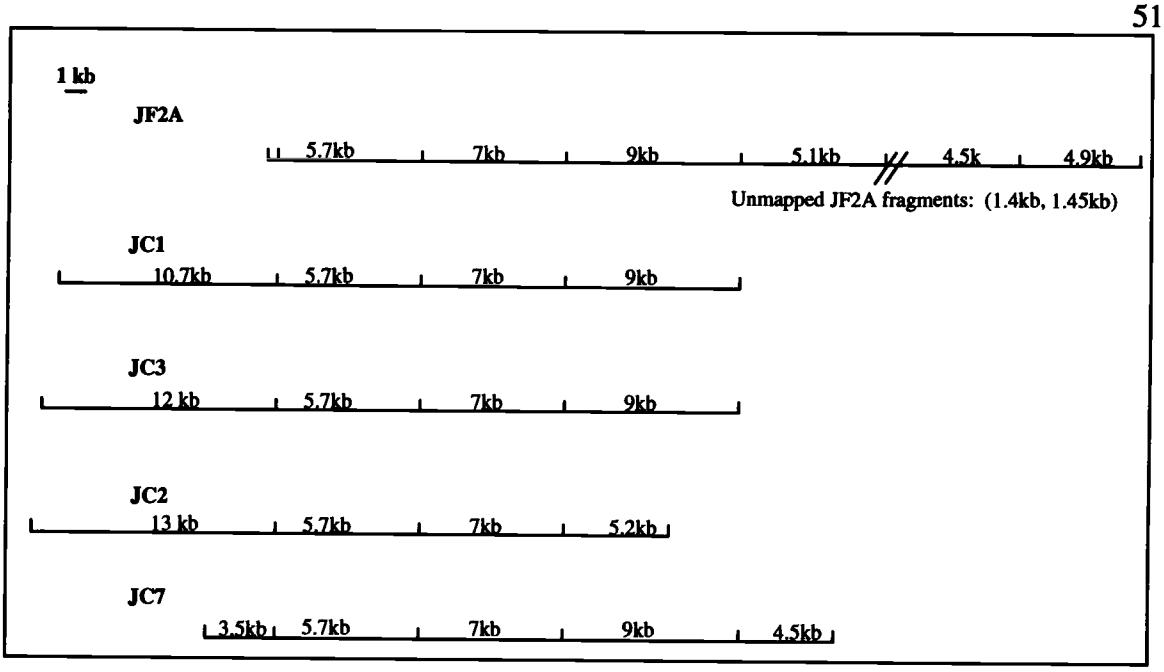


Figure 2.3. *Bam*HI restriction maps of the five overlapping cosmids/fosmid inserts containing the mixed PKS/NRPS gene cluster from *J. splendens*

The 5.2 kb *Bam*HI fragment of JC2 had been of initial interest because it strongly hybridized to labeled NRPS probes and was only found in clone JC2. This cosmid was not amendable to manipulation, however, as it exhibited propensity for instability and rearrangement. This instability was seen in attempts to subclone the JC2-5.2 kb *Bam*HI fragment. Second generation cultures of JC2-5.2 subclones produced plasmids with altered restriction patterns, presumably resulting from plasmid rearrangement. Restriction evidence, such as common 1.8 kb *Sma*I and 7.8 kb *Eco*RI fragments in digests of both JC1-9 kb and JC2-5.2 kb subclones, suggested that the JC2-5.2 kb fragment was the terminal portion of JC2 and overlapped the 9 kb fragments of JC1, JC3 and JC7. The positioning of the JC2-5.2 kb fragment was confirmed by sequencing the terminal ends of the insert off of the pBKS vector. Sequencing this fragment also proved to be quite difficult. Multiple attempts to sequence the ends of the JC2-5.2 kb *Bam*HI fragment were conducted by MWG Biotech, however no useful sequence was obtained. In a final attempt to obtain end sequence for this fragment, purified first generation plasmid was submitted for sequencing at the Center for Gene Research and Biotechnology Central Services Laboratory at Oregon State University. Fortunately the sequencing was

successful and subsequent alignment with the terminal sequence of the J7-9 subclone confirmed the common identity of these two fragments.

Verification that the 5.2 kb fragment is contained in the 9 kb fragment eliminated the need to obtain stable constructs of the former. This did not, however, explain the instability of the of the smaller 5.2 kb fragment. Plasmid instability is frequently encountered with GC rich DNA from *Streptomyces*, but the cause of the instability is not clear.[18] One possible explanation is that the observed instability was a result of the intense secondary structure found in both the 5.2 and 9 kb fragments. The structure was such that only single stranded, unidirectional sequence could be obtained for some regions of the 9 kb fragment. The secondary structure may have interfered with the replication machinery and resulted in rearrangement and loss of genetic material in the 5.2 kb fragment.

In order to verify the organization of the cosmid maps, a final round of sequencing was performed to confirm the location of the *Bam*HI restriction sites and the junctions between the various *Bam*HI fragments. These sequencing reads, which used cosmid template, confirmed the organization of the four cosmid maps.

Sequencing the Mixed PKS/NRPS Gene Cluster

Three subclones were initially selected for end sequencing at MWG Biotech. The first *Bam*HI fragment selected was the 7 kb band that showed strong hybridization to both NRPS and PKS probes in Southern analysis. The JC1-7kb subclone was sequenced off of the T3 and T7 promoters of the pBKS vector. The initial sequence showed strong sequence similarity to known PKSs on one side of the insert and strong sequence similarity to known NRPSs on the other side. The co-localization of both of these biosynthetic pathway elements to the same 7 kb fragment suggested their involvement in a common biosynthetic pathway and prompted the sequencing of the remainder of the insert.

The 9 kb *Bam*HI fragment was the second fragment to be sequenced. Even though MWG experienced some difficulty sequencing this clone, and the resulting sequence was not of high quality, two NRPS A-domains were found. In addition, one of the two motifs appeared to match the sequence of the JasF2 PCR-generated probe whose

gene product was predicted to activate a β-amino acid. The presence of a putative β-amino acid activating domain is consistent with the expected architecture of the jaspamide biosynthetic gene cluster and encouraged the continued sequencing of this insert.

The results of the initial sequencing of the 7 kb *Bam*HI fragment identified a juncture between the PKS and NRPS portions of the cloned pathway. Due to the mapped location of this fragment, it was expected that the upstream cosmid region would contain additional PKS modules. For this reason, the 5.7 kb *Bam*H1 fragment was submitted for complete sequencing. This fragment had been mapped to the region directly upstream of the 7 kb PKS containing fragment (Figure 2.3). The complete sequence of the 5.7 kb *Bam*H1 determined the size to be 5673 bp and confirmed the presence of additional PKS domains.

The terminal 11-13 kb *Bam*HI fragment was also sequenced. Initial sequencing was conducted with the 11 kb JC1 subclone, and the upstream, non-overlapping region of JC3-13, was sequenced by primer walking.

End sequencing of the terminal 4.5 kb *Bam*H1 fragment of JC7 revealed several conserved NRPS adenylation motifs and appears to be a continuation of the third adenylation domain found on the 9 kb fragment. Initial sequence of the other end of this fragment also showed high similarity to NRPS condensation domains, however no characteristic conserved sequence motifs could be found. The JC7-4.5 kb subclone proved to be one of the most difficult to sequence. Initially, transposon sequencing was attempted because it had been used with some success to sequence the larger subclones. Problems associated with secondary structure that were also encountered in the adjacent 9kb fragment seem to be present in this fragment and prevented efficient transposon introduction. Consequently, a combination of primer walking and shotgun cloning was necessary to sequence this fragment.

The presence of a terminal condensation domain in JC7-4.5, as well as the desire to identify other candidates for the jaspamide biosynthetic gene cluster prompted us to rescreen the second Fijian *Japsis* genomic library. Although this screening did not identify any new gene clusters, as discussed earlier, we were able to use the same library

to search for additional fosmids that overlapped the previously characterized JC1-7 cosmids. Screening of the library with a fragment of the JC7-9 kb subclone identified a single overlapping fosmid. Mapping of this new cosmid, JF2A, showed that it contained a larger fragment that overlapped with the JC7-4.5 kb *Bam*HI fragment (Figure 2.3). The new fosmid, JF2A, has seven *Bam*HI restriction fragments and includes the 9 and 7 kb bands found in the JC1-7 cosmids. JF2A had additional 5.7, 5.1,4.9, 4.1 and 1.4 kb *Bam*HI fragments. The 4.9 kb fragment aligned with the 11-12 kb BamHI fragments from JC1-7 and represented the upstream terminus of the fosmid insert. The 5.7, 7, and 9 kb *Bam*HI fragments were confirmed to be identical to the corresponding bands in JC1-7 through subcloning and end sequencing. The 5.1, 4.1 and 1.4 kb bands were all unique to this construct and represented regions downstream of the JC7-4.5kb fragment (Figure 2.3). This presented an opportunity to gain additional sequence information and determine if additional NRPS modules were located downstream of this region.

Sequencing of the three fragments proved to be extremely challenging. Transposon sequencing used for the bulk of the JC1-7 subclones proved to be fairly reliable and successful. When transposon sequencing was attempted with these fragments, however, the transposons were only able to integrate into a few locations. Consequently, a combination of shotgun sequencing and primer walking was required to sequence these subclones. The 5.1 kb subclone was found to overlap with the JC7-4.5 kb fragment. As this was the most downstream fragment in the JC cosmid clones, obtaining sequence for this region was a priority. Sequence on the upstream side of this fragment was found to overlap with JC7-4.5 kb band, while the downstream side did not. Additionally, it was not possible to read more than 350 bp into the subclone insert from the JC7 side due to the presence of secondary structure. More than 3000 bp of the downstream side of the same 5.1 kb fragment was sequenced, however the corresponding end of the JC7-4.5kb fragment was not found. This suggests that the terminal sequence of the Fijian and Pohnpei gene clusters were not identical or that a rearrangement had occurred on one of the cosmids. The other two downstream fragments were equally difficult to sequence. The 1.4 kb fragment was successfully sequenced, however it did not display any sequence homology to bacterial biosynthetic genes. The 4.1 kb fragment

was not successfully sequenced despite repeated attempts and further subcloning. The limited amount of sequence that was generated from the terminal ends of the fragment did not exhibit any similarity to biosynthetic gene clusters, suggesting that end of the gene cluster was contained on the JF2A 5.1 kb *Bam*HI fragment.

Sequence Analysis of the Mixed PKS/NRPS Gene Cluster

The approximately 39 kb of contiguous sequence obtained for the mixed *Jaspis* PKS/NRPS gene cluster contains both PKS and NRPS elements as well as a series of ORFs that appear to be involved in synthesis or modification of the polyketide starter unit. The overall GC content for the gene cluster was approximately 60% and approximately 70% in the third codon position. The stop codon for the PKS ORF and the start codon for the NRPS ORF overlap, indicating that they are co-transcribed off of a single operon. The organization of the *Jaspis* PKS/NRPS gene cluster is shown in Figure 2.4.

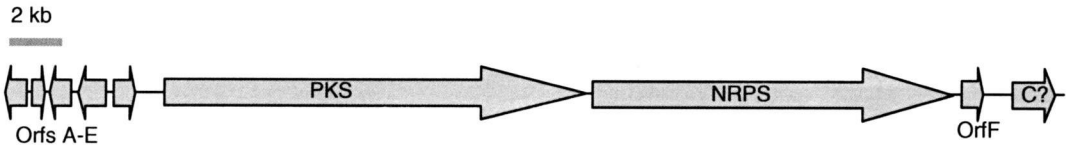


Figure 2.4. Organization of the *Jaspis* gene cluster.

PKS genes

The PKS portion of the pathway encodes three complete PKS extension modules (Figure 2.5A). Overall, the PKS gene region had high sequence similarity to PKS genes from *Nostoc punctiforme* (COG3321). Initial analysis of the domain arrangement for the PKS region of the gene cluster revealed several alterations from the predicted architecture of the jaspamide PKS gene cluster (Figure 2.5, B).

Typically, a type I PKS gene cluster will have a loading module containing an AT domain at the *N*-terminus of the protein.[19,20] Such a loading module was not found in the *Jaspis* cluster. Unlike other gene clusters, however, the *Jaspis* PKS has an additional ACP domain at the *N*-terminus of the first extension module. Similar PKS architecture is

was not successfully sequenced despite repeated attempts and further subcloning. The limited amount of sequence that was generated from the terminal ends of the fragment did not exhibit any similarity to biosynthetic gene clusters, suggesting that end of the gene cluster was contained on the JF2A 5.1 kb *Bam*HI fragment.

Sequence Analysis of the Mixed PKS/NRPS Gene Cluster

The approximately 39 kb of contiguous sequence obtained for the mixed *Jaspis* PKS/NRPS gene cluster contains both PKS and NRPS elements as well as a series of ORFs that appear to be involved in synthesis or modification of the polyketide starter unit. The overall GC content for the gene cluster was approximately 60% and approximately 70% in the third codon position. The stop codon for the PKS ORF and the start codon for the NRPS ORF overlap, indicating that they are co-transcribed off of a single operon. The organization of the *Jaspis* PKS/NRPS gene cluster is shown in Figure 2.4.

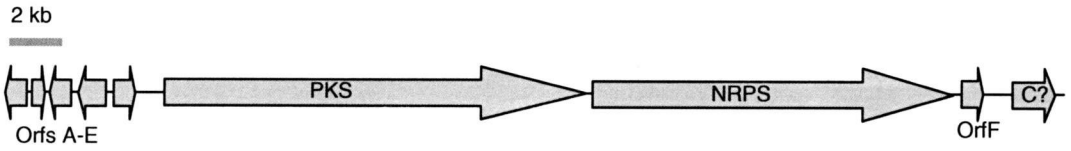


Figure 2.4. Organization of the *Jaspis* gene cluster.

PKS genes

The PKS portion of the pathway encodes three complete PKS extension modules (Figure 2.5A). Overall, the PKS gene region had high sequence similarity to PKS genes from *Nostoc punctiforme* (COG3321). Initial analysis of the domain arrangement for the PKS region of the gene cluster revealed several alterations from the predicted architecture of the jaspamide PKS gene cluster (Figure 2.5, B).

Typically, a type I PKS gene cluster will have a loading module containing an AT domain at the *N*-terminus of the protein.[19,20] Such a loading module was not found in the *Jaspis* cluster. Unlike other gene clusters, however, the *Jaspis* PKS has an additional ACP domain at the *N*-terminus of the first extension module. Similar PKS architecture is

seen in the myxothiazol biosynthetic gene cluster, however the initial modules in that system also contain two tandem AT domains following the initial ACP domain.[14] The terminal ACP in this cluster would presumably be loaded with a starter unit from an unidentified AT.

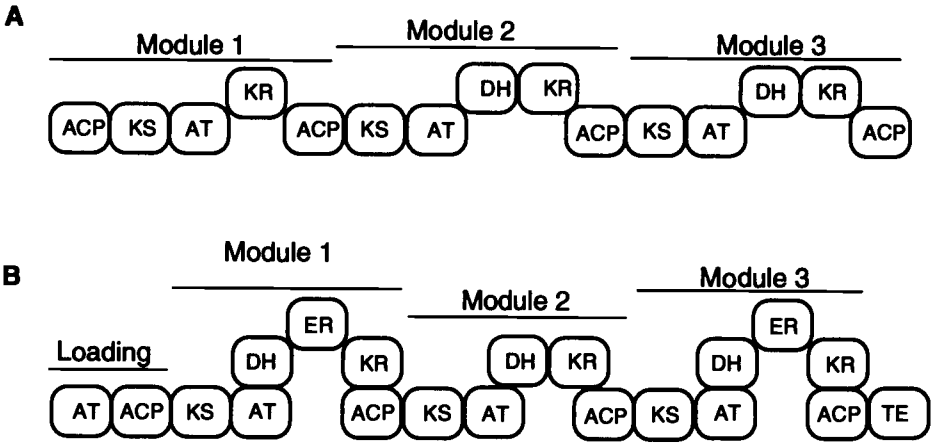


Figure 2.5. A. Domain organization of the PKS region of the cloned *Jaspis* gene cluster. B. Predicted organization of the jaspamide biosynthetic gene cluster.

The first PKS module is typical of PKS modules and contains a KS, AT, KR, and ACP domains. Each of the domains within the first module contains the appropriate conserved motifs that are in good agreement with the consensus motifs (Table 2.2).

Table 2.2. Comparison of sequence motifs for *Jaspis* PKS modules with consensus sequence.

	KS	AT	KR	DH	ACP
Consensus	TVDTGCSSxL	GHSxG	GxGxxGxxxA	LxxHxxxGxxxxP	LGxDS
Jaspis PKS Mod1	ALDTACSSSL	GHSVG	GLGVLGLRVA	N/A	MGMDS
Jaspis PKS Mod2	IADTACSSSL	GHSVG	GLGALGLQVA	LADYRVCDAVVVP	LGMDS
Jaspis PKS Mod3	YVDTACSSSL	GHSVG	GLGALGLQVA	LKHHRVFERVILP	MGFDS

In addition to the KS, AT, KR and ACP domains, the second PKS module contains a region that has high sequence similarity to DH domains, including MxaC and MtaD.[14,21] It is not known if this is a functional DH-domain however, because the conserved LxxHxxxGxxxxP sequence motif is altered to LxxYxxxDxxxxP (Table 2.3). Despite this difference, the putative DH-domain has 43% identity, 58% similarity with MxaC (*Stigmatella aurantiaca* Sga15) and 41% identity, 57% similarity with MtaD (*Stigmatella aurantiaca* DW4/3-1).

Table 2.3. Comparison of DH domain sequence motifs. Alignment of DH-domain conserved sequence motifs from *Jaspis* PKS modules 2 and 3 with DH motifs from myxalamid (MxaC) and myxathiazol (MtaD) domains.

Consensus	**LxxHxxxGxxxxP**
MxaC DH1	LaeHrvqGvpalP
MxaC DH2	LddHrlyGtvvvP
MtaD DH	IddHrvnGtvvlP
***Jaspis* M2 DH**	LadYrvcDavvvP
Jaspis M3 DH	LkhHrvfErvilP

The third PKS module contains the core KS, AT, and ACP domains as well as a DH and a KR domain. Like the second module, the DH domain found within the third PKS module has a substitution of an acidic residue for the conserved glycine in the sequence motif, however the conserved histidine residue is present (Table 2.3).

In addition to domain architecture, substrate specificity predictions can be used to gain information about the biosynthetic product of a PKS gene cluster. The type of extender unit (malonyl or methylmalonyl-CoA) incorporated by PKS modules is determined by the AT-domain. Through sequence comparisons, Leadlay *et al.* were able to identify sequence motifs that would predict the substrate specificity of PKS AT domains.[22] Using this method the substrate specificity for each of the three *Jaspis* PKS modules was predicted. The AT domains in the *Jaspis* cluster do not perfectly match either malonyl-CoA (acetate incorporating) or methylmalonyl-CoA (propionate incorporating) motifs (Table 2.4). The first substrate specificity motifs of the two *Jaspis* AT domains most closely match acetate incorporating domains. The third AT motif shares equal similarity to both types of domains. More recently, groups lead by Leadlay and Müller identified an additional sequence motif in AT domains that appears to correlate with the specificity for either malonyl-CoA or methylmalonyl-CoA.[21,22] All three of the AT domains identified in the *Jaspis* cluster contain the HAFH motif which corresponds to acetate incorporation. Taken together, it is predicted that all three AT modules incorporate acetate into the polyketide chain.

Table 2.4: Comparison of acyltransferase consensus motifs with corresponding regions from *Jaspis* PKS

Mal-CoA consensus	ETGYA	QVALFXLL	GHSVG	xxAFH
MeMal-CoA consensus	RVDVV	MXSXAAhW	GHSQG	DYASH
Jas AT1	QTAYT	EYALAALW	GHSVG	HAFH
Jas AT2	QTAYT	EYALSALW	GHSVG	HAFH
Jas AT3	QTAYT	EYALSALW	GHSVG	HAFH

NRPS genes

The *Jaspis* mixed gene cluster was also predicted to encode a three module NRPS (Figure 2.6, A). The domain organization of this NRPS contains a condensation domain upstream of the first A-domain. This domain is located between the end of the PKS portion of the pathway and beginning of the NRPS portion and likely serves to transfer the completed polyketide chain to the first NRPS module. Condensation domains have been located at the NRPS/PKS juncture of other mixed NRPS/PKS pathways, such as in microcystin synthetase,[23] and is believed to be involved in the transfer of the product between the two enzymes. A PKS-type phophopantetheine binding motif, LGxDS, is also located between the C-domain and the A-domain of the first module, however it is not known if this post-translational modification of this motif occurs or not.

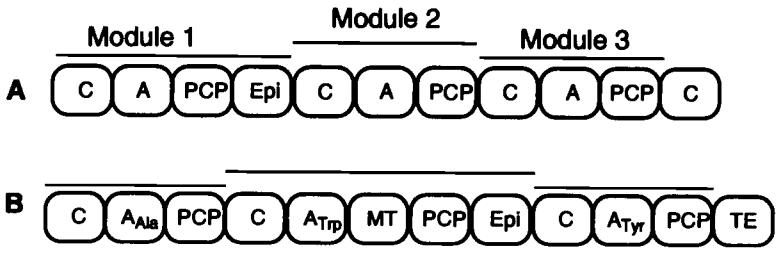


Figure 2.6. A. Domain organization of the NRPS region of the cloned *Jaspis* cluster. B. Predicted organization of the jaspamide biosynthetic

The second and third NRPS modules contain the standard C-A-PCP domain organization and do not contain any auxiliary domains. The A-domain associated with the first NRPS module, however, has several interesting features. The module 1 A-domain corresponds to the JasF2 PCR product. This domain has the altered A4 motif of FxDS. This altered aspartate position has been seen in adenylation domains that activate

β-amino acids. An example of this is the FxbB protein of exochelin synthetase, which activates a β-alanine.[24] Also associated with the first complete module is an epimerization domain, which suggests the introduction of a D-amino acid at this position. Mutational analysis of a PCP-domain by Linne *et al.* determined that an alteration in the conserved phosphopantetheine binding motif (LGGHS) to LGGDS is required for epimerization of the bound NRPS intermediate.[25] The PCP-domain associated with the first Jaspis *NRPS* module contains this alteration and suggests that the epimerization domain is functional.

Although a thioesterase domain has not been located within the *Jaspis* gene cluster, an additional condensation domain is located at the *C*-terminal end of module 3 where a thioesterase domain would typically be located if this is the actual end of the structural genes. Only the first five of the seven conserved condensation domain motifs could be located for the *C*-terminal condensation domain motif and its involvement in the biosynthetic pathway is not known.

The amino acid selectivity for each of the three adenylation domains was predicted using the method developed by Challis *et al.* and Stachelhaus *et al.*[15,16] This method was developed based on the crystal structure of a phenylalanine activating A-domain from gramicidin synthetase, GrsA.[26] The crystal structure allowed for the identification of the substrate binding pocket as well as several active site amino acids involved in ligand binding.[26] Key interactions were found to include ATP binding by Asp413 and Lys517. Lys517 and Asp235 were found to be important in binding the substrate amino acid carboxylate and amino groups, respectively. The amino acid sequences of other NRPS A-domains of known substrate specificity were then aligned with GrsA and resulted in the identification of active site amino acid residues that are involved in substrate selectivity.[15,16] Identification of these residues through sequence alignments and subsequent comparison to A-domains of known function can be used to predict substrate specificity. The A-domains for each of the three *Jaspis* NRPS modules were aligned with GrsA and the binding pocket amino acids were identified and are shown in Table 2.5.

Table 2.5. Substrate specificity 'codons' for the *Jaspis* A-domains. Numbering is according to GrsA.

Position	235	236	239	278	299	301	322	330	517	Predicted substrate
Jas ModA1	V	D	W	V	I	S	L	A	K	β-amino acid/Ser
Jas ModA2	D	V	W	F	I	S	L	I	K	Ser/hydrophobic
Jas ModA3	D	I	L	Q	L	G	L	I	K	Gly

Based on alignment alone, the first and second *Jaspis* cluster A-domains most closely match the binding pockets of serine activating domains including the nostopeptolide NosA serine activating module.[27] While the alignment of these two domains is quite similar to those activating serine, one key amino acid is not in agreement. Comparison of serine activating domains reveals a nearly invariant histidine residue at position 278 (Figure 2.7, B). This histidine is believed to hydrogen bond with the serine hydroxyl.[16] *Jaspis* domains A1 and A2 contain hydrophobic residues in positions 278 as well as at the highly selective position 239. This suggests hydrophobic amino acid binding, rather than a serine, may be facilitated in these two domains. Consequently sequence analysis would predict that JasA1 activates a hydrophobic β-amino acid.

The selectivity-conferring residues of the third *Jaspis* A-domain predict the activation of a glycine residue. While the binding pocket residues are identical with those of glycine activating domains, it is interesting to note that the substitution of residue Ile330 for valine would alter the specificity to 3-hydroxy-4-methyl phenylalanine, as is seen in saframycin Mx1 synthetase.[28]

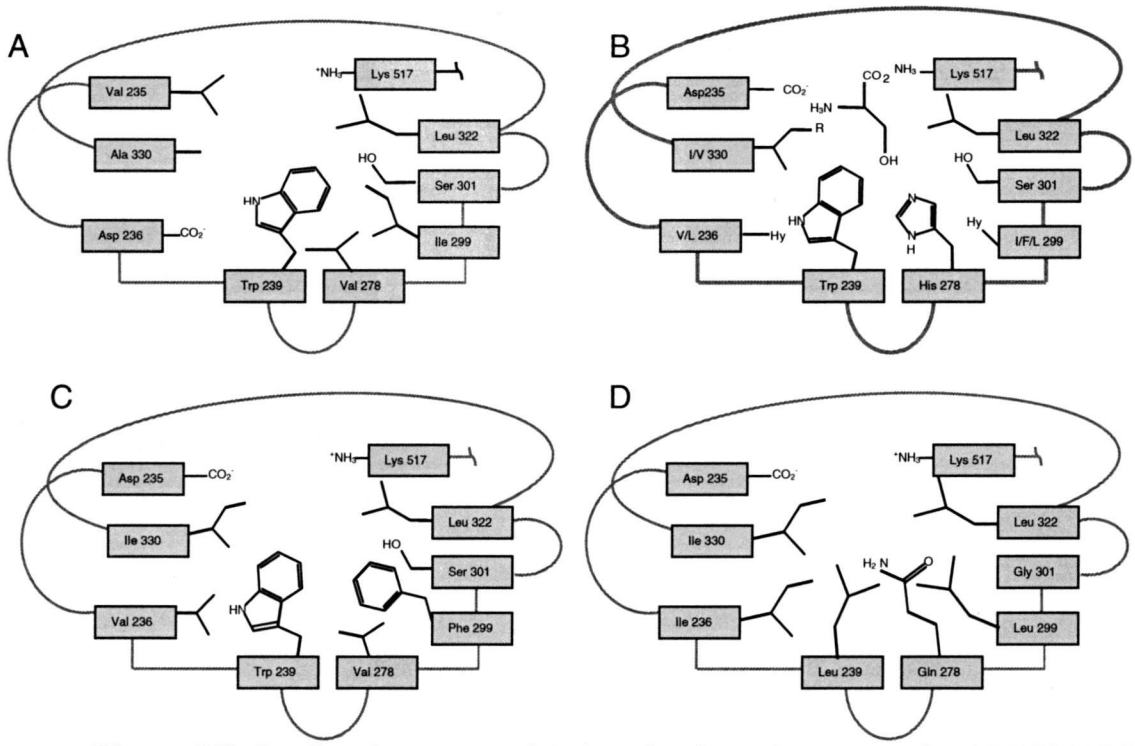


Figure 2.7. Predicted structure of A-domains from the *Jaspis* mixed PKS/NRPS gene cluster. A. Module 1 A-domain. B. Representative serine activating A-domain with bound substrate. C. Module 2 A-Domain D. Module 3 A-Domain

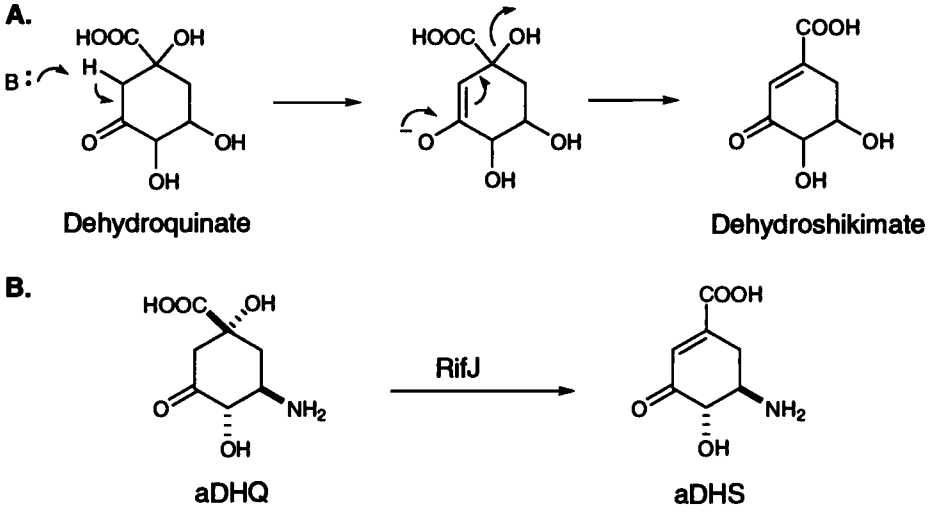
Amino acid specificity for the first and second *Jaspis* A-domains can only be speculated, however, as there are not A-domains with identical "codons" that have been biochemically characterized. This method of substrate prediction is not without limitations and each new domain must be evaluated on an individual basis. Marahiel's work identified a variation in the importance of residues in different regions of the binding pocket.[16] In addition, the significance of a given residue appears to vary according to the type of amino acid being activated. For example, the specificity for small amino acids is most greatly affected by residue 322 near the top of the binding pocket while residues 239 and 278 near the bottom of the pocket tend to affect the selectivity for large amino acids.

Additional ORFs Associated with the Mixed Gene Cluster

Located upstream of the PKS portion of the *Jaspis* gene cluster are five ORFs that may be involved in the synthesis or modification of a polyketide starter unit. The first ORF, *orf*A, is located at the upstream terminus of the cosmid inserts. Although *orf*A is

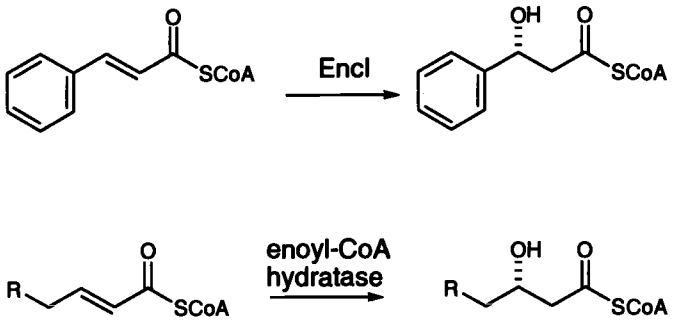
not completely contained on the cosmid inserts, sequence analysis of the partial gene product found it belongs to the conserved amidohydrolase family, pfam04909.5.[29] The closest matches in the public databases to OrfA are a series of hypothetical proteins from *Streptomyces avermitilis* and *Mycobacterium avium*. The closest protein of known function is 2-amino-3-carboxylmuconate semialdehyde reductase. OrfA also shares 27% identity and 46% similarity with BarH from the barbamide biosynthetic gene cluster. BarH is predicted to be involved in decarboxylative formation of the thiazole moiety of barbamide.[30]

A conserved domain search with OrfB revealed that it is a member of the dehydroquinase class II protein family, pfam01220.11.[29] OrfB is most similar to dehydroquinate dehydratase from *Thermoanaerobacter tengcongensis* and shares 43% identity, 65% similarity. Dehydroquinate dehydratases are involved in the shikimate biosynthetic pathway and catalyze the formation of dehydroshikimate from dehydroquinate through an enolate intermediate (Scheme 2.1).[31] The rifamycin B biosynthetic pathway contains a protein, RifJ, that is also analogous to dehydroquinate dehydratase. RifJ is involved in the synthesis of the rifamycin B starter unit, 3-amino-5-hydroxybenzoate (3,5-AHBA).[32] Specifically, RifJ catalyzes the formation of 5-deoxy-5-amino-3-dehydroshikimate (aDHS) from 5-deoxy-5-amino-dehyroquinate (aDHS) (Scheme 2.1).[33,34]



Scheme 2.1. A. Mechanism of a type II dehydroquinate dehydratase. B. Reaction catalyzed by RifJ.

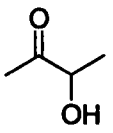
The third ORF, *orfC*, is predicted to encode a member of the enoyl-CoA hydratase/isomerase family of proteins. This is a diverse protein family that contains hydratases, isomerases and racemases. The closest match in the databases to OrfC is an unknown protein from *Mesorhizobium loti* MAFF303099 with 41% identity and 56% similarity. Enoyl-CoA hydratase from *Archaeoglobus fulgidus* is the closest OrfC homolog of known function and shares 38% identity and 57% similarity. Enoyl-CoA hydratases are involved in lipid metabolism. An enoyl-CoA hydratase type enzyme, EncI, has been identified in the biosynthetic pathway of the type II aromatic polyketide enterocin and cloned from the producing marine bacterium *Streptomyces maritimus*.[32,35] EncI is believed to be a cinnamoyl-CoA hydratase involved in the biosynthesis of the benzoyl-CoA enterocin starter unit (Scheme 2.2).[36]



Scheme 2.2. Reactions catalyzed by enoyl-CoA hydratase and enterocin biosynthetic enzyme EncI

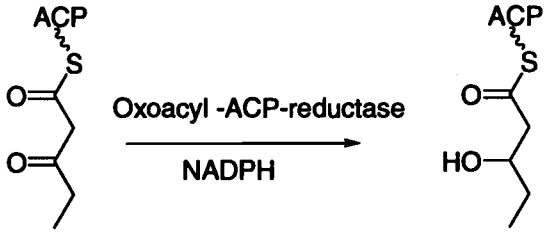
The product of the fourth ORF, OrfD is related to a number of proteins involved in acetoin catabolism. The closest match in the public databases for OrfD is a hypothetical protein from *Chloroflexus aurantiacus*. OrfD shares sequence homology with AcoX, a protein involved in acetoin catabolism in several bacteria including *Alcaligenes eutrophus* (40% identity, 59% similarity) and *Chloroflexus aurantiacus* (46% identity, 61% similarity).[37,38] Mutational studies have demonstrated that AcoX is essential for acetoin catabolism, however the function of this protein is unknown.[37,38] Additionally, a search in the conserved domain database revealed that OrfD belongs to the ATP-NAD^+ kinase protein family that catalyze the phosphorylation of NAD^+ to yield $NADP^+$ (pfam01513.11) however OrfD contains imperfect NAD kinase conserved sequence motifs XXX-XGGDG-XL (LVTLGGDGTNL) and TPTGSTAY

(TGAGGTTV) where X is a hydrophobic residue.[29,39] Although the function of OrfD cannot be predicted it is of interest to note that a kinase is involved in the biosynthesis of the rifamycin precursor AHBA. Kanosamine 6-kinase has been identified in the rifamycin gene cluster and is involved in the biosynthesis of AHBA.[40] Orfs B and D may perform similar functions in the *Jaspis* PKS/NRPS pathway.



Acetoin

Like Orfs C and D, OrfE is similar to enzymes involved in fatty acid metabolism. *orf*E is located just upstream of the PKS gene and has 38% identity and 57% similarity to 3-oxoacyl-(ACP) reductase, FabG, from *Bacillus halodurans*.[41] Oxoacyl-ACP-reductase catalyzes the formation of 3-hydroxyacyl-ACP in the sequential reduction of oxoacyl-ACP to acyl-ACP (Scheme 2.3).[42]



Scheme 2.3. Reaction catalyzed by oxoacyl-ACP-reductase

In addition to the genes for biosynthesis enzymes, a gene for a putative regulatory protein has also been identified within the *Jaspis* PKS/NRPS gene cluster (Figure 2.4). The *orf*F gene product, OrfF, is similar to the SyrP family of regulatory proteins. BLAST search analysis of this OrfF showed that it has 46% identity and 61% similarity to a SyrP-like protein in the bleomycin biosynthetic cluster of *Streptomyces verticillus*.[43] OrfF also has 37% identity and 56% similarity with SyrP. The SyrP protein is believed to be involved in the regulatory phosphorylation cascade that is associated with the control of syringomycin biosynthesis in *Pseudomonas syringae*.[44]

On the terminus of the JC7 cosmid insert, just downstream of *orf*F, is a region that has significant homology to NRPS condensation domains through BLAST analysis. Despite the similarity, however, this region was not found to encode any of the conserved

C-domain sequence motifs. The inability to obtain additional sequence for this downstream region prevented further analysis of this putative condensation domain. An interesting feature of this putative C-domain is the separation from the other biosynthetic genes by a regulatory gene, *orf*F. The insertion of ORF into the end of an NRPS gene cluster is unusual; however the tyrocidine biosynthetic operon provides a precedent for this type of organization. The tyrocidine operon contains two *orf*Fs (*tyc*D and *tyc*E) that code for ABC transporters.[45] These two ORFs separate the thioesterase domain from the rest of the biosynthetic genes located on *tyc*A, B, and C.

Discussion

The first step in gaining access to the incredible biosynthetic potential of marine organisms is the development of applicable molecular genetic techniques. In this regard, the initial aim of this project was to develop methods for isolation of metagenomic DNA from a sponge assemblage, construct a genomic library and then used the library to try and identify and clone the jaspamide biosynthetic gene cluster. Molecular techniques for the identification of biosynthetic gene clusters from cultivable organisms have been well established.[5,7,46] Adaptation of these techniques for use with complex assemblages of uncultivable organisms, however, requires the consideration of several factors.

One factor that must be considered is the often limited supply of marine tissue sample that is available for DNA extraction. The scarcity of some organisms and the difficulties associated with the collection of many marine samples often dictates that only a limited supply of material will be available for genetic analysis. Another factor that may contribute to this need to work from small tissue samples is the heterogeneous nature of the invertebrate-microbial assemblages. The presence of an interesting bioactive metabolite in one collection does not guarantee that the same compound will be present in all similar specimens.[47] This will be particularly important in situations where an associated microbe is involved in the production of the metabolite. In these situations it is necessary to confirm the presence of the target compound prior to molecular investigation. These factors demonstrate the need to develop techniques that allow for the maximum recovery of high molecular weight genomic DNA from these limited tissue samples. The techniques developed for the isolation of metagenomic DNA from *Jaspis*

splendens in this work has also been applied to the isolation of gDNA from other marine sources including other sponge families, tunicates, and cyanobacteria. Once high quality gDNA has been isolated, amplification of gene probes and generation of genomic libraries can be conducted according to standard molecular biology techniques.

Several features need to be designed into the gene probes for their optimal effectiveness in identifying the desired gene cluster. The probes need to be specific enough to identify the desired gene cluster while minimizing the number of false positives resulting from hybridization to other similar genes. This issue becomes important when the possible number of related gene clusters within an invertebrate-microbe assemblage is considered. The extreme potential of microbes to produce secondary metabolites was demonstrated by the genomic sequencing of several antibiotic producing bacteria.[48,49] An example is the avermectin producer, *Streptomyces avermitilis*, which contains 24 PKS and NRPS gene clusters.[48] This demonstrates the number of pathways that must be sorted through in just one organism. When looking at symbiotic marine organisms, this metabolic potential must then be multiplied by the number of microbes found within the invertebrate tissue. Thus, the designed gene probes must be specific enough to identify the correct gene cluster amid potentially hundreds of similar gene clusters. Another feature of the probes is that they need to be able to hybridize to the gene of interest, although the sequence of the target gene is not known. For this reason, it was necessary to select target genes that are not only likely to be found within the desired gene cluster, but also belong to a gene family with regions of high sequence similarity. An obvious feature of the two gene probe requirements is the paradox created by the need to be specific enough to identify only the desired gene cluster but also be general enough to hybridize with a gene of unknown sequence is not known. The opposing features of these two requirements must be balanced in order to optimize the likelihood of identifying the target gene cluster while minimizing the number of clones that will need to be sequenced.

A method to improve target gene specificity in the degenerate PCR primers is to consider the codon bias of the host organism. This allows for the design of primers with minimal degeneracy and increases the hybridization specificity. This presents a unique

problem with marine organisms where it is not known which organism, among potentially hundreds of organisms, harbors the target gene. Often this will require the sequential screening with degenerate primers pairs designed with alternate codon biases. This same bias, however, can be used to the advantage of the researcher if the producing organism is known or can be predicted. In this work, an early hypothesis was that a marine bacterium, possibly a myxobacteria or related bacteria, was likely to be the true producer of jaspamide. For this reason, the degeneracy of some primers used for gene probe amplification were biased to myxobacteria-like high GC codons. Although the gene cluster that was identified contained only 60% GC overall, the percentage of GC in the third codon position was higher (~70%) and is consistent with the codon bias of the primers.

Using these techniques, metagenomic libraries were constructed from two jaspamide producing *Jaspis splendens* specimens. One concern in regard to the cloning of antibiotic gene clusters is that heterologous expression, without the necessary resistance genes, may kill the host cell. In an attempt to reduce possible gene toxicity, an inducible fosmid vector system was used to generate the second Fijian library. The advantage of using fosmids is that they are maintained as only a single copy in the host cells and are often used when the product of a cloned gene may be toxic to the heterologous host. By maintaining only a single copy, toxicity of a cloned gene product is less likely to kill the host cell. The use of fosmid systems increases the chances for the successful creation of genomic libraries that do not have a bias against potentially toxic genes and reduces the likelihood of cosmid rearrangement. A disadvantage of typical fosmid systems, however, is that in order to purify large amounts of fosmid for sequencing and subcloning it is necessary to grow large volumes of cells. This typically makes fosmid systems more difficult to work with as compared to cosmid systems. The fosmid vector system, pCC1Fos™, used in this work could be maintained as a single copy vector under normal growth conditions, such as would be used during the screening of the library, but that could be induced to a high copy number when it was time to purify the fosmid from the host.

Screening of these libraries with PKS and NRPS gene probes resulted in the identification of a single mixed PKS/NRPS gene cluster. Given the microbial diversity within the sponge sample, this was not expected. Sequencing of this gene cluster proved to be a difficult task, as frequent regions of high secondary structure would result in aborted sequencing runs. For this reason, a variety of techniques needed to be applied in order to accomplish the complete sequencing of the gene cluster. These sequencing difficulties also resulted in only single strand coverage of some regions of the *Jaspis* gene cluster.

Sequence analysis alone cannot confirm or reject this pathway as being the jaspamide biosynthetic cluster, however the majority of the evidence suggests that it is not. One factor that contributes to this conclusion is the architecture of the NRPS gene. The amino acid sequence of nonribosomal peptides is nearly always the same as the order on the chromosome of the modules encoding their incorporation and processing; the so-called colinearity principle.[50] The predicted order of amino acid incorporation in the identified pathway is β-(hydrophobic) amino acid, hydrophobic amino acid (or serine), glycine. The NRPS encoding the jaspamide tripeptide, L-alanyl-*N*-methyl-2-bromo-D-tryptophanyl-D-β-tyrosine, should have the β-amino acid incorporating module furthest downstream (Figure 2.8, A). The NRPS in the cluster described here is predicted to form a peptide having a D-β-amino acid at the *N*-terminus (Figure 2.8, B). As alanine and glycine only differ in the presence of a methyl group, it is conceivable that the third adenylation domain could activate an alanine residue. The binding pocket of the second module, however, is quite different than would be predicted for tryptophan in terms of both polarity and steric requirements. Consequently, it is less likely that this domain would activate the tryptophan required for jaspamide biosynthesis. The prediction that the PKS AT domains incorporate malonyl-CoA rather than methylmalonyl-CoA is also inconsistent with the predictions for the jaspamide gene cluster.

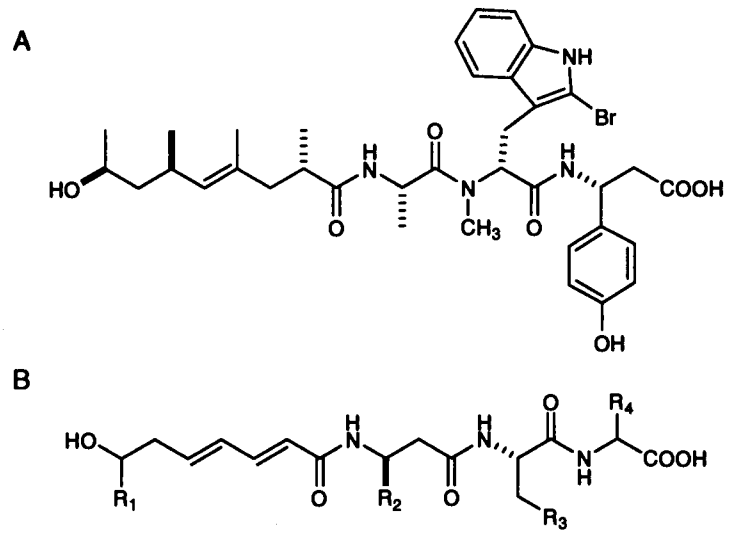


Figure 2.8. Comparison of the expected linear precursor of jaspamide (A) with the expected linear precursor of the cloned *Jaspis* PKS/NRPS (B).

Identification of auxiliary domains in the cloned NRPS pathway has provided little support for the production of jaspamide by this cluster. The proposed jaspamide pathway would require the epimerization domain, as well as an *N*-methylation domain, to be located in the second, tryptophan activating, module (Figure 2.4). The pathway that has been cloned has the epimerization domain in the first module. An *N*-methylation domain is also not found in the cloned gene cluster and supports the belief that this may not be the jaspamide biosynthetic cluster.

Another absent feature of the *Jaspis* PKS/NRPS cluster is a thioesterase domain. Typically these domains are located at the end of the last module. It is possible that either of the two final condensation domains is performing an equivalent function and catalyzing the final release (and possibly cyclization) of the mixed NRPS/PKS product. A third set of elements that would be expected to flank the biosynthetic gene cluster and have not been located are genes for accessory proteins such as a phosphopantetheinyl transferase and tailoring enzymes such as a halogenase.

Even though the features of the cloned pathway do not match those predicted for the jaspamide biosynthesis, sequence analysis alone can not rule out this is the correct cluster. The cloned pathway is clearly a mixed NRPS/PKS system. The differing rearrangement may simply indicate the presence of an unprecedented biosynthetic mechanism. Sequence analysis is able to provide only a limited amount of information

and deciphering the role of this PKS/NRPS gene cluster in the secondary metabolism of the *Jaspis* assemblage requires additional biochemical characterization.

Materials and Methods

General

Routine molecular biology procedures including DNA manipulation, growth and maintenance of *E. coli* were conducted according to standard techniques.[51] Restriction enzymes, T4 DNA ligase and *taq* polymerase were purchased from various suppliers and used according to manufacturer's protocols. QIAprep® spin miniprep and QIAquick® gel extraction kits (Qiagen) were used for DNA purification. Sequencing of vector inserts was conducted by MWG Biotech and the Center for Gene Research and Biotechnology (CGRB), Oregon State University.

Sequence analysis was conducted with Vector NTI (Informax) software. BLAST search analysis was conducted in National Center for Biotechnology Information (NCBI) public databases.[29,52]

Samples of *Jaspis splendens* and other marine invertebrates used in this work were kindly provided by Professor Tadeusz F. Molinski, University of California, Davis; Mary Kay Harper, University of Utah; and Dr. Valerie Bernan, Wyeth Research (Pearl River, NY).

Isolation of Total Genomic DNA from Sponge Tissue: Method 1.

Fifty to eighty mgs of minced sponge tissue was weighed into a 1.5 mL microcentrifuge tube and tissue washed once with 10.3% sucrose and twice with TE buffer (10 mM Tris, 1 mM ethylenedinitrilo-tetraacetic acid (EDTA) pH 8.0). Cells were lysed by incubation in 600 µL lysis solution (17% sucrose, 33 mM Tris, 0.1 mM EDTA, 8 mg/mL lysozyme, pH 8.0) at 37 °C for 1-2 h, until the sample became viscous. Proteinase K (0.2 mg) and 360 µL 10% SDS were added and sample was incubated at 37 °C for 2 h. Proteins were precipitated by adding 160 µL 10% hexadecyl trimethyl ammonium bromide (CTAB) and incubation at 65 °C for 10 min. Following extraction of the aqueous layer with 1.2 mL phenol/chloroform/isoamyl alcohol (25:24:1), genomic

DNA was precipitated with 300 µL isopropyl alcohol at -20 °C. Genomic DNA was harvested by centrifugation and washed with 70% ethanol prior to being dissolved in sterile water.

Isolation of Genomic DNA from Sponge Tissue: Method 2

Genomic DNA was prepared using a PureGene Genomic DNA purification kit for yeast and Gram-positive bacteria (Gentra Systems) with a slightly modified protocol. All reagents were provided in the kit. 30-50 mg of sponge tissue was weighed into a 1.5 mL microcentrifuge tube and homogenized with microscissors and a microfuge tube homogenizer. Tissue was resuspended in 300 µL cell suspension solution with 1.5 µL lytic enzyme solution and incubated at 37 °C for 30 min. Tissue sample was pelleted by centrifugation and resuspended in 300 µL cell lysis solution then incubated at 80 °C for 5 min. After cooling to room temperature, 1.5 µL RNAseA solution was added and the sample was incubated for 1 h at 37 °C. The sample was cooled to room temperature and 100 µL protein precipitation solution was added with brief (20 second) vortexing, followed by incubation at 37 °C for 1 h. Following centrifugation, the supernatant was transferred to a new 1.5 mL microcentrifuge tube and 1.5 µL glycogen solution was added. Genomic DNA was precipitated by adding 300 µL isopropanol, harvested by centrifuging, washed with 70% ethanol, and dissolved in 30 µL DNA hydration solution.

Amplification of PKS Gene Probes

Isolated gDNA was either partially digested with *Sau*3A or mechanically sheared by passing through a 20 gauge hypodermic needle and purified by ethanol precipitation. PCR was carried out in a total volume of 50 µL containing 60 ng gDNA template, 1X Promega Thermophilic Buffer, 2.5 mM $MgCl_2$, 0.4 mM dNTP mix, 5% DMSO, 50 pmol forward primer 5' MIGIGARGCIHWISMIAGTGAYCCICARCIA -3', 50 pmol reverse primer 5' GGRTCICCIARISWIGTICCIGTICCRTG -3' (where M is A or C; I is inosine; R is G or A; H is A,C, or T; W is A or T; and S is G or C) , and 5 units Taq polymerase (Roche). Gel purified PCR products were ligated into pGEMT*easy* and used to transform *E. coli* DH5α or JM109 competent cells.

Amplification of NRPS Gene Probes

Isolated gDNA was either partially digested with *Sau*3A or mechanically sheared by passing through a 20 gauge hypodermic needle and purified by ethanol precipitation. PCR was carried out in a total volume of 50 μL containing 60 ng gDNA template, 1X Promega Thermophilic Buffer, 2.5 mM $MgCl_2$, 0.4 mM dNTP mix, 5% DMSO, 50 pmol forward primer LGA2F1, 5'GCNGGYGGYGC NTAYGTNCC3', 50 pmol reverse primer LGA8R1, CCNCGDATYTTNACYTG 3', and 5 units Taq polymerase (Roche). Gel purified PCR products were ligated into pGEMT*easy* and used to transform *E. coli* DH5α or JM109 competent cells.

Preparation of a Jaspis Cosmid Library

A cosmid library of *Jaspis* genomic DNA was constructed using the pWEB™ Cosmid Cloning Kit (Epicentre). Purified gDNA was end repaired with no additional shearing and fractionated on a 1% low melting point agarose gel. The approximately 40 kb band was excised. gDNA was purified from the gel by gelase (Roche) digestion followed by ammonium acetate precipitation and ethanol precipitation of the gDNA. Prepared gDNA was ligated to the linearized, dephosphorylated pWEB vector at room temperature overnight. The ligation mixture was then packaged using MaxPlax™ Lambda Packaging Extract. E. *coli* EPI305 cells were infected with the packaged phages and plated on LB-ampicillin plates. The titer (15000 cfu/mL) was calculated after overnight incubation. The remaining packaged phages were then used to infect EPI305 cells and plated to obtain 300-400 colonies per 150 mm plate.

Preparation of a Second Jaspis Fosmid Library

A second genomic library was constructed using the Copy Control™ Fosmid Library Kit (Epicentre). The protocol used was the same as for production of the cosmid library with the following modifications: gDNA was purified from the sizing gel with the QIAquick® gel extraction kit (Qiagen) and packaged fosmids were transformed into *E. coli* ECI 300 competent cells.

Screening the Cosmid Library

Duplicates of the cosmid library were transferred to nylon membranes through standard colony lift procedures.[17] Lifted colonies were allowed to regenerate by incubating the membranes on a fresh LB-ampicillin plate for 3-5 h. (regeneration of the fosmid library was conducted on LB-ampicillin plates containing 0.01% arabinose to induce high copy number.

The colony lifts were processed for hybridization by consecutively placing them on top of a series of 3M Whatman papers saturated with each of the following solutions for the time indicated: (1) 10% sodium dodecyl sulfate (SDS), 5 min; (2) 0.5 M NaOH, 1.5 M NaCl, 15 min; (3) 0.5 M Tris, 1.5 M NaCl, 15 min; (4) 2X SSC (300 mM NaCl, 30 mM sodium citrate), 10 min. The DNA was fixed to the membrane by placing the membranes into an 80 °C vacuum oven for one hour. Prior to prehybridization the membranes were placed into hybridization tubes and washed in a solution containing 50 mM Tris, pH 8.0, 1 M NaCl, 1 mM EDTA, and 0.1% SDS in a minihybridization oven at 42°C.

Preparation of Digoxygenin Labeled Probes

NRPS and PKS probes were amplified by PCR and purified as described above. A total of 600 ng of each purified PCR product was labeled with digoxygenin using the Dig Labeling Kit (Roche) according to manufacturer's instructions.

Colony Hybridization

The colony lifts were prehybridized with 50 mL of prehybridization solution per tube at 60 °C for at least two h in a minihybridization oven. The prehybridization solution contained 5X SSC (750 mM NaCl, 75 mM sodium citrate, pH 7.0), 1X Roche Blocking Agent, 0.1% lauroyl sulfate, and 0.02% SDS. The hybridization solution contained prehybridization solution with 10 ng/mL Digoxygenin labeled probe. Membranes were hybridized for 36 h. in a minihybridization oven at 65 °C.

Washing and Detection of Hybridized Membranes.

Following hybridization, the membranes were washed successively with each of the following solutions for the time indicated: (1) twice in 2x SSC, 0.1% SDS at RT for 10 min; (2) twice in 0.1% SDS, 1x SSC (150 mM NaCl, 15 mM sodium citrate, pH 7.0) at 68 °C, 15 min; (3) once in 0.1% SDS, 0.5 x SSC (75 mM NaCl, 7.5 mM sodium citrate, pH 7.0) at room temperature (RT) for 10 min. Detection of hybridized probes was carried out using a digoxygenin wash and detection buffer kit (Roche) according to manufacturer's instructions. Following reaction of antidigoxygenin-AP (alkaline phosphatase) with the hybridized Digoxygenin labeled probes, the membranes were exposed to disodium 3-[4-methoxyspiro {1,2-dioxetane-3,2'-(5'-chloro) tricyclo [3.3.1.1]decan}-4-yl] phenyl phosphate (CSPD). Chemiluminescence detection of hybridized probes, resulting from the AP-CSPD reaction, was detected by exposing the membranes to film.

Southern Blots

Purified cosmid DNA was digested with either *Not*1 or *Bam*H I. The digests were elecrophoresed on a 1% agarose gel overnight at 35-40 V. Gel was washed successively with 0.25 N HCl for 15 min, then in denaturation solution (1.5 M NaCl, 0.5 M NaOH) for 35 min. Finally, the gel was washed with neutralization solution (1.5 M NaCl, 1 M Tris, pH 8.0) for 35 min. The DNA was then transferred to a Hybond N+ nylon membrane (Pharmacia) with a BioRad vacuum blotter using 10X SSC (1.5 M NaCl, 150 mM sodium citrate, pH 7.0) for 1 h. The DNA was fixed to the membrane by baking it in a vacuum oven at 80 °C for 1 h. Prehybridization, hybridization, washing and detection of the membranes were performed as described for the colony lifts.

Subcloning of JC1,JC2,JC3,JC7and JF2A.

Each *Bam*HI fragment for each of the four cosmids was subcloned into the *Bam*HI site of pBluescript KS + (Stratagene). Ligations were conducted at 16 °C for 24 h with T4 DNA ligase. Ligated plasmids were used to transform *E. coli* DH5α competent cells and plated onto LB/IPTG/Xgal ampicillin selective plates.

References

(1) Rudd, B. A.; Hopwood, D. A. Genetics of actinorhodin biosynthesis by *Streptomyces coelicolor* A3(2). *J. Gen. Microbiol.* **1979**, *114*, 35-43.

(2) Beppu, T. Signal transduction and secondary metabolism: prospects for controlling productivity. *Trends. Biotechnol.* **1995**, *13*, 264-269.

(3) Chater, K. F. Genetic regulation of secondary metabolic pathways in Streptomyces. *Ciba. Found. Symp.* **1992**, *171*, 144-156; discussion 156-162.

(4) Malpartida, F.; Hopwood, D. A. Molecular cloning of the whole biosynthetic pathway of a Streptomyces antibiotic and its expression in a heterologous host. *Nature* **1984**, *309*, 462-464.

(5) Lacalle, R. A.; Tercero, J. A.; Jimenez, A. Cloning of the complete biosynthetic gene cluster for an aminonucleoside antibiotic, puromycin, and its regulated expression in heterologous hosts. *Embo. J.* **1992**, *11*, 785-792.

(6) Yin, X.; O'Hare, T.; Gould, S. J.; Zabriskie, T. M. Identification and cloning of genes encoding viomycin biosynthesis from *Streptomyces vinaceus* and evidence for involvement of a rare oxygenase. *Gene* **2003**, *312*, 215-224.

(7) Malpartida, F.; Hallam, S. E.; Kieser, H. M.; Motamedi, H.; Hutchinson, C. R.; Butler, M. J.; Sugden, D. A.; Warren, M.; McKillop, C.; Bailey, C. R. Homology between Streptomyces genes coding for synthesis of different polyketides used to clone antibiotic biosynthetic genes. *Nature* **1987**, *325*, 818-821.

(8) Kieser, T. B., M.; Buttner, M.; Chater, K.; and Hopwood, D. *Practical Streptomyces Genetics*; 1*st* ed.; The John Innes Foundation: Nowich, 2000; 613.

(9) Neilan, B. A.; Dittmann, E.; Rouhiainen, L.; Bass, R. A.; Schaub, V.; Sivonen, K.; Borner, T. Nonribosomal peptide synthesis and toxigenicity of cyanobacteria. *J. Bacteriol.* **1999**, *181*, 4089-4097.

(10) Rajendran, N. Identification and cloning of a gene locus encoding peptide synthetase of *Pseudomonas fluorescens* by two sets of PCR primers. *Z. Naturforsch [C]* **1999**, *54*, 105-109.

(11) Borchert, S.; Patil, S. S.; Marahiel, M. A. Identification of putative multifunctional peptide synthetase genes using highly conserved oligonucleotide sequences derived from known synthetases. *FEMS Microbiol. Lett.* **1992**, *71*, 175-180.

(12) Gaitatzis, N.; Hans, A.; Muller, R.; Beyer, S. The mtaA gene of the myxothiazol biosynthetic gene cluster from *Stigmatella aurantiaca* DW4/3-1 encodes a phosphopantetheinyl transferase that activates polyketide synthases and polypeptide synthetases. *J. Biochem. (Tokyo)* **2001**, *129*, 119-124.

(13) Beyer, S.; Kunze, B.; Silakowski, B.; Muller, R. Metabolic diversity in myxobacteria: identification of the myxalamid and the stigmatellin biosynthetic gene cluster of *Stigmatella aurantiaca* Sg a15 and a combined polyketide-(poly)peptide gene cluster from the epothilone producing strain *Sorangium cellulosum* So ce90. *Biochim. Biophys. Acta.* **1999**, *1445*, 185-195.

(14) Silakowski, B.; Schairer, H. U.; Ehret, H.; Kunze, B.; Weinig, S.; Nordsiek, G.; Brandt, P.; Blocker, H.; Hofle, G.; Beyer, S.; Muller, R. New lessons for combinatorial biosynthesis from myxobacteria. The myxothiazol biosynthetic

gene cluster of *Stigmatella aurantiaca* DW4/3-1. *J. Biol. Chem.* **1999**, *274*, 37391-37399.

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

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

(17) Berger, S. L.; Kimmel, A. R. Guide to molecular cloning techniques. *Methods enzymol.* **1987**, *152*, 410-413.

(18) Taylor, D. P.; Eckhardt, T.; Fare, L. R. Gene cloning in the actinomycetes. *Ann. N. Y. Acad. Sci.* **1983**, *413*, 47-56.

(19) Donadio, S.; Staver, M. J.; McAlpine, J. B.; Swanson, S. J.; Katz, L. Modular organization of genes required for complex polyketide biosynthesis. *Science* **1991**, *252*, 675-679.

(20) August, P. R.; Tang, L.; Yoon, Y. J.; Ning, S.; Muller, R.; Yu, T. W.; Taylor, M.; Hoffmann, D.; Kim, C. G.; Zhang, X.; Hutchinson, C. R.; Floss, H. G. Biosynthesis of the ansamycin antibiotic rifamycin: deductions from the molecular analysis of the rif biosynthetic gene cluster of *Amycolatopsis mediterranei* S699. *Chem. Biol.* **1998**, *5*, 69-79.

(21) Silakowski, B.; Nordsiek, G.; Kunze, B.; Blocker, H.; Muller, R. Novel features in a combined polyketide synthase/non-ribosomal peptide synthetase: the myxalamid biosynthetic gene cluster of the myxobacterium *Stigmatella aurantiaca* Sga15. *Chem. Biol.* **2001**, *8*, 59-69.

(22) Haydock, S. F.; Aparicio, J. F.; Molnar, I.; Schwecke, T.; Khaw, L. E.; Konig, A.; Marsden, A. F.; Galloway, I. S.; Staunton, J.; Leadlay, P. F. Divergent sequence motifs correlated with the substrate specificity of (methyl)malonyl-CoA:acyl carrier protein transacylase domains in modular polyketide synthases. *FEBS Lett.* **1995**, *374*, 246-248.

(23) Christiansen, G.; Fastner, J.; Erhard, M.; Borner, T.; Dittmann, E. Microcystin biosynthesis in planktothrix: genes, evolution, and manipulation. *J. Bacteriol.* **2003**, *185*, 564-572.

(24) Yu, S.; Fiss, E.; Jacobs, W. R., Jr. Analysis of the exochelin locus in *Mycobacterium smegmatis*: biosynthesis genes have homology with genes of the peptide synthetase family. *J. Bacteriol.* **1998**, *180*, 4676-4685.

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

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

(27) Hoffmann, D.; Hevel, J. M.; Moore, R. E.; Moore, B. S. Sequence analysis and biochemical characterization of the nostopeptolide A biosynthetic gene cluster from Nostoc sp. GSV224. *Gene* **2003**, *311*, 171-180.

(28) Pospiech, A.; Bietenhader, J.; Schupp, T. Two multifunctional peptide synthetases and an O-methyltransferase are involved in the biosynthesis of the DNA-binding antibiotic and antitumour agent saframycin Mx1 from *Myxococcus xanthus*. *Microbiology* **1996**, *142 (Pt 4)*, 741-746.

(29) Marchler-Bauer A, A. J., DeWeese-Scott C, Fedorova ND, Geer LY, He S, Hurwitz DI, Jackson JD, Jacobs AR, Lanczycki CJ, Liebert CA, Liu C, Madej T, Marchler GH, Mazumder R, Nikolskaya AN, Panchenko AR, Rao BS, Shoemaker BA, Simonyan V, Song JS, Thiessen RA, Vasudevan S, Wang Y, Yamashita RA, Yin JJ, and Bryant SH CDD: a curated Entrez database of conserved domain alignments. *Nucleic Acids Res.* **2003**, *31*, 383-387.

(30) Chang, Z.; Flatt, P.; Gerwick, W. H.; Nguyen, V. A.; Willis, C. L.; Sherman, D. H. The barbamide biosynthetic gene cluster: a novel marine cyanobacterial system of mixed polyketide synthase (PKS)-non-ribosomal peptide synthetase (NRPS) origin involving an unusual trichloroleucyl starter unit. *Gene* **2002**, *296*, 235-247.

(31) Gourley, D. G.; Shrive, A. K.; Polikarpov, I.; Krell, T.; Coggins, J. R.; Hawkins, A. R.; Isaacs, N. W.; Sawyer, L. The two types of 3-dehydroquinase have distinct structures but catalyze the same overall reaction. *Nat. Struct. Biol.* **1999**, *6*, 521-525.

(32) Moore, B. S.; Hertweck, C. Biosynthesis and attachment of novel bacterial polyketide synthase starter units. *Nat. Prod. Rep.* **2002**, *19*, 70-99.

(33) Yu, T. W.; Muller, R.; Muller, M.; Zhang, X.; Draeger, G.; Kim, C. G.; Leistner, E.; Floss, H. G. Mutational analysis and reconstituted expression of the biosynthetic genes involved in the formation of 3-amino-5-hydroxybenzoic acid, the starter unit of rifamycin biosynthesis in *Amycolatopsis mediterranei* S699. *J. Biol. Chem.* **2001**, *276*, 12546-12555.

(34) Kim, C. G.; Kirschning, A.; Bergon, P.; Zhou, P.; Su, E.; Sauerbrei, B.; Ning, S.; Ahn, Y.; Breuer, M.; Leistner, E.; Floss, H. G. Biosynthesis of 3-Amino-5-hydroxybenzoic Acid, the Precursor of mC_7N Units in Ansamycin Antibiotics. *J. Am. Chem. Soc.* **1996**, *118*, 7486-7491.

(35) Piel, J.; Hertweck, C.; Shipley, P. R.; Hunt, D. M.; Newman, M. S.; Moore, B. S. Cloning, sequencing and analysis of the enterocin biosynthesis gene cluster from the marine isolate '*Streptomyces maritimus*': evidence for the derailment of an aromatic polyketide synthase. *Chem. Biol.* **2000**, *7*, 943-955.

(36) Xiang, L.; Moore, B. S. Characterization of benzoyl coenzyme A biosynthesis genes in the enterocin-producing bacterium "*Streptomyces maritimus*". *J. Bacteriol.* **2003**, *185*, 399-404.

(37) Priefert, H.; Hein, S.; Kruger, N.; Zeh, K.; Schmidt, B.; Steinbuchel, A. Identification and molecular characterization of the *Alcaligenes eutrophus* H16 aco operon genes involved in acetoin catabolism. *J. Bacteriol.* **1991**, *173*, 4056-4071.

(38) Kruger, N.; Oppermann, F. B.; Lorenzl, H.; Steinbuchel, A. Biochemical and molecular characterization of the *Clostridium magnum* acetoin dehydrogenase enzyme system. *J. Bacteriol.* **1994**, *176*, 3614-3630.

(39) Kawai, S.; Mori, S.; Mukai, T.; Hashimoto, W.; Murata, K. Molecular characterization of *Escherichia coli* NAD kinase. *Eur. J. Biochem.* **2001**, *268*, 4359-4365.

(40) Arakawa, K.; Muller, R.; Mahmud, T.; Yu, T. W.; Floss, H. G. Characterization of the early stage aminoshikimate pathway in the formation of 3-amino-5-hydroxybenzoic acid: the RifN protein specifically converts kanosamine into kanosamine 6-phosphate. *J. Am. Chem. Soc.* **2002**, *124*, 10644-10645.

(41) Shen, Z.; Byers, D. M. Isolation of *Vibrio harveyi* acyl carrier protein and the fabG, acpP, and fabF genes involved in fatty acid biosynthesis. *J. Bacteriol.* **1996**, *178*, 571-573.

(42) Winter, E.; Brummel, M.; Schuch, R.; Spener, F. Decarboxylation of malonyl-(acyl carrier protein) by 3-oxoacyl-(acyl carrier protein) synthases in plant fatty acid biosynthesis. *Biochem. J.* **1997**, *321 (Pt 2)*, 313-318.

(43) Du, L.; Sanchez, C.; Chen, M.; Edwards, D. J.; Shen, B. The biosynthetic gene cluster for the antitumor drug bleomycin from *Streptomyces verticillus* ATCC15003 supporting functional interactions between nonribosomal peptide synthetases and a polyketide synthase. *Chem. Biol.* **2000**, *7*, 623-642.

(44) Zhang, J. H.; Quigley, N. B.; Gross, D. C. Analysis of the syrP gene, which regulates syringomycin synthesis by *Pseudomonas syringae pv. syringae*. *Appl. Environ. Microbiol.* **1997**, *63*, 2771-2778.

(45) Mootz, H. D.; Marahiel, M. A. The tyrocidine biosynthesis operon of Bacillus brevis: complete nucleotide sequence and biochemical characterization of functional internal adenylation domains. *J. Bacteriol.* **1997**, *179*, 6843-6850.

(46) Becker, J. E.; Moore, R. E.; Moore, B. S. Cloning, sequencing, and biochemical characterization of the nostocyclopeptide biosynthetic gene cluster: molecular basis for imine macrocyclization. *Gene* **2004**, *325*, 35-42.

(47) Akee, R.; Cartner, L.; McCloud, T.; Colin, P.; Newman, D. Distribution and Preparative Isolation of Jaspamide, a Marine Anticancer Cyclic Peptide. *American Society of Pharmacognosy- 38th Annual Meeting-July 26-30*: Iowa City, IA, 1997.

(48) Omura, S.; Ikeda, H.; Ishikawa, J.; Hanamoto, A.; Takahashi, C.; Shinose, M.; Takahashi, Y.; Horikawa, H.; Nakazawa, H.; Osonoe, T.; Kikuchi, H.; Shiba, T.; Sakaki, Y.; Hattori, M. Genome sequence of an industrial microorganism *Streptomyces avermitilis*: deducing the ability of producing secondary metabolites. *Proc. Natl. Acad. Sci. U S A* **2001**, *98*, 12215-12220.

(49) Pradella, S.; Hans, A.; Sproer, C.; Reichenbach, H.; Gerth, K.; Beyer, S. Characterisation, genome size and genetic manipulation of the myxobacterium *Sorangium cellulosum* So ce56. *Arch. Microbiol.* **2002**, *178*, 484-492.

(50) Konz, D.; Marahiel, M. A. How do peptide synthetases generate structural diversity? *Chem. Biol.* **1999**, *6*, R39-48.

(51) Sambrook, J.; Russell, D. *Molecular Cloning A Laboratory Manual*; 3 ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, 2001.

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

Chapter 3
Expression of Adenylation Domains

Introduction

Sequence analysis of the cloned *Jaspis* mixed PKS/NRPS gene cluster alone was not enough to clearly determine if the cluster was involved in the biosynthesis of jaspamide. In order to further investigate the function of this gene cluster, biochemical characterization of the pathway needed to be completed. Characterization of each of the three types of biosynthesis genes found within the *Jaspis* cluster, however, would provide differing amounts of information into the nature of the gene cluster.

The PKS portion of the gene cluster was not chosen for biochemical studies because it was expected to provide only limited information on the function of the biosynthetic pathway. This is largely due to the manner in which PKSs generate the structural diversity seen in their products. The structure of a polyketide PKS product is largely determined by the choice of extender units for each elongation step and the varying degree of reduction that occurs on the incorporated ketide. As PKSs typically utilize only two types of extender units (malonyl or methylmalonyl CoA) and the degree of ketide reduction can be predicted based on the presence of auxiliary domains, analysis of the substrate specificity of the A•PCP-domains was not likely to provide new information. The ORFs located upstream of the PKS gene are predicted to be involved in biosynthesis of the PKS starter unit, and could provide significant insight into the nature of the produced metabolite. A large amount of structural diversity seen in polyketides is generated by the type of molecule used as a starter unit. Although the type of chemistry catalyzed by these domains can be predicted, both the substrates and the order of operation for each of these enzymes can not. This would make characterization of these genes quite difficult.

The remaining option for initial biochemical characterization of the *Jaspis* mixed PKS/NRPS gene cluster was the NRPS. One method commonly used to characterize

NRPS gene clusters is the amplification and heterologous expression of the adenylation domains (A-domains).[1-4] It has been found that the A-domains are largely responsible for the substrate specificity of these enzymes and characterization should provide the most useful information about the biosynthetic pathway and the nature of the metabolite produced.[1,5] Once solubly expressed, the adenylation domains can be analyzed for their substrate specificity through ATP-$^{32}PP_i$ exchange assays. Through these assays it is possible to determine the nature, or in some cases the identity, of the amino acid substrate specified by a given A-domain. From this information, a better estimation of the proposed metabolite produced by this gene cluster can be made.

Results and Discussions

Design and Construction of A-domain Expression Vectors.

The expression of adenylation domains from NRPS gene clusters has been met with varying success. Problems associated with the heterologous expression of A-domains have ranged from the inability to see any expression to the observed inactivity of the overexpressed protein. When Recktenwald *et al.* first attempted to overexpress A-domains from the balhimycin gene cluster in *E. coli* they were unable to detect any production of the desired proteins.[6] Unlike the balhimycin A-domains, low levels of soluble expression were achieved for an A-domain from the pristinamycin gene cluster however the protein was found to be catalytically inactive.[7]

Lack of soluble overexpressed A-domains is a significant obstacle in the heterologous expression of NRPS gene clusters. Many approaches have been taken to try and overcome this issue with varying success. One approach is to use alternate surrogate host cells for expression. This approach was successfully used for the functional expression of balhimycin A-domains in a *Streptomyces* host system.[6] Another approach is to coexpress accessory proteins that can promote proper protein folding. Duitman *et al.* found that coexpression of the GroEL/ES chaperones allowed for the soluble expression of MycB, an A-domain that is involved in mycosubtilin biosynthesis.[8]

Improper definition of A-domain boundaries has been proposed to be a common cause for the observed insolubility and inactivity of heterologously expressed A-

domains.[9] NRPSs are multimodular, multifunctional proteins containing multiple catalytic domains within a single ORF. The size of native NRPSs are often too large to make heterologous expression of the entire ORF practical. In addition, the presence of more than one module on a multifunctional protein makes characterization of individual activities difficult. Consequently, detailed characterization of an NRPS requires the amplification and heterolgous expression of individual catalytic domains. Excision of these domains from the NRPS modules necessarily requires the identification of the domain boundaries, which often can not be clearly defined by sequence analysis alone. The definition of domain boundaries has generated difficulties not only in the heterologous expression of individual A-domains, but also has generated a number of problems with combinatorial biosynthesis of nonribosomal peptides.[7,9-11]

Many researchers have avoided this problem by expressing domains that are at, or near, the 5'-terminus of the ORF. This allows for use of the native start codon and has been successfully used to express a number of A-domains involved in the biosynthesis of antibiotics such as bleomycin and barbamide.[4,12]

Expression starting with the native start codon is not always practical with large NRPS systems or with all A-domains. In ORFs that contain multiple extension modules, one or more A-domain are internally located within the large multifunctional protein. The earliest attempts to heterologously express internal A-domains resulted in the formation of a "100 aa" rule.[9] This method was used to define the boundaries of five internal A-domains from tyrocidine synthetase. This method defines the *N*-terminus of the A-domain to be located 95-100 amino acids upstream of the conserved A2 sequence motif (KAGGA). Similarly, the C-terminus of the domain is defined as being 100 amino acids downsteam of the conserved A3 sequence motif (RIELGEIE).[9]

The successful expression of a number of A-domains from various NRPS gene clusters allows for modification of this method. Strict utilization of the 100 aa rule does not account for the variation seen in non-conserved regions of the A-domains. These regions may contain variable numbers of amino acids and result in the incorrect definition of domain boundaries. Sequence alignments, however, can be used to compensate for these differences because the presence of highly conserved sequence

motifs serve as alignment anchors. Comparison of the domain boundaries used for the successful characterization of other A-domains with the target A-domain provides an alternate method for identifying putative A-domain boundaries. This method was used to define A-domain boundaries for the *Jaspis* NRPS.

The *Jaspis* NRPS has two internal A-domains, as well as an *N*-terminal A-domain that is preceded by a C-domain. In order to design expression constructs for the three *Jaspis* adenylation domains, each domain was aligned with several successfully expressed A-domains from other gene clusters. Only A-domains that are internally located in their respective ORFs were used for alignment and included balhimycin Mod 4-8[6] and bleomycin NRPS 6.[12] Using these alignments, the boundaries of the adenylation domains was determined as is shown in Figure 3.1.

A-domain expression constructs that contain the corresponding PCP-domain, as well as those that do not, have been reported in the literature. Since we wanted to optimize our chances for obtaining soluble protein we initially designed two expression constructs for each of the three A-domains, resulting in expression both with and without the corresponding PCP-domains. As the *N*-terminal A-domains tend to be easier to heterologously express, expression constructs for the first A-domain, both with and without the corresponding downstream PCP domain, were constructed for expression from the native ORF start site. These constructs contained the upstream C domain found at the beginning of the ORF, however similar architecture is found in one characterized bleomycin A-domain suggesting that functional expression was possible.[12]

Once the boundaries for the expression constructs had been determined, PCR primers were designed to amplify the corresponding region from the cloned NRPS gene cluster. Constructs were made in a pET20b+, pET28a+, and pET41b+ expression systems (Table 3.1).

Table 3.1. Expression vectors used in this work

Vector	Features	His Tag	Antibiotic Resistance
pET20b+	Optional pelB leader	*C*-Terminal	Ampicillin
pET28a+		*N*-Terminal	Kanamycin
pET41b+	*N*-Terminal GST	*C*-Terminal	Kanamycin

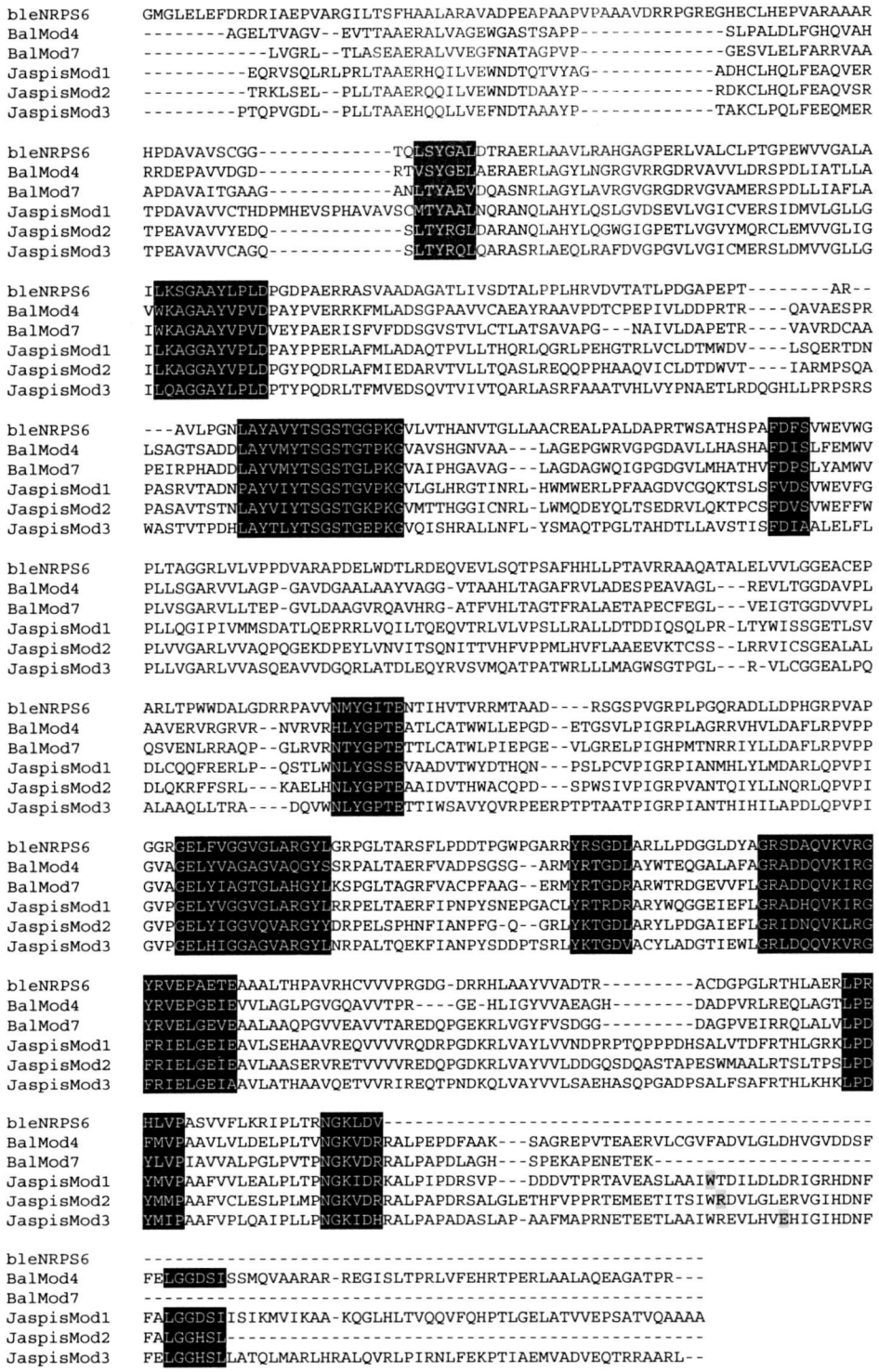


Figure 3.1. Alignment of *Jaspis* adenylation domain expression constructs with successfully expressed domains from other NRPS clusters. Highly conserved sequence motifs are highlighted in black. Stop sites used for expression of just the A-domains are highlighted in grey. bleNRPS6 is from the bleomycin gene cluster while BalMOD4 and BalMod7 are from the balhimycin gene cluster.

Expression and Purification of A-domains

One issue that must be addressed when using a heterologous expression system is the codon bias of both the source and host organisms. Expression of genes containing codons that are rare in the surrogate host can result in translational errors and result in less active or inactive protein.[13] Codons that are not commonly found in *E. coli* genes include the arginine codons AGA and AGG as well as the proline codon CCC. In *E. coli* K12 these codons occur at a rate of only1.2 codons /1000 for AGG of and 5.5 codons/1000 for CCC.[14] Several of these codons are frequently used in the cloned *Jaspis* PKS/NRPS gene cluster. The CCC codon, in particular, is frequently used in the *Jaspis* cluster and is found at a rate of up to 29 codons/1000 in the first A-domain. In order to overcome the *E. coli* codon bias, Rosetta (DE3) pLysS *E. coli* cells, available from Invitrogen, were used for the heterologous expression of the *Jaspis* NRPS A-domains. The Rosetta cell line contains extra copies of genes for rare *E. coli* tRNAs and allows for the improved expression of GC rich genes.

Once the boundaries for the expression constructs had been determined, PCR primers were designed to amplify the corresponding region from the cloned NRPS gene cluster. The original constructs were made in the pET20b+ expression system (Figure 3.2). Expression from this vector results in the addition of a *C*-terminal His_6-tag as well as an optional pelB leader to the target protein.

The cloning sites used in the construction of the initial expression plasmids removed the *N*-terminal pelB leader and allowed for expression of the A-domain, or A-PCP-domain, with a *C*-terminal His_6-tag (Figure 3.2, A). Only the first two A-domains were cloned into this vector because initial expression studies did not result in soluble protein even after several rounds of optimization, including expression at 37, 25, and 18 °C and inducing with varying levels of IPTG. When the initial expression studies were unsuccessful, a new series of expression constructs were generated in pET20b+ that utilized alternate cloning sites and resulted in expression of the proteins with both the *C*-terminal His_6-tag and an *N*-terminal pelB leader (Figure 3.2B). The pelB leader has two functions as a fusion protein. First, it provides several *E. coli* codons that allow for the efficient initiation of translation and may enhance the heterologous expression of genes

containing high GC content or unusual codons. The pelB leader also serves as a signal sequence that directs the transport of the protein to the periplasmic space.[15] Export of proteins into the periplasm has been shown to improve solubility by providing an environment that is more amendable to protein folding.[16] The pelB leader is typically cleaved from the target protein by signal peptidase during export.

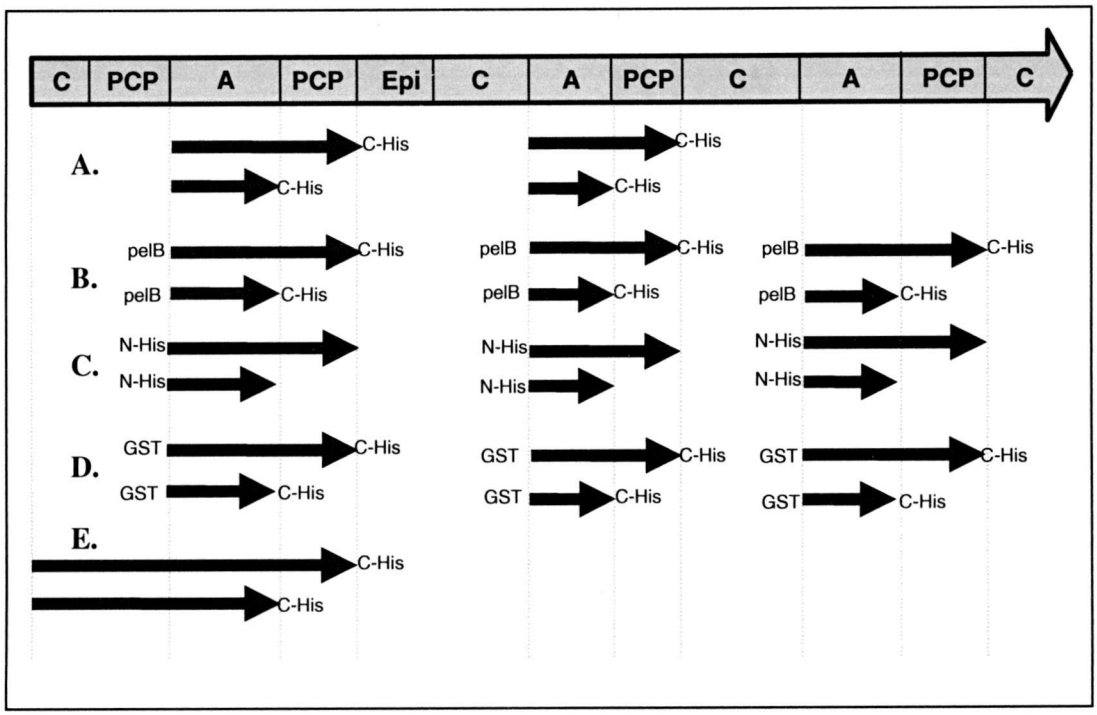


Figure 3.2. A-domain expression constructs used in this work. A. Construction of *C*-terminal His_6-tag proteins in pET20b+. B. Construction of *N*-terminal pelB fusion and *C*-terminal His_6-tag proteins in pET20b+. C. Construction of *N*-terminal GST fusion and *C*-terminal His_6-tag proteins in pET41b+. D. Construction of *N*-terminal His_6-tag proteins in pET28a. E. Construction of *C*-terminal His_6-tag proteins from the native start codon in pET20b+

Expression of the pET20pelB-A-domain and A•PCP-domain constructs in *Rosetta* (DE3) pLysS did result in increased expression of the desired protein, however solubility was still an issue. Our limited success with the pelB fusions led us to construct two additional series of expression plasmids with alternate *N*-terminal fusions. One possibility for the observed insolubility may result from inefficient cleavage of the pelB fusion and subsequent interference of the fusion with A-domain folding. In order to avoid possible interference of the fusion protein, and retain the translational advantage of

an *N*-terminal leader, a second set of expression constructs were made in pET28a (Figure 3.2,C). Expression from pET28A results in the fusion of an *N*-terminal His_6-tag that is significantly smaller than the pelB fusion.

An alternate approach was to use a larger soluble fusion protein that can promote solubility of the target protein. Expression with large fusion proteins has been shown to increase the yield of soluble target protein.[17,18] The pET41b+ expression vector was used to test this possibility (Figure 3.2, D). Overexpression in pET41b+ results in the production of an *N*-terminal fusion of a 25 kDa glutathione S transferase (GST) as well as a *C*-terminal His_6-tag. An added advantage of the GST-tag is that it can be used for protein purification in conjunction with glutathione affinity chromatography.

Clear overexpression of the A-domains was achieved with the pET28 and pET41 expression systems, however it was not clear if any of the protein was soluble. The levels of soluble protein were not sufficient for easy visualization by SDS-PAGE. In order to check for soluble protein, the cleared lysate from the initial expressions was passed over a nickel affinity resin to purify any His-tagged protein that may be present. It was found that the *N*-terminal His_6-tag fusion proteins did not bind well to the chromatography resin and could be eluted with low concentrations of imidazole. Consequently, only partial purification of the A-domains could be achieved. SDS-PAGE analysis of the resulting protein showed the presence of several bands (Figure 3.3 A). The size of one of the bands was consistent with the predicted size of the adenylation (or A•PCP) domains,

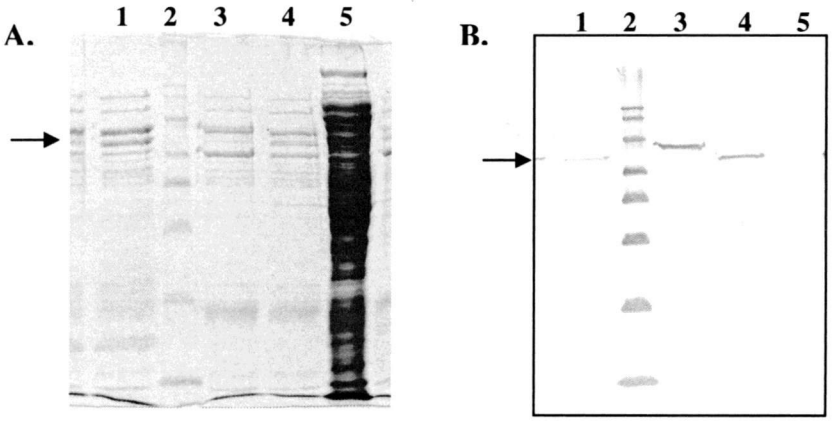


Figure 3.3. Proteins expressed from pET28a constructs with an *N*-terminal His_6-Tag fusion protein in *E. coli* 'Rosetta'. Lane: **1**. Purified M3AT **2.** Magic Mark Western Ladder (Invitrogen) **3.** Purified M1AT **4.** Purified M2AT **5**. pET28a vector control CFE

however due to the presence of contaminating protein, the results were not clear. In order to confirm that the A-domains were soluble and had been successfully expressed, the purified proteins were analyzed by Western analysis using Novagen's His•Tag® Western detection kit. This method utilizes a His-tag mouse monoclonal antibody (IgG_1) coupled with an anti-mouse IgG alkaline phosphatase (AP) secondary antibody. Western analysis clearly demonstrated the presence of a single band at the predicted molecular weight and confirmed the presence of soluble protein for each of the three adenylation domains (Figure 3.3, B).

Although soluble protein could be obtained from both the pET28a and the pET41 constructs, the pET28a constructs were used for the *in vitro* assays. The primary reason for this was that the *N*-terminal fusions on proteins produced with the pET28a system were much smaller than the pET41 GST fusion. For the *in vitro* substrate activity assays, the GST fusion protein would need to be cleaved from the A-domain. This would require an additional purification step and the resulting A-domain would still contain a 5 kDa S-tag fusion. Given the relatively low levels of expression, these extra purification steps would further reduce the total amount of protein available for the assays. In contrast, purified protein from the pET28a expression system, although not homogeneous, could be used directly for substrate selectivity assays.

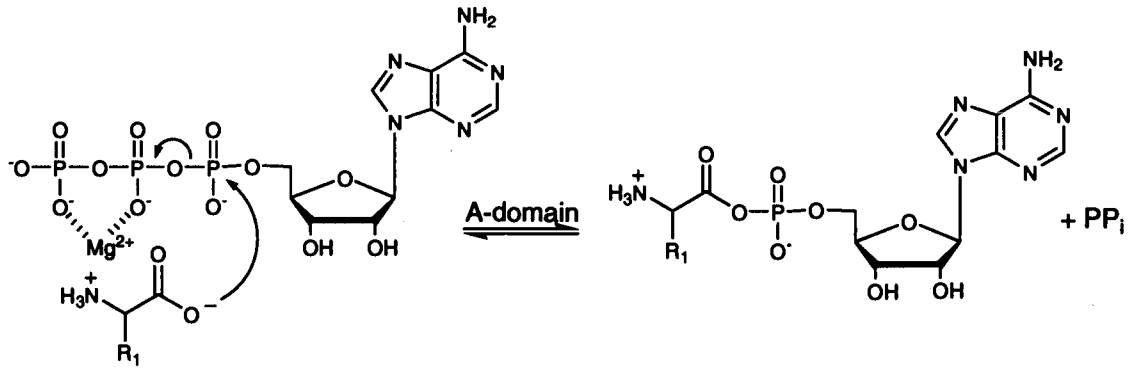
Another issue that needed to be addressed prior to *in vitro* analysis of the A•PCP-domains was the generation of *holo*-enzyme. As discussed in chapter 1, the ribosomal synthesis of peptide synthetases results in the formation of *apo* enzyme. Under native conditions the NRPS module is modified through the addition of a phosphopantetheine cofactor to the PCP-domain. Although adenylation domains have been shown to be functional in absence of phosphopantetheinylated PCP-domain the holoenzyme is more representative of the native protein.[19,20] *E. coli* does not possess a phosphopantetheinyl transferase that is capable of forming *holo* PCP-domains and this function must be supplied *in trans*. *B. subtilis* surfactin phosphopantetheinyl transferase, Sfp, exhibits very relaxed substrate specificity and is capable phosphopantetheinylating NRPS PCP-domains as well as PKS ACP-domains.[21,22] The formation of *holo*-PCP-domains is typically accomplished by coexpression of *sfp* with the NRPS PCP-domain in *E. coli*.

B. subtillus sfp in a pUC8 expression vector was obtained as a gift from Professor P. Zuber and used to generate an *E. coli* cell line for the expression of adenylation domains. The plasmid pUC8*sfp* was introduced into *E. coli* Rosetta (DE3) pLysS and competent cells were then generated from the resulting cell line, Rosetta (DE3)pLysS/pUC8*sfp* and transformed with the three pET28a expression constructs, p28M1AT, p28M2AT, and p28M3AT.

In order to purify an adequate quantity of protein for *in vitro* analysis, one liter cultures of each construct were used. It was found that higher levels of soluble protein could be obtained from cultures where protein expression was not induced by addition of IPTG. Also, it was necessary to grow the cells at 18-20 °C in order to get soluble production of the A•PCP-domains.

Determination of Adenylation Domain Substrate Specificity

The substrate specificity for each of the three *Jaspis* A-domains was determined through an ATP-$^{32}PP_i$ exchange assay. This assay takes advantage of the reversibility of the adenylation reaction and has been described by Lee *et al.* and Stachelhaus *et al.*[1,23] Enzyme is combined with a potential substrate amino acid in the presence of radiolabeled inorganic pyrophosphate ($^{32}PP_i$). If the enzyme can use the amino acid as a substrate it will form the corresponding acyladenylate and inorganic pyrophosphate as shown in Scheme 3.1. The acyladenylate can undergo the reverse reaction to reform the free amino acid and ATP. One consequence of the reversibility of this reaction is the incorporation of the $^{32}PP_i$ into ATP. The assay is incubated for such a time as to allow equilibrium between the forward and reverse reactions to occur. Following quenching, the ATP in the assay mixture is purified by binding to charcoal. Any radioactivity not incorporated into ATP is then washed away and the amount of bound radioactivity, corresponding to the amount of $AT^{32}P$, is counted with a scintillation counter. The amount of $AT^{32}P$ formed corresponds to the ability of the A-domain to activate the given amino acid. The relative activity of the enzyme with the various amino acids can then be determined and used to predict substrate specificity.



Scheme 3.1. Activation of an amino acid by the A-domain

Each adenylation domain was tested for its ability to activate the twenty common amino acids as well as several nonproteinogenic amino acids. In the first set of assays partially purified A•PCP-domains were used. In order to minimize the use of radioactive pyrophosphate and to conserve protein, initial assays were conducted with multiple amino acid substrates. The activity of each A•PCP-domain with the various amino acid mixes (counts per minute, cpm) was converted to relative percent activity (Figure 3.4). In addition to the three A•PCP-domains, two control A-domains from the tyrocidine biosynthetic gene cluster, ProA2 and TyrA7, were also analyzed.[9] It was found that all three *Jaspis* A-domains had the highest level of activity with the aliphatic amino acid mix (Mix 1: glycine, alanine, valine, leucine, isoleucine, and proline). The second highest level of activity for M1AT and M2AT was seen with the basic amino acid mixture (Mix 2: lysine, arginine, glutamine, asparagine, and histidine) followed by the Mix 4 (serine, methionine, cysteine, threonine, glutamate, and aspartate). M3AT, however had the second highest level of activity with Mix 4, followed by aliphatic Mix 2. The presence of β-tyrosine in jaspamide, as well as the predicted substrate specificity of the first A-domain, prompted us to test for the ability each of the three A-domains to activate this non-proteinogenic amino acid. As seen in Figure 3.4, none of the three domains had significant activity with β-tyrosine.

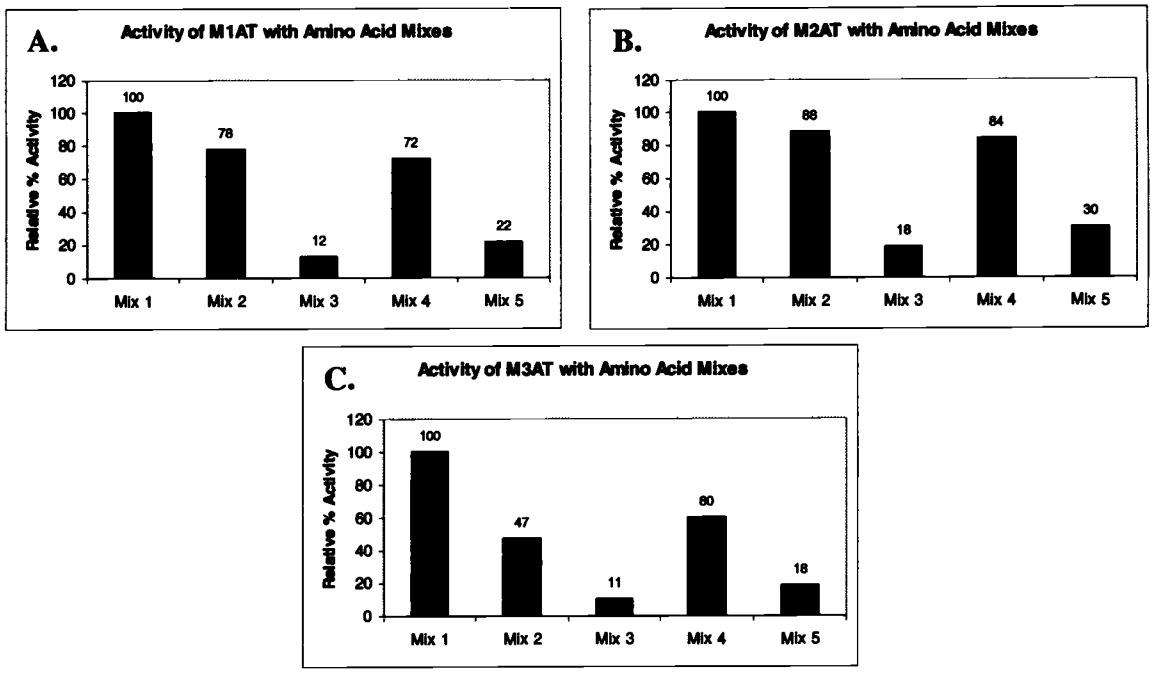


Figure 3.4. Relative activity of A•PCP-domains with combinations of amino acids. A. Activity of the A•PCP-domain from the first NRPS module (M1AT). B. Activity of the A•PCP-domain from the second NRPS module (M2AT). C. Activity of the A-PCP-domain from the third NRPS module (M3AT). Mix 1: glycine, alanine, valine, leucine, isoleucine, and proline. Mix 2: lysine, arginine, glutamine, asparagine, and histidine. Mix 3: β-tyrosine. Mix 4: serine, methionine, cysteine, threonine, glutamate, and aspartate. Mix 5: tryptophan and phenylalanine

In order to further investigate substrate specificity, each of the three A•PCP-domains was assayed with individual amino acids from the mixtures that appeared to contain the best substrates. The results of these assays are shown in Figure 3.5. The highest level of activity for all three A-domains was seen with leucine, however differences between the domains were seen in substrates with the second highest levels of activation. M1AT exhibits adenylation activity with valine, isoleucine and lysine. M3AT followed a similar trend and activates threonine, isoleucine and valine in addition to leucine. M2AT has interesting substrate selectivity as it activates lysine, as well as valine and isoleucine nearly as well as it activated leucine. M2AT also had the broadest substrate specificity in that it demonstrated the same level of adenylation with four amino acids. In order to investigate this further, M2AT was assayed with three lysine homologs (ornithine, 2,3-diaminoproprionic acid (DAP), and 3,4-diaminobutryic acid (DAB)) however significant levels of activation were not observed (Figure 3.5).

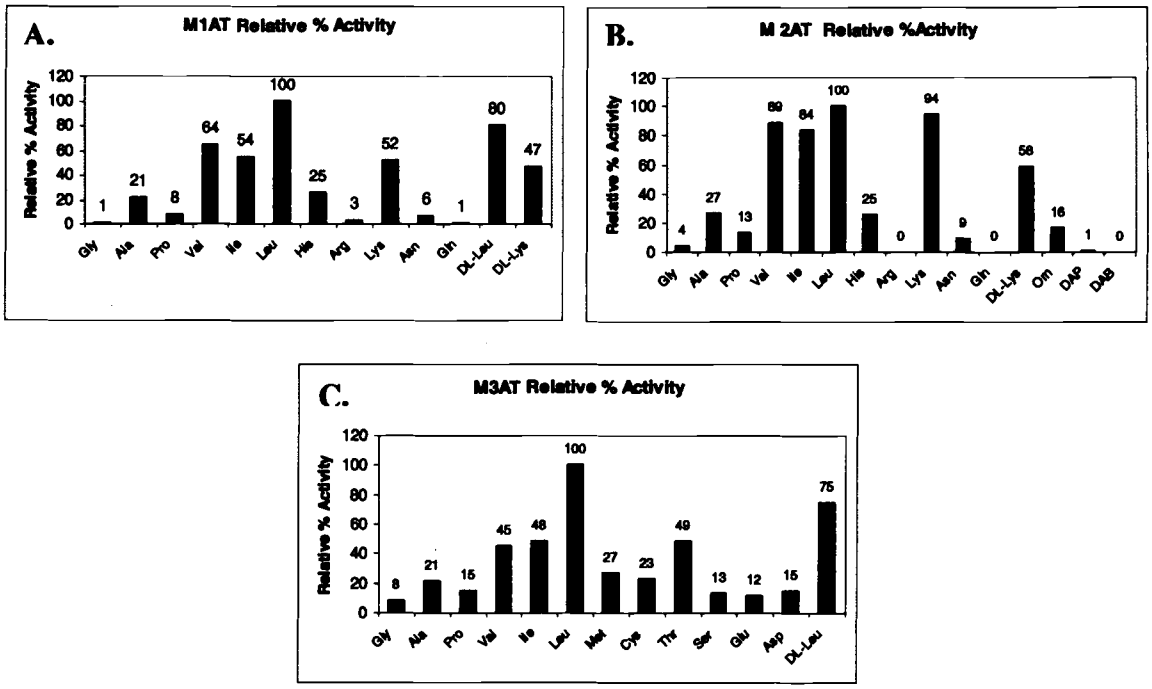


Figure 3.5. Relative activity of A•PCP-domains with individual amino acids. A. Activity of the A•PCP-domain from the first NRPS module (M1AT). B. Activity of the A•PCP-domain from the second NRPS module (M2AT). C. Activity of the A•PCP-domain from the third NRPS module (M3AT).

The stereoselectivity of each A•PCP-domain was also investigated by incubation of the enzymes with D,L-leucine. M2AT was also assayed with D,L-lysine. In all cases the level of activity was lower than the activity with just the L-isomer, indicating that the L-isomer is the preferred substrate.

In order to confirm the results of these substrate selectivity assays, the experiments were repeated with A-domains not fused with PCP-domains. Although the individual domains within a peptide synthetase module carry out specific catalytic functions, they do not function as autonomous units. This has been demonstrated in a number of studies by the groups of Marahiel and Walsh.[1,24,25] Other catalytic domains within an NRPS module can have an influence on the identity of the amino acid being incorporated into the product. Although the most significant influence on amino acid specificity was seen with the C-domain, we decided to repeat the *in vitro* ATP-$^{32}PP_i$ exchange assays using expression constructs that coded for only the A-domain of each module. Construction of pET28M1A, pET28M2A and p28M3A was achieved as

described above. As these constructs did not contain the PCP-domain, coexpression with *sfp* was not necessary.

ATP-$^{32}PP_i$ exchange assays using just the A-domains confirmed the overall trends seen in the first round of experiments, however there were some slight differences in substrate specificity (Figure 3.6). Preliminary assays with amino acid mixtures indicated that M2A retained the same pattern of substrate selectivity as was seen with M2AT. M1A and M3A both exhibited a shift in substrate selectivity when expressed in absence of the PCP-domain. In both cases, the preferred substrates shifted to the amino acid mix that had the second highest level of activity in the A•PCP-domain assays (Figure 3.4).

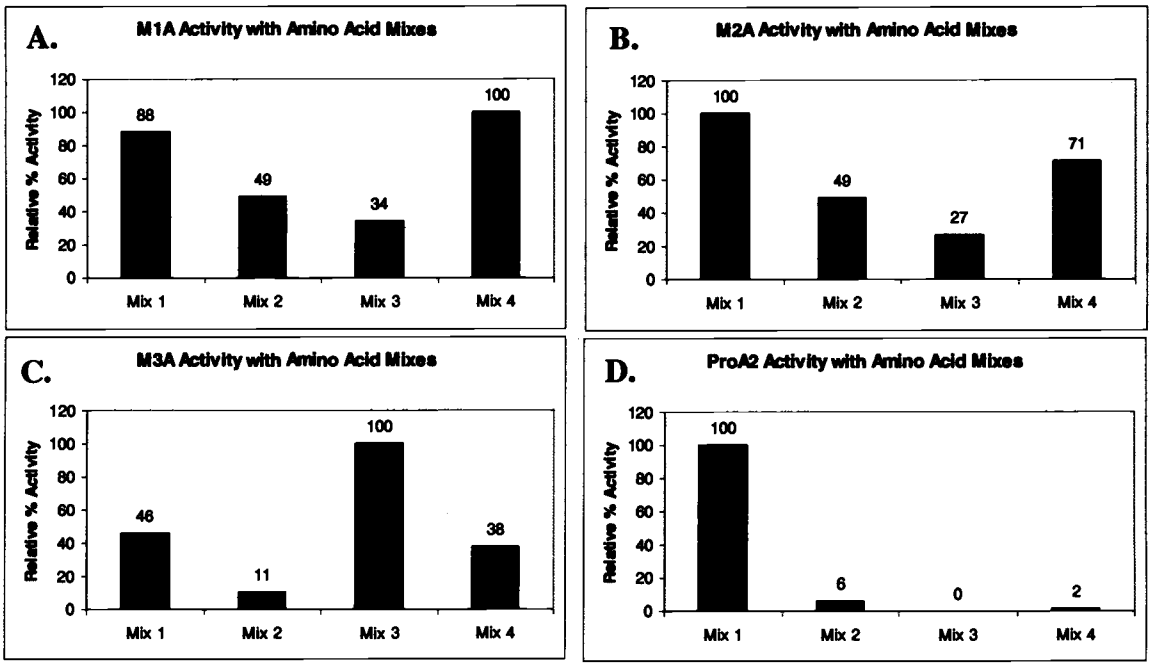


Figure 3.6. Relative activity of *Jaspis* A-domains with combinations of amino acids. A. Activity of the A- domain from the first NRPS module (M1A). B. Activity of the A-domain from the second NRPS module (M2A). C. Activity of the A -domain from the third NRPS module (M3A). D. Activity of the control A domain, ProA2 from the tyrocidine biosynthetic gene cluster. Mix 1: glycine, alanine, valine, leucine, isoleucine, and proline. Mix 2: serine, methionine, cysteine, threonine, glutamate, and aspartate. Mix 3: α-tyrosine, β-tyrosine, tryptophan and phenylalanine. Mix 4: lysine, arginine, glutamine, asparagine, and histidine

These observed shifts in substrate selectivity for M1A and M3A were confirmed in subsequent assays with individual amino acids. (Figure 3.7) M1A was found to preferentially activate lysine, as opposed to leucine that was the preferred substrate of

M1AT. The order of observed activity among the branch chain aliphatic amino acids was also slightly different and M1A showed the highest level of activity with isoleucine followed by valine and then leucine. When assayed with lysine homologs, M1A demonstrated only moderate activity with ornithine. One additional lysine analog, β-lysine, was also assayed with M1A. As discussed in chapter 2, the binding pocket of this first A-domain has the highly conserved aspartate residue, involved in binding the substrate amino group, shifted by one position. This led us to speculate that β-lysine may be a better substrate for M1A than lysine. The results of the assays, however showed that lysine was a better substrate than β-lysine.

M3A also demonstrated a significant shift in substrate specificity and had the highest level of activity with lysine. Surprisingly, the best substrate for M3AT, leucine, was found to be sixth in the order of M3A substrate preference, falling behind isoleucine, valine, proline, and threonine.

Although not apparent in preliminary assays with the amino acid mixtures, M2A also exhibited an alteration in substrate specificity, albeit the degree of change was not as severe as was seen for M1A and M3A. The preferred substrate for M2A was still a nonpolar branch chain amino acid, only valine was the preferred substrate followed by lysine, isoleucine and leucine. Aside from this shift in leucine activation, the relative order of substrate selectivity for M2A was unchanged.

Discussion

Analysis of the ATP-$^{32}PP_i$ exchange data does not clearly identify a single preferred substrate for any of the three domains. Unlike the *Jaspis* A-domains, the control proteins ProA2 and TyrA7 both exhibit significant substrate specificity for the natural substrate amino acid. ProA2, for example, had more than a 10 fold increase in activity with the cognate amino acid, proline, relative to other amino acids (Figure 3.6, D). This large difference in observed activity is consistent with the selectivity of other A-domains reported in the literature.[12,26] The *Jaspis* cluster A-domains, on the other hand, did not show significant substrate specificity with any of the amino acids tested. In addition, the overall level of adenylation activity for the three modules was in the same

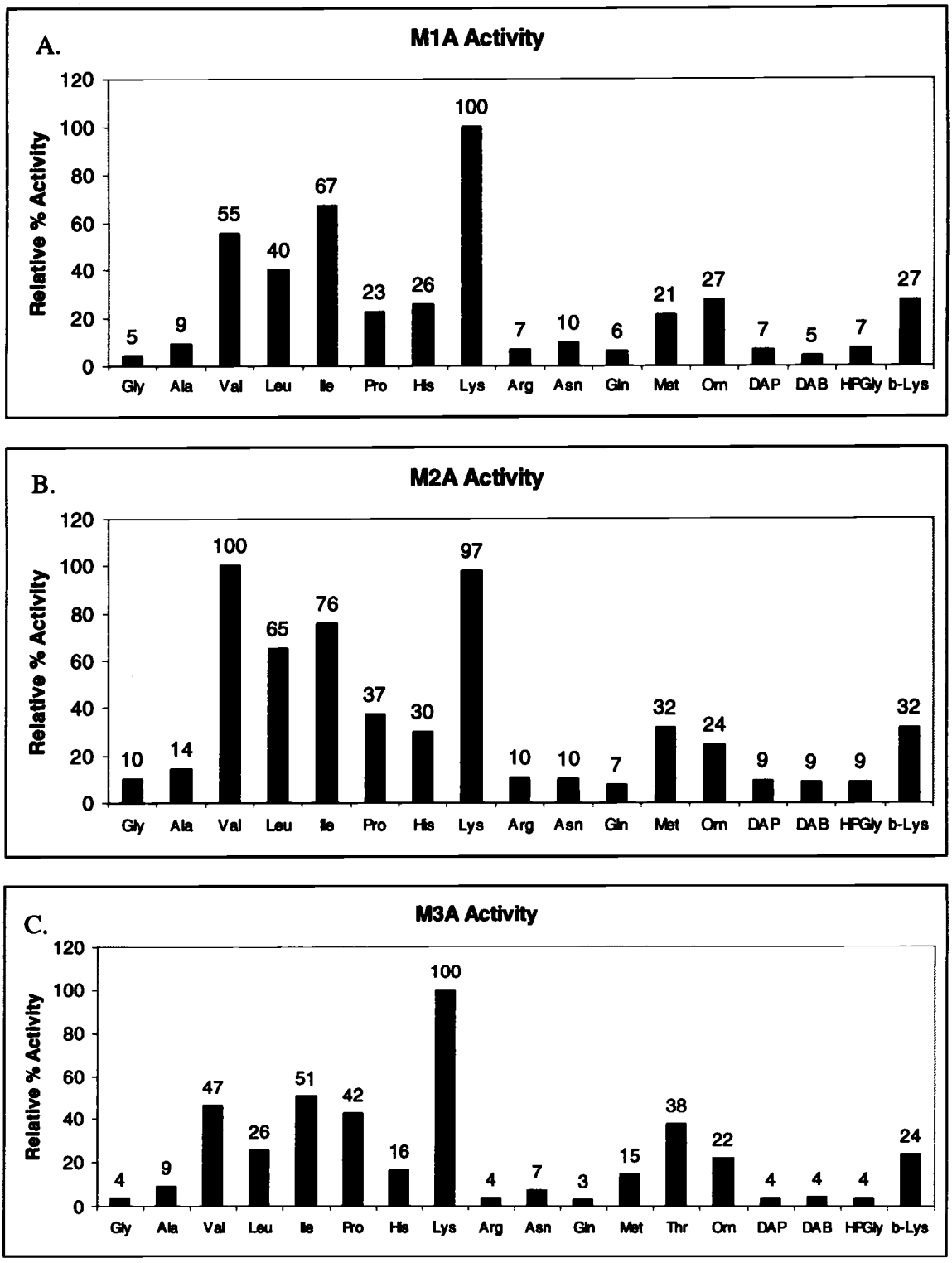


Figure 3.7. Relative activity of A•PCP-domains with individual amino acids. A. Activity of the A- domain from the first NRPS module (M1A). B. Activity of the A-domain from the second NRPS module (M2A). C. Activity of the A-domain from the third NRPS module (M3A). DAP is 2,3-diaminoproprionic acid, DAB is 3,4-diaminobutyric acid, HPGly is hydroxyphenylglycine

range as the control ProA2 and TyrA7 domains with noncognate amino acids. This suggests that the true substrates for the *Jaspis* PKS/NRPS A-domains have not been identified.

Although the true substrate amino acids for these A-domains could not be predicted, the overall trends in substrate selectivity can provide insight into the type of amino acid that may be activated. In all three cases, the trend was for the activation of linear or branched chain amino acids. It is interesting that the *Jaspis* A-domains also activate lysine, as it would not be expected that a binding pocket that promotes binding of a nonpolar aliphatic amino acid would provide favorable binding interactions for the basic lysine side chain.

Although general trends as to the nature of amino acids that are activated by these domains can be determined through ATP-$^{32}PP_i$ exchange assays, the substrate specificity of an *in vitro* assay does not always correspond to the *in vivo* findings. This scenario was observed by Chang *et al.* in amino acid selectivity studies with A-domains of the barbamide gene cluster. Although barbamide A_D was shown to equally activate both valine and leucine *in vitro*, only leucine is incorporated *in vivo*.[4] These results demonstrate the limitations of characterizing a single catalytic domain that has been removed from the native protein and environment. This also emphasizes the importance of *in vivo* characterization of enzyme and enzyme systems even in this genomic era.

The biochemical characterization of the *Jaspis* cluster A-domains does support the conclusion that it does not encode jaspamide biosynthesis. A-domains for the jaspamide cluster would be expected to activate tyrosine (or β-tyrosine), tryptophan, and alanine. Even if the *in vitro* assays were not entirely representative of the intact enzyme, the trend for the first two domains to activate aromatic amino acids would still be expected. Analysis of the substrate selectivity data supports the information obtained from analysis of the gene cluster and clearly shows that this mixed pathway is unlikely to be involved in the biosynthesis of jaspamide.

Localization of the Mixed PKS/NRPS Gene Cluster

The involvement of bacteria associated with marine invertebrates in the biosynthesis of many marine natural products has been frequently discussed in recent literature.[27-30] In an effort to answer this question as it pertains to this work, attempts were made to try and determine the origin of the cloned biosynthetic gene cluster. The successful identification of the source organism would not only be of academic interest, but it could also provide the opportunity to identify the metabolite that is produced by the *Jaspis* PKS/NRPS gene cluster.

Microbiologists at Wyeth Research, as well as in our own laboratory, have been able to successfully culture a number of bacterial species from Fijian *J. splendens* samples. The successful culture of these microbes provided us with the opportunity to identify the source organism, as well as potentially associate a secondary metabolite for this 'orphan' gene cluster. PCR primers were designed to amplify the region of the *Jaspis* gene cluster between the PKS and NRPS genes (Figure 3.8). In order to reduce the likelihood of amplifying a non-related PKS or NRPS gene clusters, PCR primers were specifically designed off of gene regions that are not highly conserved within PKS and NRPS genes. These primers were then used to screen genomic DNA isolated from the cultured microbes by PCR. Interestingly, PCR products were amplified from two different Actinomycetes that were cultured from the *J. splendens* sponge, a *Rhodococcus sp.* and a *Salinospora sp.* Sequencing confirmed that the amplified products corresponded to the *J. splendens* mixed PKS/NRPS gene cluster. Although this was not expected, it is not an unprecedented phenomenon. Many microorganisms, particularly within the Actinomycetes family, have been found to produce identical or related metabolites. Blasticidin S, for example, has been isolated from 3 different *Streptomyces* species.[31-33] Similarly, the streptothricin F gene cluster has been identified and cloned from two different *Streptomyces* species.[34-36] Horizontal gene transfer is believed to be responsible for this observation and is a common feature of many bacterial species.[37] Many researchers take advantage of this bacterial property to introduce DNA into bacterial species that are otherwise difficult to transform. [38,39]

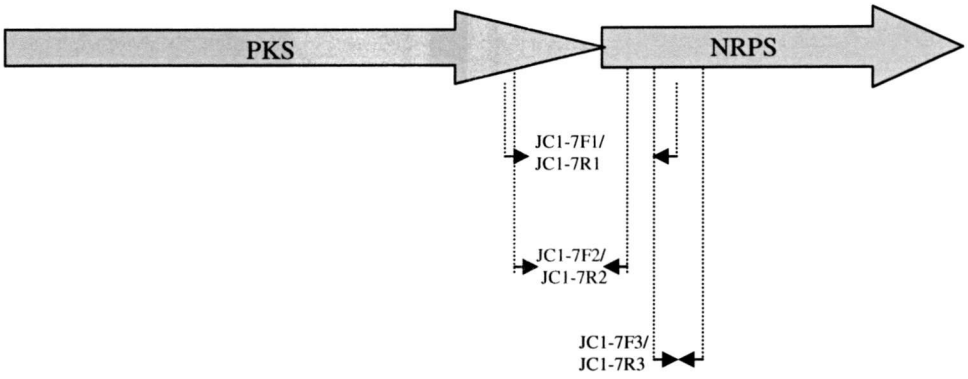


Figure 3.8. Region of the *Jaspis* mixed PKS/NRPS gene cluster amplified by specific PCR primers.

Given the ability of microorganisms to transfer genetic material, it is not surprising that two closely related bacteria, both Actinomycetes, isolated from the same environment possess the same biosynthetic gene cluster. One interesting feature, however, is that this gene cluster was identified in sponges that were collected at different locations and in different years. This suggests that either the bacteria, or the gene cluster, are widely dispersed within the South Pacific, or that this bacterium commonly forms associations with *Jaspis* sponges.

An additional observation is that this specific gene cluster was cloned twice from *Jaspis* metagenomic libraries and, in each instance, it was the only mixed PKS/NRPS gene cluster to be identified. Several explanations may account for this observation. First, this may simply be a result of the observed stability differences between the genomic DNA of various marine organisms. Initial studies on the purification of genomic DNA from marine sources found that the gDNA from some organisms appeared to be more fragile than others. This 'fragile' gDNA was more likely to degrade during isolation and would result in a lower yield of high molecular weight DNA. A similar observation was made by A. Mintie in our laboratory who was involved in the culturing of microorganisms from marine sources, particularly the *Jaspis* sponge. It was found that purification of gDNA from several isolated bacterial strains was much more difficult to isolate than others (A. Mintie, personal communication). The type of degradation observed was consistent with gDNA preparation from some intact marine organisms. The differential degradation of genomic DNA from various bacteria within a sponge

assemblage may unintentionally result in the formation of genomic libraries that are biased towards organisms with stable gDNA. If the *Rhodococcus* and *Salinospora* gDNA is more resilient than that of other *Jaspis* associated bacteria, generation of the genomic library could result in the better representation of these species than other bacteria. Consequently, the opportunities to identify and clone biosynthetic gene clusters from these organisms would be enhanced.

Significance

The results of the substrate specificity assays supported the results of sequence analysis and confirm that the identified mixed PKS/NRPS gene cluster that has been isolated from the marine sponge, *J. splendens*, is likely not involved in the biosynthesis of jaspamide. Although the specific aim of this project, cloning of the jaspamide biosynthetic gene cluster, was not accomplished, the overall goal to develop a method for cloning biosynthetic gene clusters from symbiotic marine organisms was successful.

One factor that likely contributed to our inability to identify the jaspamide gene cluster was the biological complexity of the *J. splendens* tissue sample. At the beginning of this work *J. splendens* was reported to be a simple symbiotic system that contained only a small number of associated microbes (M.K. Harper, personal communication with M. Zabriskie). This was one of the features that made the *J. splendens* system attractive for our initial work in this field. The limited number of microbes associated with the sponge sample reduces the number of genomes that must be sorted through in search for the desired gene cluster. As the number of associated microbes increases so does the size of the genomic library that must be generated to represent all of the organisms. In a system where only two or three genomes need to be considered a smaller library, such as the approximately 20,000 colonies screened in this work, provided adequate genome coverage.

More recently, work by researchers at Wyeth Reserch (Dr.Valerie Bernan and Jeff Janso), as well as by Ann Mintie working in our laboratory, suggested that *J. splendens* may not have a simple symbiosis after all. The Wyeth group was able to culture 62 bacterial species including 3 *Actinomycetes* and 2 fungi from the same *J.*

splendens sample used in our work. In addition, 16S ribosomal DNA analysis identified a number of other bacteria present in the sample that could not be cultured. These results clearly show that the incredible microbial diversity that is present and reveal that genomic library used in this work was not large enough to represent even single coverage of the bacterial genomes contained within the sample. The metagenomic size of the cultured actinomycetes alone represents an estimated 24 Mb, based on an average actinomycetes genome of 8 Mb.[40-42] When the genomes of the associated fungi, eubacteria, uncultured bacteria and sponge are added to this number it becomes apparent that our genomic library, of approximately 700-800 Mb, was not large enough to statistically represent the metagenome of the *Jaspis* assemblage.

Although this approach has not been successful, the search for the jaspamide biosynthetic gene cluster has not been abandoned. A modified approach based on the similarity of jaspamide to other natural products is being applied. One conclusion from this work is that the practical identification of the jaspamide gene cluster from such a complex assemblage of microorganisms will require the development of more specific gene probes. The similarity of jaspamide to natural products isolated from terrestrial *Chondramyces* may make this possible. The degree of similarity between jaspamide and chondramide (Figure 1.2) suggests that the biosynthetic gene clusters are also similar in architecture. An alternate approach to the identification of the jaspamide gene cluster is to first identify the chondramide biosynthetic gene cluster and then use that cluster to generate specific gene probes. As the chondramide producing bacteria, *Chondromyeces crocatus*, can be cultured in the laboratory, issues such as excessive genome size do not need to be addressed. Recently the chondramide biosynthetic gene cluster was cloned from the chondramide producing bacteria (R. Muller, personal communication to M. Zabriskie). In collaboration with Dr. Muller, our laboratory is planning on using sequence information from the chondramide gene cluster to reprobe the *J. splendens* genomic library in an attempt to identify the jaspamide gene cluster. In addition to guiding the continued search for the jaspamide biosynthetic gene cluster, the molecular techniques developed in this work can not only be applied to the *Jaspis* system, but also to other marine organisms in the search for biosynthesis gene clusters.

Materials and Methods

General

Routine molecular biology procedures including DNA manipulation, growth and maintenance of *E. coli*, transformations, competent cell preparation and SDS-PAGE analysis were conducted according to standard techniques.[43] Restriction enzymes, T4 DNA ligase and *taq* polymerase were purchased from various suppliers and used according to manufacturer's protocols. QIAprep® spin miniprep and QIAquick® gel extraction kits (Qiagen, Valencia, CA) were used for DNA purification. Sequencing of vector inserts was conducted by the Center for Gene Research and Biotechnology, Oregon State University. Plasmids containing control A-domain constructs, ProA2 and TyrA7, were gifts from Prof. M. Marahiel,[9] pUC8sfp was a gift from P. Zuber.[19,44]

Construction of A-domain Expression Vectors

The three adenylation domains of the *J. splendens* mixed PKS/NRPS gene cluster were amplified by PCR for cloning into pET series expression vectors (Novagen, Madison, WI). The sequences of PCR primers used to amplify the A-domains are shown in Table 3.2; specific primer pairs used for cloning into each of the vectors are described below. PCR was carried out in a total volume of 50 μL containing 10 ng JC7 template, 1X Promega Thermophilic Buffer, 1.25 mM $MgCl_2$, 0.4 mM dNTP mix (MBI), 5% dimethylsulfoxide, 50 pmol forward primer, 50 pmol reverse primer, and 5 units Platinum Taq DNA polymerase (Invitrogen, Carlsbad, CA). Gel purified PCR products were digested with appropriate restriction enzymes then ligated into a similarly restricted vector (pET20b+, pET28a+ or pET41a+). The resulting plasmids were used to transform *E.coli* JM109 (Promega, Madison, WI) and the correct construction was confirmed by sequencing. Expression constructs containing just the A-domain were then used to transform *E.coli* Rosetta (DE3)pLysS competent cells (Novagen) for expression. Expression constructs carrying both the A-and PCP-domains were used to transformed *E.coli* Rosetta (DE3)pLysS/pUC8sfp.

Table 3.2. PCR primers used for the amplification of A-domains. Introduced restriction sites are underlined.

Primer	Sequence	Rest. site
M1A*NdeI*For-	GGAATTCCATATGGAGCAACGCGTTTCGCAACTTCGA	*Nde*I
M1ST*EcoR*For-	CTCGAGGGAATTCGATGAACTGGGTTGAATCT	*EcoR*I
M1A*EcoR*For-	CTCGAGGGAATTCGGAGCAACGCGTTTCGCAA	*EcoR*I
M1ARev-	TATTTCTCGAGCCAGATGGCGGCGAGGCTTGC	*Xho*I
M1ATRev-	TATTCTCGAGTTACGCTGCCGCCTGCACCGT	*Xho*I
M2AFor-	GGAATTCCATATGACCCGAAAGCTATCGGAACTG	*Nde*I
M2A*Bam*For-	GAATTCGGATCCGACCCGAAAGCTATCGGAA	*Bam*HI
M2ATRev-	TATTCTCGAGTAGACTGTGAGCCGTATCGAG	*Xho*I
M2ATRev-	TATTCTCGAGTTATAGACTGTGAGCCGTATC	*Xho*I
M3Ab*EcoR*For-	CTCGAGGGAATTCGCCGACGCAGCCGGTC	*EcoR*I
M3Ab*Nde*For-	GAATTCCATATGCCGACGCAGCCGGTC	*Nde*I
M3ARev-	TATTCTCGAGTTCGACGTGTAAGACCTCGCGCCA	*Xho*I
M3ATRev-	TATTCTCGAGTTAGCGTGCGGCTCTACGAGT	*Xho*I

Primer pairs of M1AFor/ M1ATRev, M1AFor/ M1ARev, M2AFor/ M2ATRev, M2AFor/ M2ARev, M1STFor/ M1ARev and M1STFor/ M1ATRev were used for cloning into similarly digested pET20b+ for expression with a *C*-terminal His_6-tag. Primer pairs of M1AEcoFor/ M1ATRev, M1AEcoFor/ M1ARev, M2ABamFor/ M2ATRev, M2ABamFor/ M2ARev, M3AbEcoFor/ M3ARev and M3AbEcoRIFor/ M3ARev were used for cloning into similarly digested pET20b+ for expression with the pelB leader and *C*-terminal His_6-tag. Primer pairs of M1AFor/ M1ATRev, M1AFor/ M1ARev, M2AFor/ M2ATRev, M2AFor/ M2ARev, M1STFor/ M1ARev and M1STFor/ M1ATRev were used for cloning into similarly digested pET28a+ for expression with an *N*-terminal His_6-tag. Primer pairs of M1AEcoFor/ M1ATRev, M1AEcoFor/ M1ARev, M2ABamFor/ M2ATRev, M2ABamFor/ M2ARev, M3AbEcoFor/ M3ARev and M3AbEcoRIFor/ M3ARev were used for cloning into similarly digested pET41b+ for expression with an *N*-terminal GST-fusion and a *C*-terminal His_6-tag.

Optimization of A-domain Expression

Seed cultures of *E. coli* Rosetta (DE3) pLysS containing the expression construct were grown overnight grown in Terrific Broth containing 50 μg/mL chloramphenicol for

maintenance of pLysS, and either 100 μg/mL ampicillin (pET20b+) or 50 μg/mL kanamycin (pET28a+, pET41b+). Overnight seed culture was used to inoculate 5 mL of Terrific broth supplemented with the appropriate antibiotics. The cells were grown to an $OD_{600} \cong 0.5$ and then protein expression was induced with isopropyl-β-D-thiogalactopyranoside (IPTG) at a final concentration of 0-1.0 mM. The cultures were then grown for an additional 3-40 h at 18-37 °C. Cells were harvested by centrifugation, resuspended in 1 mL 4 °C Talon Wash Buffer (TWB: 50 mM NaH_2PO_4, 300 mM NaCl, pH 7.5) and lysed by sonication in a Microson ultrasonic cell disruptor (four 15 s bursts at 6 watts output with 30 s cooling on ice between bursts). Cellular debris was pelleted by centrifugation at 18000 *x g* for 30 min at 4 °C. SDS-PAGE analysis and Western detection of His_6-tagged proteins (His•Tag® Western detection kit, Novagen) was used to analyze A-domain expression.

Growth and Expression of A•PCP-domains

Seed cultures of *E. coli* Rosetta (DE3) pLysS/pUC8*sfp* containing the expression construct (pM1AT28, pM2AT28, pM3AT28, ProA2, or TyrA7) were grown overnight in Terrific Broth containing 50 μg/mL kanamycin for maintenance of the pET28a construct, 50 μg/mL chloramphenicol for maintenance of pLysS, and 100 μg/mL ampicillin for maintenance of pUC8sfp.

Overnight seed culture was used to inoculate 1 L of Terrific broth, supplemented with the appropriate antibiotics, and grown for 40 h at 20 °C. Cells were harvested by centrifugation and resuspended in 30 mL 4 °C Talon Wash Buffer (TWB: 50 mM NaH_2PO_4, 300 mM NaCl, pH 7.5) supplemented with 0.2 μg/mL leupeptin, 1 μg/mL pepstatin A, and 2μg/mL phenylmethylsulfonyl fluoride (PMSF). Cells were lysed by sonication in a Microson ultrasonic cell disruptor (twelve 1 min bursts at 6 watts output with 1 min cooling on ice between bursts). Cellular debris was pelleted by centrifugation at 18000 *x g* for 40 min at 4 °C. Soluble A•PCP-domains were purified from the cleared lysate cobalt affinity chromatography (TALONspin, BD Bioscience) according to manufacturer's protocols. Following protein binding, the resin was washed with increasing concentrations of imidazole in TWB (3 x with 5 mM imidazole; 2 x with 10

mM imidazole; 1 x with each 15 mM imidazole, 20 mM imidazole, 40 mM imidazole and 250 mM imidazole.

SDS-PAGE analysis was use to identify fractions that contained A•PCP-domains. These fractions were pooled and dialyzed overnight against two changes of dialysis buffer (50 mM Tris pH 7.5, 100 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT, 1 mM EDTA, 10 % glycerol) at 4 °C (2 L for 2 h then 4 L overnight). Finally, protein was concentrated to ~0.3 μg/μL with a Centricon® Plus-20 concentrator (Millipore, Bedford, MA) and stored at 4 °C. Protein concentration was determined by Bradford analysis against BSA standards.[45]

ATP-$^{32}PP_i$ Exchange Assays

Radiolabled [^{32}P] pyrophosphate was purchased from New England Nuclear (PerkinElmer Life and Analytical Science, Boston, MA). β-Tyrosine was purchased from PepTech Corp (Burlington, MA) and 4-hydroxy-L-phenylglycine from was purchased from Fluka (Milwaukee, WI). β-Lysine was prepared by Professor M. Zabriskie. The remaining amino acids, tetrasodium pyrophosphate, magnesium chloride, ATP and activated charcoal were purchased from Sigma-Aldrich (St.Louis, MO).

ATP-32PP_i exchange assays were conducted as described by Luo *et al.*.[46] A final reaction volume of 100 μL assay buffer containing 50 mM Tris (pH 7.5), 10 mM $MgCl_2$, 4 mM ATP,0 or 2 mM amino acid and 14 μg enzyme was initiated with the addition of 1 mM-0.3 μCi (16.06 Ci/mmol) tetrasodium [^{32}P]pyrophosphate and incubated at 37 °C for one h. Reactions are quenched by adding 0.5 ml quench solution (1.6% w/v activated charcoal, 4.46% w/v tetrasodium pyrophosphate, and 3.5% perchloric acid in water). The charcoal-bound radioactivity was pelleted by centrifugation, washed twice in 1 mL wash solution (4.46% w/v tetrasodium pyrophosphate, and 3.5% perchloric acid in water), and resuspended in 0.5 mL sterile water. Resuspended charcoal was added to 0.5 mL scintillation fluid (ScintiSafe Gel, Fisher) and charcoal-bound radioactivity was determined by liquid scintillation counting (LSC) on a Beckman LS 6500.

Design of Specific Jaspis PKS/NRPS PCR Primers.

Specific PCR primers used to amplify the *Jaspis* mixed PKS/NRPS gene cluster from *Rhodococcus* and *Salinospora* were as follows: JC1-7F (GCAAAACCAAGCCAAAGCCG)/ JC1-7R (TAGCCGCTGGTGGGTCAACA) amplifies a 3172 bp fragment of the PKS/NRPS junction just prior to the first conserved KR motif to the region between the conserved NRPS adenylation motifs A2 and A3. A second primer pair, JC1-7F2 (ATTGTGCATGCCGCAGGGGTT) / JC1-7R2 (TTGGCAAATCGAGCACGGGTA), amplifies a 1711 bp fragment of the PKS/NRPS junction between the second conserved ketoreductase motif and the region between C4 and C5 of the first NRPS condensation domain. PCR reactions were conducted by A. Mintie.

References

(1) Stachelhaus, T.; Marahiel, M. A. Modular structure of peptide synthetases revealed by dissection of the multifunctional enzyme GrsA. *J. Biol. Chem.* **1995**, *270*, 6163-6169.

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

(3) Becker, J. E.; Moore, R. E.; Moore, B. S. Cloning, sequencing, and biochemical characterization of the nostocyclopeptide biosynthetic gene cluster: molecular basis for imine macrocyclization. *Gene* **2004**, *325*, 35-42.

(4) Chang, Z.; Flatt, P.; Gerwick, W. H.; Nguyen, V. A.; Willis, C. L.; Sherman, D. H. The barbamide biosynthetic gene cluster: a novel marine cyanobacterial system of mixed polyketide synthase (PKS)-non-ribosomal peptide synthetase (NRPS) origin involving an unusual trichloroleucyl starter unit. *Gene* **2002**, *296*, 235-247.

(5) Keating, T. A.; Suo, Z.; Ehmann, D. E.; Walsh, C. T. Selectivity of the yersiniabactin synthetase adenylation domain in the two-step process of amino acid activation and transfer to a holo-carrier protein domain. *Biochemistry* **2000**, *39*, 2297-2306.

(6) Recktenwald, J.; Shawky, R.; Puk, O.; Pfennig, F.; Keller, U.; Wohlleben, W.; Pelzer, S. Nonribosomal biosynthesis of vancomycin-type antibiotics: a heptapeptide backbone and eight peptide synthetase modules. *Microbiology* **2002**, *148*, 1105-1118.

(7) de Crecy-Lagard, V.; Blanc, V.; Gil, P.; Naudin, L.; Lorenzon, S.; Famechon, A.; Bamas-Jacques, N.; Crouzet, J.; Thibaut, D. Pristinamycin I biosynthesis in

Streptomyces pristinaespiralis: molecular characterization of the first two structural peptide synthetase genes. *J. Bacteriol.* **1997**, *179*, 705-713.

(8) Duitman, E. H.; Hamoen, L. W.; Rembold, M.; Venema, G.; Seitz, H.; Saenger, W.; Bernhard, F.; Reinhardt, R.; Schmidt, M.; Ullrich, C.; Stein, T.; Leenders, F.; Vater, J. The mycosubtilin synthetase of *Bacillus subtilis* ATCC6633: a multifunctional hybrid between a peptide synthetase, an amino transferase, and a fatty acid synthase. *Proc. Natl. Acad. Sci. U S A* **1999**, *96*, 13294-13299.

(9) Mootz, H. D.; Marahiel, M. A. The tyrocidine biosynthesis operon of Bacillus brevis: complete nucleotide sequence and biochemical characterization of functional internal adenylation domains. *J. Bacteriol.* **1997**, *179*, 6843-6850.

(10) Duerfahrt, T.; Doekel, S.; Sonke, T.; Quaedflieg, P. J.; Marahiel, M. A. Construction of hybrid peptide synthetases for the production of alpha-l-aspartyl-l-phenylalanine, a precursor for the high-intensity sweetener aspartame. *Eur. J. Biochem.* **2003**, *270*, 4555-4563.

(11) Doekel, S.; Marahiel, M. A. Dipeptide formation on engineered hybrid peptide synthetases. *Chem. Biol.* **2000**, *7*, 373-384.

(12) Du, L.; Sanchez, C.; Chen, M.; Edwards, D. J.; Shen, B. The biosynthetic gene cluster for the antitumor drug bleomycin from *Streptomyces verticillus* ATCC15003 supporting functional interactions between nonribosomal peptide synthetases and a polyketide synthase. *Chem. Biol.* **2000**, *7*, 623-642.

(13) Kane, J. F. Effects of rare codon clusters on high-level expression of heterologous proteins in Escherichia coli. *Curr. Opin. Biotechnol.* **1995**, *6*, 494-500.

(14) Nakamura, Y.; Gojobori, T.; Ikemura, T. Codon usage tabulated from international DNA sequence databases: status for the year 2000. *Nucleic Acids Res.* **2000**, *28*, 292.

(15) Keen, N. T.; Tamaki, S. Structure of two pectate lyase genes from *Erwinia chrysanthemi* EC16 and their high-level expression in *Escherichia coli*. *J. Bacteriol.* **1986**, *168*, 595-606.

(16) Matthey, B.; Engert, A.; Klimka, A.; Diehl, V.; Barth, S. A new series of pET-derived vectors for high efficiency expression of *Pseudomonas* exotoxin-based fusion proteins. *Gene* **1999**, *229*, 145-153.

(17) Uhlen, M.; Moks, T. Gene fusions for purpose of expression: an introduction. *Methods Enzymol.* **1990**, *185*, 129-143.

(18) Smyth, D. R.; Mrozkiewicz, M. K.; McGrath, W. J.; Listwan, P.; Kobe, B. Crystal structures of fusion proteins with large-affinity tags. *Protein Sci.* **2003**, *12*, 1313-1322.

(19) Weinreb, P. H.; Quadri, L. E.; Walsh, C. T.; Zuber, P. Stoichiometry and specificity of in vitro phosphopantetheinylation and aminoacylation of the valine-activating module of surfactin synthetase. *Biochemistry* **1998**, *37*, 1575-1584.

(20) Walsh, C. T.; Gehring, A. M.; Weinreb, P. H.; Quadri, L. E.; Flugel, R. S. Post-translational modification of polyketide and nonribosomal peptide synthases. *Curr. Opin. Chem. Biol.* **1997**, *1*, 309-315.

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

(22) Kealey, J. T.; Liu, L.; Santi, D. V.; Betlach, M. C.; Barr, P. J. Production of a polyketide natural product in nonpolyketide-producing prokaryotic and eukaryotic hosts. *Proc. Natl. Acad. Sci. U S A* **1998**, *95*, 505-509.

(23) Lee, S. G.; Lipmann, F. Tyrocidine synthetase system. *Methods Enzymol.* **1975**, *43*, 585-602.

(24) Belshaw, P. J.; Walsh, C. T.; Stachelhaus, T. Aminoacyl-CoAs as probes of condensation domain selectivity in nonribosomal peptide synthesis. *Science* **1999**, *284*, 486-489.

(25) Clugston, S. L.; Sieber, S. A.; Marahiel, M. A.; Walsh, C. T. Chirality of peptide bond-forming condensation domains in nonribosomal peptide synthetases: the C5 domain of tyrocidine synthetase is a (D)C(L) catalyst. *Biochemistry* **2003**, *42*, 12095-12104.

(26) Cheng, Y. Q.; Tang, G. L.; Shen, B. Identification and localization of the gene cluster encoding biosynthesis of the antitumor macrolactam leinamycin in *Streptomyces atroolivaceus* S-140. *J. Bacteriol.* **2002**, *184*, 7013-7024.

(27) Engel, S.; Jensen, P. R.; Fenical, W. Chemical ecology of marine microbial defense. *J. Chem. Ecol.* **2002**, *28*, 1971-1985.

(28) Haygood, M. G.; Schmidt, E. W.; Davidson, S. K.; Faulkner, D. J. Microbial symbionts of marine invertebrates: opportunities for microbial biotechnology. *J. Mol. Microbiol. Biotechnol.* **1999**, *1*, 33-43.

(29) Proksch, P.; Edrada, R. A.; Ebel, R. Drugs from the seas - current status and microbiological implications. *Appl. Microbiol. Biotechnol.* **2002**, *59*, 125-134.

(30) Salomon, C. E.; Magarvey, N. A.; Sherman, D. H. Merging the potential of microbial genetics with biological and chemical diversity: an even brighter future for marine natural product drug discovery. *Nat. Prod. Rep.* **2004**, *21*, 105-121.

(31) Sugiyama, M.; Takeda, A.; Paik, S. Y.; Nimi, O.; Nomi, R. Acetylation of blasticidin S by its producing actinomycetes. *J. Antibiot. (Tokyo)* **1986**, *39*, 827-832.

(32) Larsen, S. H.; Berry, D. M.; Paschal, J. W.; Gilliam, J. M. 5-Hydroxymethylblasticidin S and blasticidin S from *Streptomyces setonii* culture A83094. *J. Antibiot. (Tokyo)* **1989**, *42*, 470-471.

(33) Takeuchi, S.; Hirayama, K.; Ueda, K.; Sakai, H.; Yonehara, H. Blasticidin S, a new antibiotic. *J. Antibiot. (Tokyo)* **1958**, *11*, 1-5.

(34) Thiruvengadam, T. K.; Gould, S. J.; Aberhart, D. J.; Lin, H. Biosynthesis of Streptothricin F. 5. Formation of β-Lysine by *Streptomyces* L-1689-23. *J. Am. Chem. Soc.* **1983**, *105*, 5470-5476.

(35) Jackson, M. D.; Gould, S. J.; Zabriskie, T. M. Studies on the formation and incorporation of streptolidine in the biosynthesis of the peptidyl nucleoside antibiotic streptothricin F. *J. Org. Chem.* **2002**, *67*, 2934-2941.

(36) Fernandez-Moreno, M. A.; Vallin, C.; Malpartida, F. Streptothricin biosynthesis is catalyzed by enzymes related to nonribosomal peptide bond formation. *J. Bacteriol.* **1997**, *179*, 6929-6936.

(37) Lopez, J. V. Naturally mosaic operons for secondary metabolite biosynthesis: variability and putative horizontal transfer of discrete catalytic domains of the epothilone polyketide synthase locus. *Mol. Genet. Genomics* **2003**, *270*, 420-431.

(38) Flett, F.; Mersinias, V.; Smith, C. P. High efficiency intergeneric conjugal transfer of plasmid DNA from *Escherichia coli* to methyl DNA-restricting streptomycetes. *FEMS Microbiol. Lett.* **1997**, *155*, 223-229.

(39) Mazodier, P.; Petter, R.; Thompson, C. Intergeneric conjugation between *Escherichia coli* and *Streptomyces* species. *J. Bacteriol.* **1989**, *171*, 3583-3585.

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

(41) Omura, S.; Ikeda, H.; Ishikawa, J.; Hanamoto, A.; Takahashi, C.; Shinose, M.; Takahashi, Y.; Horikawa, H.; Nakazawa, H.; Osonoe, T.; Kikuchi, H.; Shiba, T.; Sakaki, Y.; Hattori, M. Genome sequence of an industrial microorganism *Streptomyces avermitilis*: deducing the ability of producing secondary metabolites. *Proc. Natl. Acad. Sci. U S A* **2001**, *98*, 12215-12220.

(42) Bamas-Jacques, N.; Lorenzon, S.; Lacroix, P.; De Swetschin, C.; Crouzet, J. Cluster organization of the genes of *Streptomyces pristinaespiralis* involved in pristinamycin biosynthesis and resistance elucidated by pulsed-field gel electrophoresis. *J. Appl. Microbiol.* **1999**, *87*, 939-948.

(43) Sambrook, J.; Russell, D. *Molecular Cloning A Laboratory Manual*; 3 ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, 2001.

(44) Nakano, M. M.; Corbell, N.; Besson, J.; Zuber, P. Isolation and characterization of sfp: a gene that functions in the production of the lipopeptide biosurfactant, surfactin, in *Bacillus subtilis*. *Mol. Gen. Genet.* **1992**, *232*, 313-321.

(45) Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **1976**, *72*, 248-254.

(46) Luo, L.; Burkart, M. D.; Stachelhaus, T.; Walsh, C. T. Substrate recognition and selection by the initiation module PheATE of gramicidin S synthetase. *J. Am. Chem. Soc.* **2001**, *123*, 11208-11218.

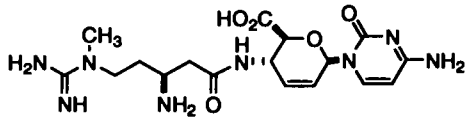
Chapter 4

General Introduction to Blasticidin S Biosynthesis

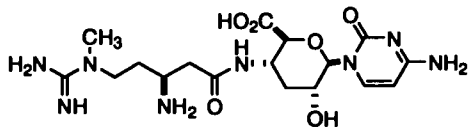
Introduction

Blasticidin S (BS) is a *Streptomyces griseochromogenes* metabolite first isolated in 1958[1] by Takeuchi while screening a number of *Streptomyces* for antifungal activity against *Piricularia oryzae.* In 1961 BS became commercially available in Japan for use as an agricultural fungicide to replace mercury-based antifungal compounds.[2] BS was used on a large scale for the prevention of rice blast caused by the fungus *P. oryzae*, a major rice disease in much of Asia and South America.[3,4] BS is a broad spectrum antibiotic with toxicity to prokarya, archaea and eukarya.[5] In addition to its antifungal activity, BS has been found to possess antiviral,[6,7] antibacterial,[8] and antitumor[9] properties. The use of blasticidin S as an agricultural fungicide has diminished in recent years,[2] however new applications for BS have been developed based on the identification and cloning of several BS resistance genes.

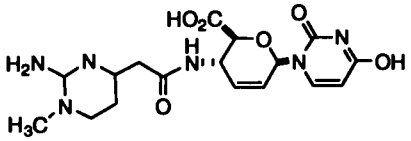
In addition to blasticidin S, several structurally related compounds have been isolated from *S. griseochromogenes* and include blasticidin H (**4.2**),[10] cytomycin (**4.3**),[9] and leucylblasticidin S (**4.4**).[11] Since its original isolation from *S. griseochromogenes*, BS has been isolated from other *Streptomyces* species including *S. setonii* and *S. morookaensis* (*Streptoverticillium sp.* 4673).[12,13]



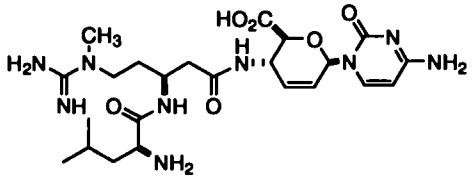
Blasticidin S
4.1



Blasticidin H
4.2

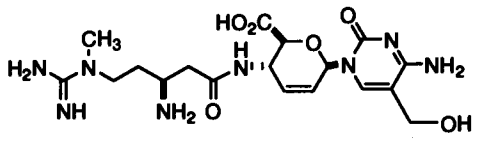


Cytomycin
4.3

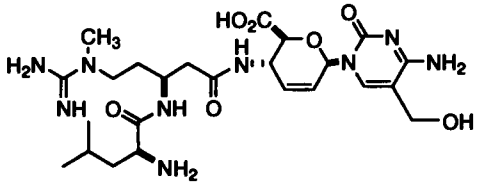


Leucylblasticidin S
4.4

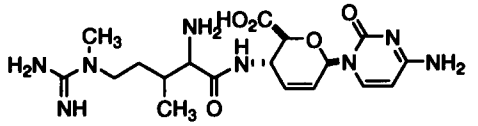
A number of structurally related natural products have also been isolated from other bacterial species and include 5-hydroxymethyl blasticidin S from *S. setonii* (**4.5**),[13] 5-hydroxymethyl leucyblasticidin S (Sch 36605) (**4.6**) from *S. sp.* Sch 36605, [14,15] arginomycin (**4.7**) from *S. arginesis,*[16] and mildiomycin (**4.8**) from *Streptoverticillium rimofaciens* B-98891.[17]



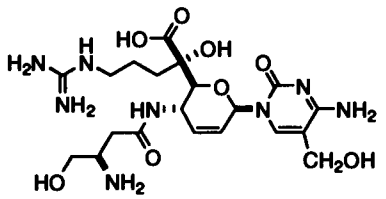
5-Hydroxymethyl blasticidin S
4.5



5-Hydroxymethyl leucylblasticidin S (Sch36605)
4.6



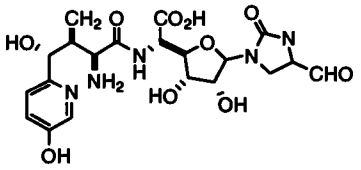
Arginomycin
4.7



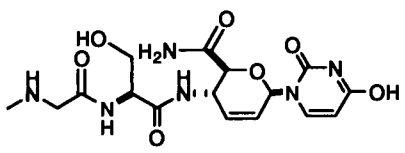
Mildiomycin
4.8

The structure of blasticidin S was determined through a series of chemical degradation and NMR studies[18-20] and confirmed through x-ray diffraction studies.[21] The structure of BS was found to be composed of two major moieties, blastidic acid and an aminosugar nucleoside (cytosinine).[20] BS is a member of the peptidyl nucleoside family of compounds. Peptidyl nucleoside antibiotics are a family of diverse compounds that generally contain a heterocyclic base, amino sugar and an amino acid or peptidyl moiety.[22] Members of this family have diversity in both structure and biological activity. Representative peptidyl nucleoside antibiotics include nikkomycin X,[23] puromycin, gougerotin, and streptothricin F.[22,24-26]

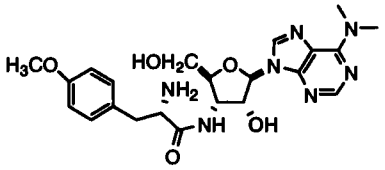
Nikkomycin X (**4.9**) is a *Streptomyces tendae* Tu901 metabolite with antifungal and insecticidal activity derived from its ability to inhibit chitin synthesis.[27] Gougerotin (**4.10**) is a broad spectrum antibiotic isolated from *S. gougerotii* that acts through the inhibition of protein synthesis.[22] Puromycin (**4.11**) has been isolated from *S. alboniger* and is another compound with the ability to block protein synthesis,[28]and streptothricin F (**4.12**) is a broad spectrum antibiotic isolated from *S. lavendulae.*[29]



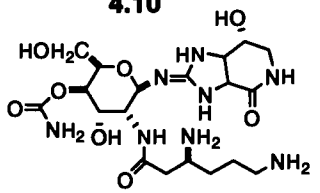
Nikkomycin X
4.9



Gougerotin
4.10



Puromycin
4.11



Streptothricin F
4.12

Blasticidin S Mechanism of Action

In addition to being used as a fungicide, blasticidin S has found several other applications within the scientific community. BS has been used as both a selectable marker system for transformed cells and in the study of ribosomal function.

Shortly after its discovery, it was found that the activity of blasticidin S arises from its ability to block protein synthesis. This activity has resulted in the use of BS to study ribosome function.[5,30-33] These studies have had the added advantage of providing insight into the BS mode of action. Early studies showed that the potent antifungal activity of blasticidin S came from its ability to block protein synthesis by binding to the large ribosomal subunit and inhibiting formation of peptide bonds.[32-34] rRNA footprinting assays later revealed that blasticidin S had a simple rRNA footprint and was proposed to function by simply providing a steric blockage to ribosome translocation.[5]

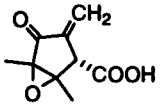
Further insight into the exact mechanism of BS activity came when Hansen *et al.* were able to crystallize the large ribosomal subunit from *Haloarcula marismortui* complexed with a variety of antibiotics.[31] They found that blasticidin S was unique among the antibiotics studied in that it was the only one to bind to the rRNA P-loop rather than one of two hydrophobic crevices. BS acts as a tRNA mimic and blocks tRNA binding to the ribosome. Key interactions in the binding of BS include base pairing of the BS cytosinine moiety with one of two rRNA residues that are critical for tRNA binding. The guanidino group of the *N*-methyl-β-arginine moiety of BS further stabilizes the binding through hydrogen bonds to backbone phosphates of the P-site rRNA.[31] The

broad spectrum antibiotic activity of BS is likely due to the catalytic importance and high degree of conservation of the rRNA residues with which BS interacts.[31]

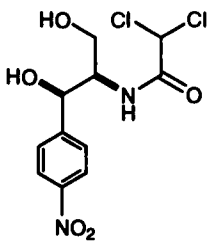
Blasticidin S Resistance Mechanisms

Antibiotic producing organisms typically require a method of self-resistance. A variety of self-resistance mechanisms have been identified in antibiotic producing Actinomycetes and include modification of the antibiotic target, and temporary inactivation of the antibiotic through chemical modification.

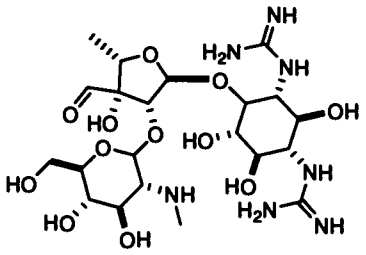
One antibiotic resistance mechanism is to prevent the accumulation of the compound inside the cell. This mechanism has been suggested to be involved in methylenomycin (**4.13**), chloramphenicol (**4.14**) and streptomycin (**4.15**) resistance by the producing organisms.[35-38] In these organisms, efflux pumps have been proposed to account for the observed antibiotic resistance. This type of resistance has been suggested to exist in *S. lividans* 66, that is highly resistant to BS *in vivo*, despite the observed high sensitivity of the *S. lividans* 66 ribosome to BS *in vitro*.[39] On agar plates, *S. lividans* 66 was able to grow with BS concentrations greater than 500 μg/mL, but purified ribosomes were inhibited by 10 μg/mL BS *in vitro*. Modification of BS, such as acetylation, was not detected and it was proposed that permeability mechanisms may explain the observations. This mechanism has not been clearly established in this system, and the existence of an alternate resistance mechanism can not be ruled out. Reduced cellular permeability to BS has also been suggested to be the mechanism of resistance in two *Saccharomyces cerevisiae* mutants.[40]



Methylenomycin A
4.13



Chloramphenicol
4.14



Streptomycin
4.15

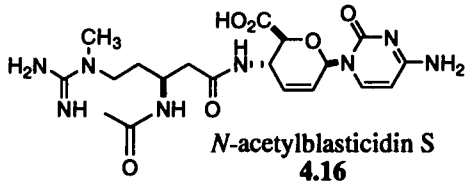
One common mechanism of antibiotic self-resistance is the modification of the antibiotic target in the producing organism. This mechanism has been demonstrated in *S.*

azureus and *S. erythreus*, the producing organisms of thiostrepton and erythromycin, respectively.[41-43] Thiostrepton and erythromycin derive their antibiotic activity by binding to the ribosome and inhibiting protein synthesis. In order to prevent binding to their own ribosomes, the producing organisms contain methylated residues within the 23S rRNA.[41,44,45] Methyltransferases have been identified that modify specific adenine residues on the 23S rRNA forming 2'-O-methyladenosine or N^6-dimethyladenosine.[41,44,45] The antibiotics do not bind strongly to methylated rRNA, which still functions properly, thus providing protection for the producing organism. Methylation of rRNA has also been proposed to be a resistance mechanism in the nosiheptide and berninamycin producing organisms, *S. actuosus* and *S. bernensis*.[46,47] Target site modification has not been observed in BS producing organisms and is not a likely resistance mechanism.

An alternative to modification of the antibiotic target is to chemically modify the antibiotic to produce an inactive or less active form of the compound. This is a common mechanism of self-resistance in Actinomycetes. Two common types of chemical modifications in *Streptomyces* include *N*-acetylation and *O*-phosphorylation. The biosynthesis of several antibiotics, including puromycin and streptomycin, involve the modification of toxic precursors to produce an inactive form of the antibiotic or biosynthetic intermediate. The last step in the biosynthesis of these antibiotics includes a hydrolytic step to remove the acetate or phosphate group to generate the active metabolite. Genes encoding acetyl and phospho hydrolases have been identified in the biosynthetic gene clusters of puromycin and streptomycin, respectively.[48,49]

Chemical modification of BS has been observed in several BS producers and non producers. One such mechanism involves the acylation of the β-amino group of the β-arginine moiety. Acylation is a common means of self-resistance found in other nucleoside antibiotic producers including *S. alboniger* (puromycin) and *S. lavendulae* (streptothricin F).[50-52] Genes encoding acetyltransferases have been cloned from puromycin, neomycin, and streptothricin producing organisms, among others.[41,43,51] This mechanism of BS resistance has been detected in *Streptoverticillium sp.* JCM4673 which is also a producer of BS.[53] The gene for BS *N*-acetyltransferase, *bls*, has been cloned and

characterized from this organism.[53] BS *N*-acetyltransferase catalyzes the formation of *N*-acetylblasticidin S (**4.16**) and requires acetyl-CoA.[53] BS acetyltransferase activity was also observed in the BS producer *S. griseochromogenes,* however it was determined that this was not the primary mechanism of self-resistance in that species.[54,55]



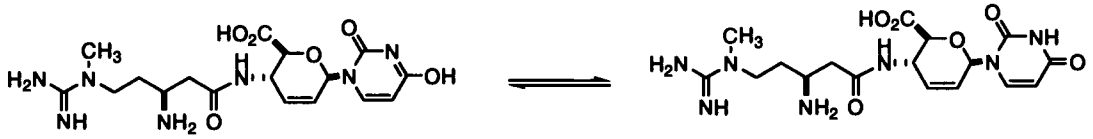
N-acetylblasticidin S
4.16

A second type of chemical modification that has been implicated in antibiotic self-resistance mechanisms is phosphorylation. Phosphotransferases, like kinases, require ATP and have been identified in the capreomycin, viomycin, chloramphenicol and neomycin producing organisms.[43,56,57] This group of enzymes catalyzes the transfer of the γ-phosphate from ATP to an oxygen on the substrate antibiotic. One example of a phosphotransferase is streptomycin 6-kinase that catalyzes the phosphorylation of the C6 hydroxyl of streptomycin.[58] This kinase is proposed to be necessary for the production of streptomycin as it is only found in *Streptomyces* strains that produce streptomycin.[41,59] Although many examples of self-resistance through phosphorylation exist, this resistance mechanism has not been detected in BS producing organisms.

The BS producer used for this work and in the biosynthesis studies described below, *S. griseochromogenes*, utilizes a novel self-resistance mechanism. Although BS acetyltransferase activity was identified in cell free extracts of *S. griseochromogenes*, neither *N*-acetylblasticidin S, nor an enzymatic activity to cleave *N*-acetylblasticidin S to BS could be detected.[54] Through studies conducted by Zhang, it was discovered the mechanism of self-resistance in *S. griseochromogenes* was through the addition of a leucine residue to the β-amine of demethylblasticidin S (DBS) to form leucyldemethylblasticidin S (LDBS).[55] The addition of a leucine to a BS biosynthetic intermediates is analogous to the acetylation or phosphorylation of other antibiotics and functions to reduce the toxicity of the compounds.[54] In addition, a leucylblasticidin S (LBS) hydrolyzing activity that resulted in the production of BS was also detected, further supporting this unique self-resistance mechanism and the intermediacy of LBS in BS biosynthesis.[11,54]

The presence of multiple self-resistance mechanisms has been observed in several antibiotic producing organisms. In addition to BS producers, multiple mechanisms of self-resistance have been identified in the neomycin and capreomycin producing organisms.[37,38] Work with the neomycin producer, *S. fradiae*, suggests that multiple mechanisms may be necessary for complete antibiotic resistance.[43] *S. fradiae* was shown to possess both neomycin acetyltransferase and phosphotransferase activities. When the genes for these two enzymes were cloned and expressed in *Pseudomonas aeruginosa* it was found that both enzymes were necessary to attain high levels of neomycin resistance. Similarly, it was found that *S. fradiae* also required both enzymes for high levels of resistance.[43] Similar results were found when the BS gene cluster was heterologously expressed in *S. lividans*.[60] It was shown that coexpression of a 2.6 kb BS resistance fragment, which is not part of the BS gene cluster, resulted in increased production of BS and BS intermediates. The detection of two resistance mechanisms in *S. griseochromogenes* has been proposed to provide an extra level of protection against toxic intermediates.[54]

A hydrolytic mechanism for resistance to BS has also been observed in non-producers. The earliest examples of BS resistance genes were identified in *Aspergillus terreus* (*bsd*)[61] and *Bacillus cereus* (*bsr*).[62] In 1972, Yamaguchi isolated a fungus, *A. terreus*, from paddy soil that was resistant to BS.[61,63] The resistance mechanism was determined to be the deamination of BS to form the uracil analog, deaminohydroxy BS (**4.17**). This conversion was found to be catalyzed by BS deaminase (BSD), which has been purified and characterized from *Aspergillus terreus* S-712. BSD was shown to be specific for BS and BS analogs and no amine hydrolysis was observed for cytosine related nucleotides or nucleosides.[64,65] Additionally, production of BSD could be induced through the addition of BS to the *A. terreus* culture medium.[61] The *bsd* gene has been cloned and sequenced,[66] and the *bsd* gene product, BSD, is a 30 kDa dimeric enzyme that contains one zinc molecule per subunit.[67]



Deaminohydroxy-blasticidin S
4.17

A second BS deaminase, BSR, has been identified in *Bacillus cereus*. Like BSD, BSR conveys BS resistance through the deamination of the cytosine moiety to create the uridine analog of BS. BSR activity was detected in BS resistant *B. cereus* and subsequently purified.[68] Characterization of BSR showed that, like BSD, it did not use cytidine or arabinosylcytosine as substrates; however it could deaminate BS, cytomycin and isoblasticidin S.[68] The Blasticidin S deaminase gene, *bsr*, was identified on a plasimid, pBSR8, from *B. cereus* and shown to confer BS resistance when introduced into *B. subtilis* and *E. coli*.[62,69]

Sequence analysis of BSD and BSR revealed that they show little similarity outside of a highly conserved region that is also shared by cytidine deaminases.[66,67] Cytidine deaminases are involved in pyrimidine salvage pathways of a number of organisms[70,71] and catalyze the similar deamination of a cytosine moiety to form the uridine analog. Like the BS deaminases, little similarity has been observed in either the sequence or tertiary structure of cytidine deaminases despite the similar activities of these enzymes.[71-73]

Cloning the genes for BS deaminases has allowed for the development of a variety of BS selectable marker systems for transformed eukaryotic cells, as well as transgenetic plants, fungi, and protozoans.[66,74-77] The high potency of BS, as well as its broad spectrum antibiotic activity has made BS a powerful tool for the selection of transformed cells. The *bsd* gene is a part of a selectable marker system used in transformed eukaryotic cells that is available commercially from Invitrogen Corp. The application of *bsr* as a resistance marker in mammalian cells has also been demonstrated.[78]

Previous Studies on Blasticidin S Biosynthesis:

The biosynthesis of BS has been extensively studied and has been reviewed.[2,22,79] The earliest studies by Seto *et al.* determined the primary precursors of blasticidin S.[80] The cytosinine moiety of was found to be derived from cytosine and D-glucose while the blastidic acid moiety was found to be derived from L-α-arginine with the methyl group being derived from L-methionine (Figure 4.1).[80] Seto also isolated the BS analogs,

leucylblasticidin S (LBS) and demethylblasticidin S (DBS) and proposed their possible biosynthetic intermediacy.[11,81] Later work by the Gould group identified a number of more advanced intermediates in BS biosynthesis through the incorporation of radiolabeled precursors in combination with cell free extract (CFE) studies.[2] These studies demonstrated the intermediacy of leucylblasticidin S.[54] Other identified precursors and intermediates included uridine-5'-diphosphoglucuronic acid (UDP-glucuronic acid),[82,83] and cytosylglucuronic acid (CGA).[82,83] Several attempts were made to utilize both whole cell and cell free systems to see incorporation of cytosinine into BS. Although initial reports by Yonehara suggested that incorporation of cytosinine into BS was observed, these results could not be duplicated by either Seto or Gould.[84,85] The intermediacy of cytosinine in BS biosynthesis has only been suggested indirectly through the identification of a cytosinine tautomerase activity in cell free extracts of *S. griseochromogenes.*[84]

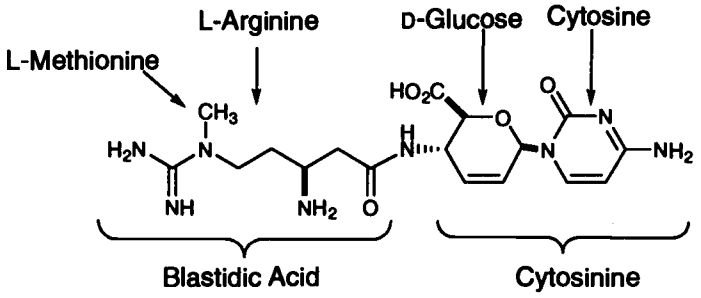


Figure 4.1. Primary precursors of Blasticidin S

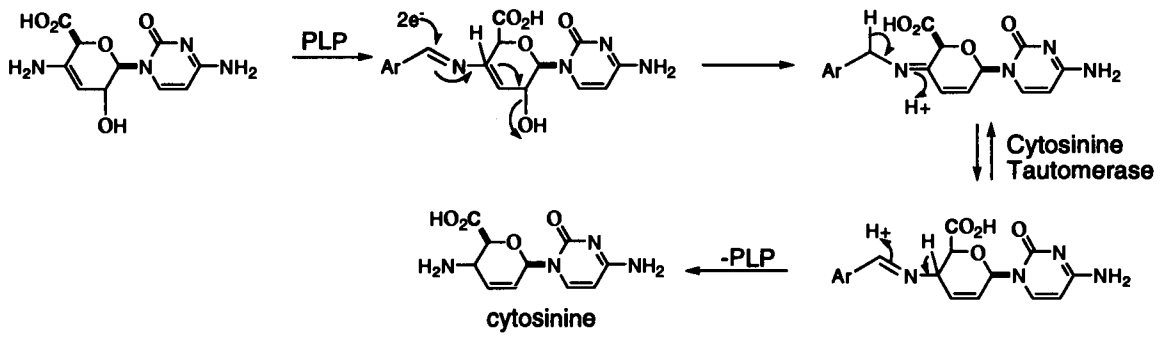
Gould *et al.* were able to identify three enzymes believed to be involved in the biosynthesis of BS. The first enzyme to be identified was found to catalyze the formation of CGA. CGA synthase was purified from *S. griseochromogenes* and was shown to form CGA from cytosine and UDP-glucuronic acid.[83] CGA synthase is specific for UDP-glucuronic acid as the sugar donor although it will accept 5-substituted cytosine analogs as well as cytosine as the acceptor.[83] Work by Kawashima supported these findings by showing the formation of a 5-fluoro analog of BS from the incorporation of 5-fluorocytosine into BS by *S. griseochromogenes.*[86]

The second enzyme activity in the BS biosynthetic pathway to be detected was a guanidino- *N*-methyltransferase. The addition of *S*-adenosyl-[$^{14}CH_3$]methionine and DBS

to *S. griseochromogenes* CFE resulted in the formation of BS, although at low levels.[87] Zhang *et al.* later demonstrated that LDBS was a better substrate for this guanidino-*N*-methyltransferase and concluded that LDBS is the likely natural substrate for this enzyme.[54]

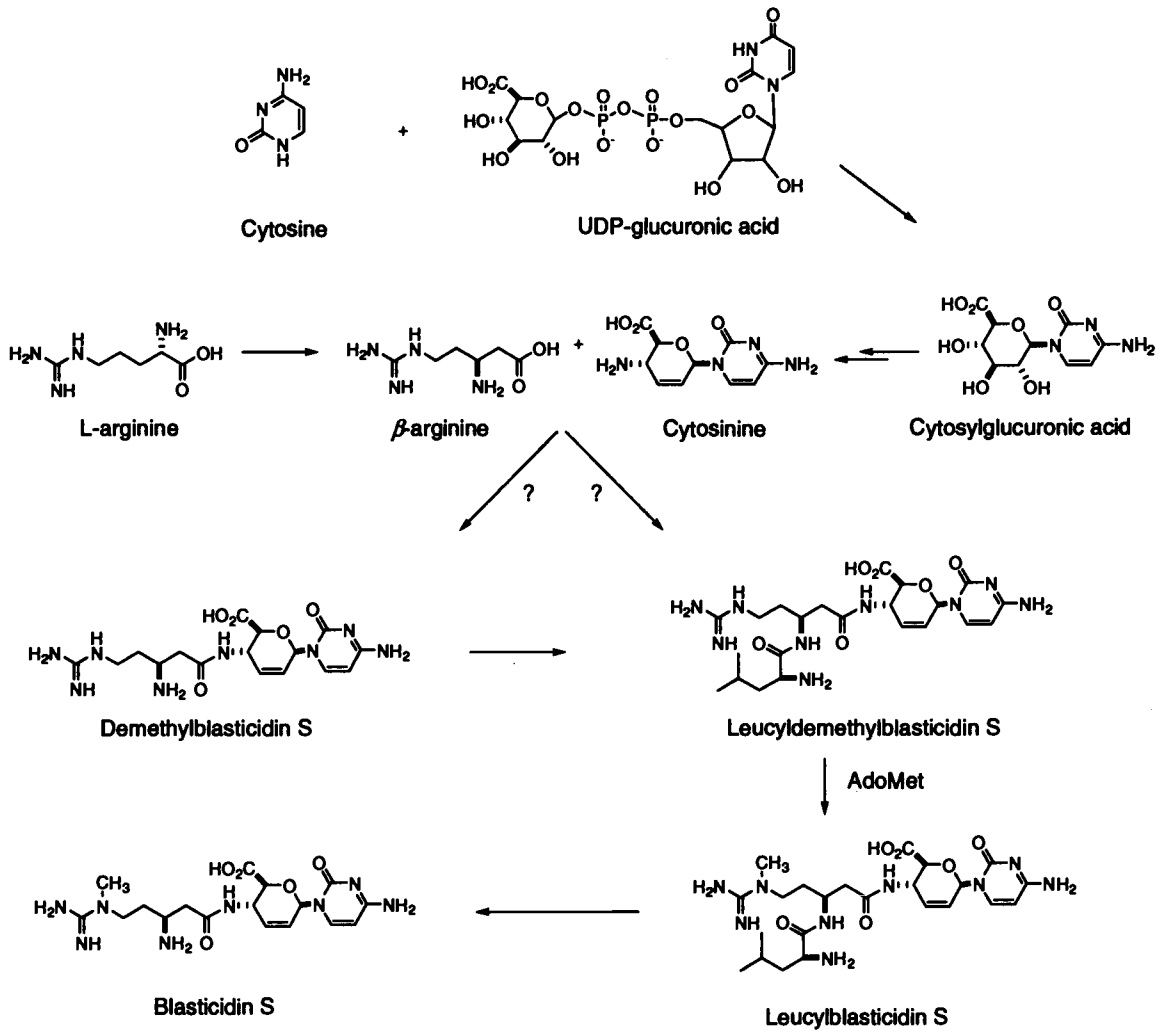
The third enzyme, and furthest along in the BS biosynthetic pathway, to be identified is a cytosinine pyridoxal phosphate tautomerase.[84] The tautomerase was discovered by Gould and Zhang through an experiment to determine the intermediacy of cytosinine in the biosynthesis of BS. While cytosinine had been proposed to be an intermediate, radiolabel incorporation studies with both *in vivo* and cell-free systems had not proven this.[84,85] Although they were not able to show the direct incorporation of cytosinine into BS, Gould and Zhang were able to demonstrate a protein dependent tautomerization of cytosinine, suggesting that a mechanism exists in *S. griseochromogenes* for the formation of this proposed intermediate.[84]

Incubation of deuterated cytosinine in CFE with pyridoxal phosphate resulted in the loss of deuterium from the C4' carbon. Correspondingly, when unlabeled cytosinine was incubated in *S. griseochromogenes* CFE prepared in D_2O, an enrichment in deuterium at C4' was observed. This evidence suggested the tautomerization of cytosinine was occurring, which was consistent with the proposed mechanism of cytosinine formation (Scheme 4.1). Based on these results, the authors concluded that a cytosinine tautomerase was present in *S. griseochromogenes* and that cytosinine was a true intermediate in BS biosynthesis.[55,84]



Scheme 4.1. Proposed mechanism for the formation of cytosinine. Reaction catalyzed by cytosinine pyridoxal phosphate tautomerase is indicated.

Based on the results from the precursor and self-resistance mechanism studies, Zhang *et al.* proposed the biosynthetic pathway to blasticidin S illustrated in Scheme 6.2.[54] Biosynthesis of BS is believed to begin with the formation of the cytosinine moiety. This begins with the production of cytosine from a nucleotide or nucleoside substrate (Scheme 4.2), as free cytosine is not a product or intermediate of primary metabolism.[70] The cytosine is then combined with glucuronic acid moiety from UDP-glucuronic acid to form cytosylglucuronic acid (CGA). Through several undefined transformations, which necessarily include deoxygenation and aminotransferase steps, the glucuronic acid moiety of CGA is converted to the aminosugar moiety of cytosinine.[88] The cytosinine is then coupled with the peptidyl portion of blasticidin S which can take one of two possible routes. It is not known whether this coupling occurs with a leucyl-β-arginine dipeptide or with just β-arginine to form either leucyldemethylblasticidin S (LDBS) or demethylblasticidin S (DBS). If the latter route is the actual path taken, the next step in the biosynthesis would be the addition of a leucine residue to form the intermediate, LDBS. The last step in the synthesis of the BS carbon skeleton is the addition of the *N*-methyl group by an *S*-adenosyl methionine dependent methyltransferase to form LBS. The final step in the biosynthesis of BS likely occurs after, or concurrent with, export of LBS from the cell and is the hydrolysis of the leucine-arginine bond to form BS.



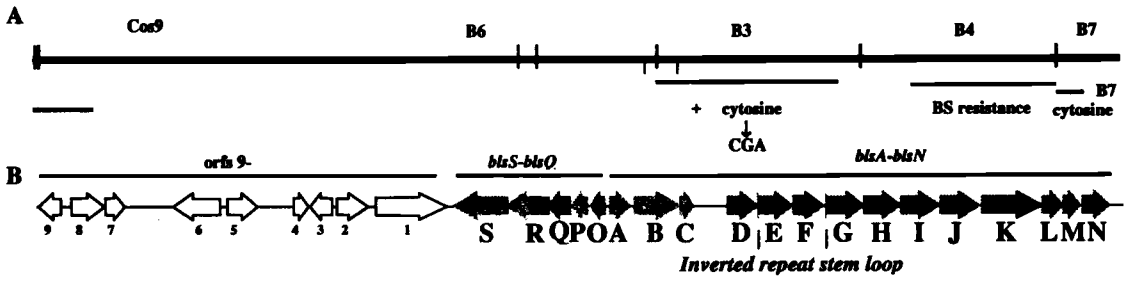
Scheme 4.2. Proposed biosynthetic pathway to blasticidin S

Cloning the Blasticidin S Biosynthesis Gene Cluster.

The blasticidin gene cluster was cloned from the BS producer *S. griseochromogenes* in 1998.[60] In prokaryotic organisms it is common for the antibiotic biosynthetic and self-resistance genes to be clustered on the chromosome.[89] This feature was exploited in order to identify the BS gene cluster. A genomic library of *S. griseochromogenes* DNA in the *Streptomyces* expression vector pIJ702 was used to transform *S. lividans* TK24. Transformants were screened for BS resistance and two small (2.6 and 4.8kb) chromosomal fragments were identified that conveyed blasticidin S

resistance to *S. lividans*. These two fragments were then used as gene probes to screen a *S. griseochromogenes* genomic cosmid library in pOJ446. Several cosmids were identified that hybridized to the 4.8 kb resistance probe and were found to produce various BS related metabolites, including CGA, DBS and LBS, when heterologously expressed in *S. lividans*.[54,60] The boundaries of the blasticidin S gene cluster were determined by comparing the range of metabolites produced with the restriction maps of other cosmids that also produced BS related metabolites.[90]

The 36.7 kb insert of one cosmid, cos9, was found to direct the production of late BS intermediates such as LBS and was subcloned and sequenced.[90] FramePlot software was used to identify open reading frames (ORFs) within the BS gene cluster. FramePlot is a program used to identify ORFs in bacterial DNA with high G+C content, such as *Streptomyces*.[91] FramePlot analysis is based on the preference for *Streptomyces* codon usage to contain a G or C in the third position and uses this feature to predict ORFs. Basic logical alignment search tool (BLAST) analysis of the identified ORFs, coupled with knowledge of BS biosynthesis from biochemical studies, was used to assign a proposed function to many of the genes located within the BS gene cluster.[92] Support for the assigned functions of several genes came from preliminary heterologous expression experiments. *Bam*HI subclones of cos9 in pIJ702 were heterologously expressed in *S. lividans*. The *S. lividans* culture broth was then monitored for the formation of BS related metabolites.[90] The blasticidin gene cluster is shown in Figure 4.2.



C

Protein	Proposed Function	Protein	Proposed Function	Protein	Proposed Function
BlsS	Oxidoreductase	BlsD	CGA synthase	BlsJ	self-resistance transporter
BlsR	unknown	BlsE	unknown	BlsK	unknown
BlsQ	transcriptional regulator	BlsF	unknown	BlsL	guanidino methyltransferase
BlsP	3-helix membrane protein	BlsG	arginine 2,3-aminomutase	BlsM	CMP Hydrolase
BlsO	2-helix membrane protein	BlsH	aminotrnasferase	BlsN	unknown
BlsA	methyltransferase	BlsI	ligase		
BlsC	regulatory protein				

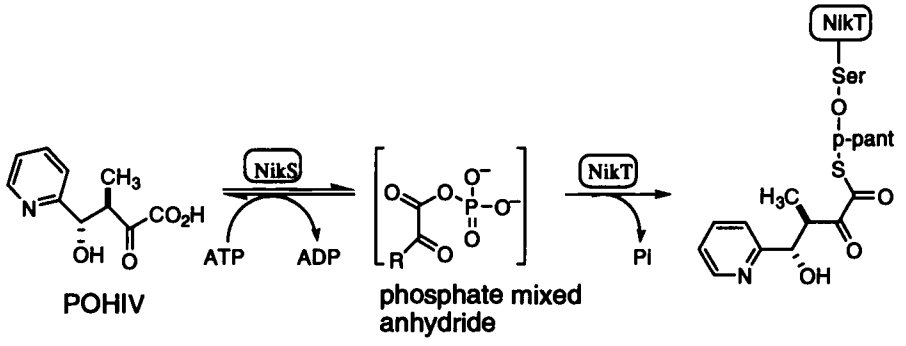
Figure 4.2. Blasticidin S Biosynthesis gene cluster. A. *Bam*HI restriction map of the BS cluster containing cos9 with functions associated with heterologously expressed fragments. B. Organization of the BS gene cluster. Filled ORFs are believed to be involved in BS biosynthesis. Unfilled ORFs are not believed to be associated with BS biosynthesis. C. Assigned functions for Bls proteins.

While many genes identified within the BS gene cluster did not show significant homology to proteins of known function, several other genes were found to have high similarity with characterized proteins and their involvement in blasticidin S biosynthesis could be predicted. *blsD* shares up to 42% similarity with glycosyltransferase genes and is believed to code for CGA synthase. A CGA synthase from *S. griseochromogenes* has been purified and was shown to form CGA from cytosine and UDP-glucuronic acid, as discussed previously.[82,83] Support for the assigned function of *blsD* came from heterologous expression studies. *S. lividans* cultures that were transformed with a 6.5 kb fragment of cos9 containing *blsD* in pIJ702 were found to produce CGA when supplemented with cytosine (Figure 4.2).[90]

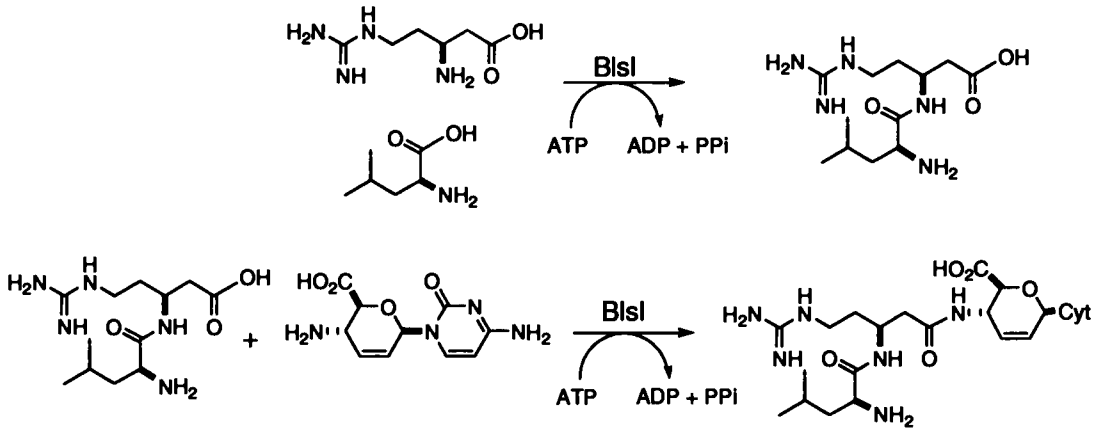
Several of the identified ORFs in the BS gene cluster are believed to be involved in formation of cytosinine from CGA. These proteins include BlsH, BlsQ, BlsR, and BlsS. BlsH has 48% similarity to an aminotransferase from *Amycolatopsis mediterranei*, RifH, and may be involved in formation of the aminosugar.[90,93] BlsQ-S are interesting in that the corresponding genes are similar to, and occur in the same order as, three genes of

unknown function found in the *Mycobacterium tuberculosis* H37Rv genome.[90,94] The similarity of *blsS* with a glucose-methanol-choline oxidoreductase suggests the involvement of these three proteins in formation of the aminosugar nucleoside.[90]

Another gene for which a biosynthetic function has been predicted based on sequence homology is *blsI*. *blsI* shares 28% identity, 40% similarity to the nikkomyin biosynthesis gene *nikS*. Based on gene disruption and complementation experiments, NikS has been shown to be a ligase that is responsible for loading 4-hydroxy-3-methyl-2-oxo-4-(2'pyridyl) butanoic acid (POHIV) onto the phosphopantetheine cofactor of NikT, an aminotransferase (Figure 4.3).[95] NikS is a member of the ATP-grasp-fold superfamily of proteins that share a common ATP binding motif and, unlike the adenylation domains of peptide synthetases, activate their substrates as the acylphosphate.[95] BlsI shares this ATP binding motif and is likely to be involved in ligation of leucine to either β-arginine or DBS to form the dipeptide or LDBS, respectively. An alternate function of this protein may be to join the amino acid or dipeptide moiety with cytosinine to form either DBS of LDBS, respectively (Figure 4.4).[90]



Scheme 4.3. Function of NikS and NikT from the nikkomycin biosynthetic pathway.



Scheme 4.4. Possible reactions catalyzed by BlsI. Top, formation of leucyl-β-arginine. Bottom, formation of leucyldemethyl blasticidin S.

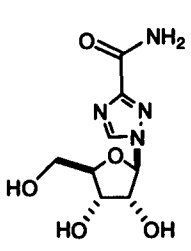
Another ORF with a readily predicted role in BS biosynthesis is BlsL. BlsL is homologous to a number of guanidino methyltransferases (KOG1709) and is likely to be involved in the methylation of the β-arginine.[90] BlsL contains the conserved guanidinoacetate methyltransferase motif as well as the highly conserved *S*-adenosylmethionine binding motif, EhGxGxGxhxxxhhΔ (h is a hydrophobic residue and Δ is a charged residue).[96] The *blsL* gene product is likely the guanidino- *N*-methyltransferase detected by Gould and by Zhang *et al.*.[54,87]

The 4.8 kb BS resistance gene probe used to identify the BS gene cluster contains *blsJ*. BlsJ is predicted to be an efflux or transport protein and is predicted to contain 11 transmembrane spanning domains.[90] Similar membrane proteins believed to be involved in antibiotic transport have been identified in other biosynthetic pathways.[37]

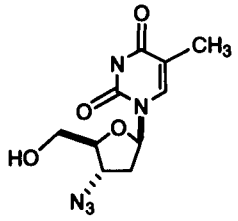
The initial steps in the biosynthesis of BS are believed to be the formation of free cytosine and β-arginine. Based on the sequence and similarity to known proteins, the functions for two ORFs located within the BS gene cluster were predicted to be involved in the formation of these BS precursors. *blsG* codes for a protein with significant sequence similarity to lysine 2,3-aminomutases and is believed to be involved in the formation of β-arginine. *blsM* is similar to several genes that code for deoxynucleoside transferases and was predicted to be involved in the formation of free cytosine. Our initial characterization of the BS gene cluster has focused on these two genes and they will be discussed more thoroughly in the following chapters.

Significance and Objectives of the Current Work.

Nucleoside antibiotics are an important class of compounds with diverse activities including antifungal, antiviral, insecticidal, and antitumor properties.[25] The biological activity of nucleoside antibiotics is typically derived from their ability to serve as nucleoside or nucleotide analogs. Nucleosides and nucleotides are involved in a wide range of metabolic functions within the cell including nucleic acid, protein, and glycan synthesis, cell signaling, energy storage, and enzymatic cofactors. The central importance and diverse intracellular functions of nucleosides has supported an ongoing interest in the development of nucleoside antibiotics. One successful nucleoside antibiotic is ribavirin, a clinically important drug used in the treatment of a number of viral diseases caused by hepatitis C virus (HCV), Lassa fever virus, and severe respiratory syncytial virus.[97-99] Another successful nucleoside antibiotic is the antiviral agent 3'-azido-2',3'-dideoxythymidine (AZT) that is an inhibitor of viral HIV-reverse transcriptase and is used in the treatment of AIDS.[100] By 1995, four nucleoside antibiotics were licensed for treatment of HIV, including AZT, however all of these compounds have toxicity issues.[100] The need for new, improved antiviral agents has supported the development of novel nucleoside antibiotics. Enzymatic synthesis of nucleoside analogs has been shown to be a useful method of preparing new antiviral compounds and has relied primarily on the use of two enzyme classes, nucleoside phosphorylases and deoxyribosyltransferases.[101] Nucleoside deaminases and lipases have also been employed in these enzymatic syntheses.[101]



Ribavirin



AZT

Despite the demonstrated utility of enzymatic synthesis, the development of new antibiotics through this approach is limited by the availability of enzymes capable of catalyzing unique chemical transformations. The expansion of this field of medicinal chemistry relies on the characterization and application of additional enzymes and

enzymatic activities. In this regard, the peptidyl nucleoside antibiotics are a particularly interesting group as they contain many unique and interesting structural features, such as unusual sugars and modified amino acids, and have demonstrated their utility as antibiotics in both agricultural and clinical settings. Members of this group have been used as agricultural fungicides (blasticidin S), insecticides (nikkomycin) and antiviral agents (ribavirin).[2,23,97]

Structural diversity among peptidyl nucleoside antibiotics is quite extensive. For example, more that two dozen peptidyl nucleoside natural products have been isolated from Gram-positive bacteria that contain a cytosine or modified cytosine moiety.[2] The structural diversity of these compounds is underlied by the common structural features shared by all peptidyl nucleosides. These similarities would suggest that both common and unique biosynthetic mechanisms are involved in the formation of these compounds. Through the characterization and elucidation of the enzymes and enzymatic mechanisms involved in these transformations, a clearer picture of the underlying chemistry can be obtained. From this information, molecular genetic manipulation of the individual genes could allow for the generation of enzymes with altered activities. This, in turn, may lead to the formation of new antibiotics with unique or altered biological activities. The number of available enzymes in the biosynthetic 'tool chest' increases with the number of characterized peptidyl nucleoside gene clusters. As more is understood about these systems, it may become possible to combine enzymes from various biosynthetic pathways to create novel engineered biosynthetic gene clusters and thus further increase the diversity of metabolite analogs that can be produced. In an ideal situation, it would be possible to utilize combinatorial biosynthesis to generate libraries of related metabolites for use in structure-activity relationship (SAR) studies and aid in drug development.

In addition to the BS gene cluster, the biosynthetic gene clusters for several peptidyl nucleoside antibiotics have been cloned and sequenced.[102-105] The puromycin gene cluster has been cloned and sequenced from *S. alboniger*.[102,106] Several individual enzymes involved in puromycin biosynthesis have been characterized including a 3'-amino-3'-dATP pyrophosphorohydratase, an ATP dehydrogenase, and an *N*-

acetylhydrolase.[48,107-109] Similarly, many of the genes involved in nikkomycin biosynthesis have also been characterized.[27,95,110] Together, the enzymes involved in the biosynthesis of these unique but related antibiotics, as well as those for BS biosynthesis, may be combined to create new engineered biosynthetic pathways.

The blasticidin S gene cluster is an excellent model system for peptidyl nucleoside biosynthesis as it is one of the least structurally complex members of this secondary metabolite family. As such, a molecular and enzymatic understanding of this system can serve as a basis for understanding the biosynthesis of more complex antibiotics. Another advantage of the BS system is the extensive amount of traditional biochemical studies that have been conducted. These studies have provided insight into BS biosynthesis and the chemical transformations involved. This information can now be used as a tool to assign functions to the various genes within the BS gene cluster and to aid in the characterization of the enzymes involved.

While there has been a tremendous amount of effort devoted to elucidating the steps in the biosynthesis of blasticidin S, the recent sequencing of the BS gene cluster affords new opportunities. Many of the biosynthetic experiments utilized by Gould and coworkers were restricted to the use of whole cell and cell-free systems due to the inability to purify adequate quantities of the biosynthetic proteins from the producing organism, *S. griseochromogenes*. The availability of the gene sequence for proteins involved in BS biosynthesis allows for the expression and complete characterization of the recombinant enzymes. An additional advantage of working with recombinant systems is the ability to utilize site directed mutagenesis to make amino acid mutations. These mutations can be used to study aspects of enzyme catalysis and may be utilized to alter substrate specificity and create new blasticidin S analogs.

The characterization of individual enzymes involved in the biosynthesis of blasticidin S started with two enzymes believed to catalyze the initial steps. One chemical conversion that has attracted attention since the discovery of BS is the formation of the β-arginine moiety. Based on its high degree of sequence similarity with lysine 2,3-aminomutase, the *blsG* gene product was predicted to be involved in the

formation of β-arginine. Chapter 5 will discuss a variety of techniques and approaches that were employed to elucidate the function of BlsG.

The first step in the biosynthesis of the nucleoside moiety of blasticidin S is believed to be the generation of free cytosine. This step would be necessary as free cytosine is not normally present within the cell. Based on sequence similarities, it was hypothesized that the *blsM* gene product was responsible for this conversion. The cloning, heterologous expression and characterization of BlsM will be discussed in Chapter 6.

References

(1) Takeuchi, S.; Hirayama, K.; Ueda, K.; Sakai, H.; Yonehara, H. Blasticidin S, a new antibiotic. *J. Antibiot. (Tokyo)* **1958**, *11*, 1-5.

(2) Gould, S. J. Blasticidin S and Related Peptidyl Nucleoside Antibiotics. *Biotechnology of Antibiotics*; Marcel Dekker, Inc.: New York, 1997; pp 703-731.

(3) Huang, K. T.; Misato, T.; Asuyama, H. Effect of Blasticidin S on Protein Synthesis of *Piricularia Oryzae*. *J. Antibiot. (Tokyo)* **1964**, *17*, 65-70.

(4) Huang, K. T.; Misato, T.; Asuyama, H. Selective Toxicity of Blasticidin S to *Piricularia Oryzae* and *Pellicularia Sasakii*. *J. Antibiot. (Tokyo)* **1964**, *17*, 71-74.

(5) Rodriguez-Fonseca, C.; Amils, R.; Garrett, R. A. Fine structure of the peptidyl transferase centre on 23 S-like rRNAs deduced from chemical probing of antibiotic-ribosome complexes. *J Mol Biol* **1995**, *247*, 224-235.

(6) Hirai, T.; Shimomura, T. Blasticidin S, An Effective Antibiotic Against Plant Virus Multiplication. *Phytopathology* **1965**, *55*, 291-295.

(7) Hirai, A.; Wildman, S. G.; Hirai, T. Specific Inhibition of TMV-RNA Synthesis by Blasticidin S. *Virology* **1968**, *36*, 646-651.

(8) Shomura, T.; Inoue, M.; Niida, T.; Hara, T. The Improved Assay Method of Blasticidin S. *J. Antibiot. (Tokyo)* **1964**, *17*, 253-261.

(9) Takana, Y.; Sakagami, Y.; Nishimura, T.; Yamaki, H.; Umezawa, H. Activity of cytomycin and blasticidin S against transplantable animal tumors. *J. Antibiot.* **1961**, *14*, 123-126.

(10) Seto, H.; Yonehara, H. Studies on the biosynthesis of blasticidin S. VI. The isolation and structure of blasticidin H. *J. Antibiot. (Tokyo)* **1977**, *30*, 1019-1021.

(11) Seto, H.; Otake, N.; Yonehara, H. Biosynthesis of blasticidin S. II. Leucylblasticidin S, a metabolic intermediate of blasticidin S biosynthesis. *Agric. Biol Chem.* **1968**, *32*, 1299-1305.

(12) Sugiyama, M.; Takeda, A.; Paik, S. Y.; Nimi, O.; Nomi, R. Acetylation of blasticidin S by its producing actinomycetes. *J. Antibiot. (Tokyo)* **1986**, *39*, 827-832.

(13) Larsen, S. H.; Berry, D. M.; Paschal, J. W.; Gilliam, J. M. 5-Hydroxymethylblasticidin S and blasticidin S from *Streptomyces setonii* culture A83094. *J. Antibiot. (Tokyo)* **1989**, *42*, 470-471.

(14) Gullo, V.; Conover, M.; Cooper, R.; Federbush, C.; Horan, A. C.; Kung, T.; Marquez, J.; Patel, M.; Watnick, A. Sch 36605, a novel anti-inflammatory compound. Taxonomy, fermentation, isolation and biological properties. *J. Antibiot. (Tokyo)* **1988**, *41*, 20-24.

(15) Cooper, R.; Conover, M.; Patel, M. Sch 36605, structure of a novel nucleoside. *J. Antibiot. (Tokyo)* **1988**, *41*, 123-125.

(16) Argoudelis, A. D.; Baczynskyj, L.; Kuo, M. T.; Laborde, A. L.; Sebek, O. K.; Truesdell, S. E.; Shilliday, F. B. Arginomycin: production, isolation, characterization and structure. *J. Antibiot. (Tokyo)* **1987**, *40*, 750-760.

(17) Feduchi, E.; Cosin, M.; Carrasco, L. Mildiomycin: a nucleoside antibiotic that inhibits protein synthesis. *J. Antibiot. (Tokyo)* **1985**, *38*, 415-419.

(18) Endo, T.; Otake, N.; Takeuchi, S.; Yonehara, H. The Structure of Blastidone, a Degradative Component of Blasticidin S. *J. Antibiot. (Tokyo)* **1964**, *17*, 172-173.

(19) Otake, N.; Takeuchi, S.; Endo, T.; Yonehara, H. Structure of cytosinine, a structural component of blasticidin S. *Tetrahedron Lett.* **1965**, *19*, 1405-1409.

(20) Otake, N.; Takeuchi, S.; Endo, T.; Yonehara, H. The structure of blasticidin S. *Tetrahedron Lett.* **1965**, *19*, 1411-1419.

(21) Swaminathan, V.; Smith, J. L.; Sundaralingam, M.; Coutsogeorgopoulos, C.; Kartha, G. Crystal and molecular structure of the antibiotic blasticidin S hydrochloride pentahydrate. *Biochim. Biophys. Acta.* **1981**, *655*, 335-341.

(22) Suhadolnik, R. J. *Nucleoside Antibiotics*; Wiley-Interscience: New York, 1970; 442.

(23) Zhang, D.; Miller, M. J. Polyoxins and nikkomycins: progress in synthetic and biological studies. *Curr. Pharm. Des.* **1999**, *5*, 73-99.

(24) Isono, K. Current progress on nucleoside antibiotics. *Pharmacol. Ther.* **1991**, *52*, 269-286.

(25) Isono, K. Nucleoside antibiotics: structure, biological activity, and biosynthesis. *J. Antibiot. (Tokyo)* **1988**, *41*, 1711-1739.

(26) Thiruvengadam, T. K.; Gould, S. J.; Aberhart, D. J.; Lin, H. Biosynthesis of Streptothricin F. 5. Formation of β-Lysine by *Streptomyces* L-1689-23. *J. Am. Chem. Soc.* **1983**, *105*, 5470-5476.

(27) Chen, H.; Hubbard, B. K.; O'Connor, S. E.; Walsh, C. T. Formation of beta-hydroxy histidine in the biosynthesis of nikkomycin antibiotics. *Chem. Biol.* **2002**, *9*, 103-112.

(28) Tercero, J. A.; Espinosa, J. C.; Lacalle, R. A.; Jimenez, A. The biosynthetic pathway of the aminonucleoside antibiotic puromycin, as deduced from the molecular analysis of the pur cluster of Streptomyces alboniger. *J Biol Chem* **1996**, *271*, 1579-1590.

(29) Jackson, M. D.; Gould, S. J.; Zabriskie, T. M. Studies on the formation and incorporation of streptolidine in the biosynthesis of the peptidyl nucleoside antibiotic streptothricin F. *J. Org. Chem.* **2002**, *67*, 2934-2941.

(30) Porse, B. T.; Rodriguez-Fonseca, C.; Leviev, I.; Garrett, R. A. Antibiotic inhibition of the movement of tRNA substrates through a peptidyl transferase cavity. *Biochem Cell Biol* **1995**, *73*, 877-885.

(31) Hansen, J. L.; Moore, P. B.; Steitz, T. A. Structures of five antibiotics bound at the peptidyl transferase center of the large ribosomal subunit. *J Mol Biol* **2003**, *330*, 1061-1075.

(32) Yamaguchi, H.; Tanaka, N. Inhibition of protein synthesis by blasticidin S. II. Studies on the site of action in *E. coli* polypeptide synthesizing systems. *J. Biochem. (Tokyo)* **1966**, *60*, 632-642.

(33) Weisblum, B.; Davies, J. Antibiotic Inhibitors of the Bacterial Ribosome. *Bacteriol. Rev.* **1968**, *32*, 493-528.

(34) Yamaguchi, H.; Yamamoto, C.; Tanaka, N. Inhibition of protein synthesis by blasticidin S. I. Studies with cell-free systems from bacterial and mammalian cells. *J. Biochem. (Tokyo)* **1965**, *57*, 667-677.

(35) Sugiyama, M.; Mochizuki, H.; Nimi, O.; Nomi, R. Mechanism of protection of protein synthesis against streptomycin inhibition in a producing strain. *J. Antibiot. (Tokyo)* **1981**, *34*, 1183-1188.

(36) Malik, V. S.; Vining, L. C. Chloramphenicol resistance in a chloramphenicol-producing *Streptomyces*. *Can. J. Microbiology* **1972**, *18*, 583-590.

(37) Cundliffe, E. Self-protection mechanisms in antibiotic producers. *Ciba. Found. Symp.* **1992**, *171*, 199-208; discussion 208-114.

(38) Cundliffe, E. How antibiotic-producing organisms avoid suicide. *Annu. Rev. Microbiol.* **1989**, *43*, 207-233.

(39) Nomura, H.; Paik, S. Y.; Izuhara, M.; Sakagami, Y. Resistant Mechanism to Blasticidin S in *Streptomyces lividans*. *Korean Biochem. J.* **1991**, *24*, 327-330.

(40) Ishiguro, J.; Miyazaki, M. Characterization of blasticidin S-resistant mutants of *Saccharomyces cerevisiae*. *Curr. Genet.* **1985**, *9*, 179-181.

(41) Cundliffe, E. Self defence in antibiotic-producing organisms. *Br. Med. Bull.* **1984**, *40*, 61-67.

(42) Cundliffe, E. Mechanism of resistance to thiostrepton in the producing-organism *Streptomyces azureus*. *Nature* **1978**, *272*, 792-795.

(43) Thompson, C. J.; Skinner, R. H.; Thompson, J.; Ward, J. M.; Hopwood, D. A.; Cundliffe, E. Biochemical characterization of resistance determinants cloned from antibiotic-producing streptomycetes. *J. Bacteriol.* **1982**, *151*, 678-685.

(44) Thompson, J.; Schmidt, F.; Cundliffe, E. Site of action of a ribosomal RNA methylase conferring resistance to thiostrepton. *J. Biol. Chem.* **1982**, *257*, 7915-7917.

(45) Skinner, R.; Cundliffe, E.; Schmidt, F. J. Site of action of a ribosomal RNA methylase responsible for resistance to erythromycin and other antibiotics. *J. Biol. Chem.* **1983**, *258*, 12702-12706.

(46) Cundliffe, E.; Thompson, J. The mode of action of nosiheptide (multhiomycin) and the mechanism of resistance in the producing organism. *J.Gen. Microbiol.* **1981**, *126*, 185-192.

(47) Thompson, J.; Cundliffe, E.; Stark, M. J. The mode of action of berninamycin and mechanism of resistance in the producing organism, *Streptomyces bernensis*. *J. Gen. Microbiol.* **1982**, *128 (Pt 4)*, 875-884.

(48) Lacalle, R. A.; Tercero, J. A.; Vara, J.; Jimenez, A. Identification of the gene encoding an N-acetylpuromycin N-acetylhydrolase in the puromycin biosynthetic gene cluster from *Streptomyces alboniger*. *J. Bacteriol.* **1993**, *175*, 7474-7478.

(49) Mansouri, K.; Piepersberg, W. Genetics of streptomycin production in Streptomyces griseus: nucleotide sequence of five genes, strFGHIK, including a phosphatase gene. *Mol Gen Genet* **1991**, *228*, 459-469.

(50) Vara, J.; Malpartida, F.; Hopwood, D. A.; Jimenez, A. Cloning and expression of a puromycin N-acetyl transferase gene from *Streptomyces alboniger* in *Streptomyces lividans* and *Escherichia coli*. *Gene* **1985**, *33*, 197-206.

(51) Kobayashi, T.; Horinouchi, S.; Uozumi, T.; Beppu, T. Purification and biochemical characterization of streptothricin acetyltransferase coded by the cloned streptothricin-resistance gene of *Streptomyces lavendulae*. *J. Antibiot. (Tokyo)* **1987**, *40*, 1016-1022.

(52) Horinouchi, S.; Furuya, K.; Nishiyama, M.; Suzuki, H.; Beppu, T. Nucleotide sequence of the streptothricin acetyltransferase gene from *Streptomyces lavendulae* and its expression in heterologous hosts. *J. Bacteriol.* **1987**, *169*, 1929-1937.

(53) Perez-Gonzalez, J. A.; Ruiz, D.; Esteban, J. A.; Jimenez, A. Cloning and characterization of the gene encoding a blasticidin S acetyltransferase from *Streptoverticillum sp. Gene* **1990**, *86*, 129-134.

(54) Zhang, Q.; Cone, M. C.; Gould, S. J.; Zabriskie, T. M. Reevaluation of the Final Steps in the Biosynthesis of Blasticidin S by *Streptomyces griseochromogenes* and Identification of a Novel Self-Resistance Mechanism. *Tetrahedron* **2000**, *56*, 693-701.

(55) Zhang, Q. Advanced Studies of Blasticidin S Biosynthesis. In *Chemistry*; Oregon State University: Corvallis, 1998; pp 159.

(56) Skinner, R. H.; Cundliffe, E. Resistance to the antibiotics viomycin and capreomycin in the *Streptomyces* species which produce them. *J. Gen. Microbiol.* **1980**, *120*, 95-104.

(57) Mosher, R. H.; Camp, D. J.; Yang, K.; Brown, M. P.; Shaw, W. V.; Vining, L. C. Inactivation of chloramphenicol by O-phosphorylation. A novel resistance mechanism in *Streptomyces venezuelae* ISP5230, a chloramphenicol producer. *J. Biol. Chem.* **1995**, *270*, 27000-27006.

(58) Sugiyama, M.; Sakamoto, M.; Mochizuki, H.; Nimi, O.; Nomi, R. Purification and Characterization of Streptomycin 6-kinase, an Enzyme Implicated in Self-protection of a Streptomycin-producing Micro-organism. *J. Gen. Micorbiol.* **1983**, *129*, 1683-1687.

(59) Walker, M. S.; Walker, J. B. Streptomycin Biosynthesis: separation and substrate specificities of phosphatases acting on guanidino-deoxy-*scyllo*-inositol posphate and streptomycin -(streptidino)-phosphate. *J. Biol. Chem.* **1971**, *246*, 7034-7040.

(60) Cone, M. C.; Petrich, A. K.; Gould, S. J.; Zabriskie, T. M. Cloning and heterologous expression of blasticidin S biosynthetic genes from *Streptomyces griseochromogenes*. *J. Antibiot. (Tokyo)* **1998**, *51*, 570-578.

(61) Yamaguchi, I.; Shibata, H.; Seto, H.; Misato, T. Isolation and purification of blasticidin S deaminase from Aspergillus terreus. *J. Antibiot. (Tokyo)* **1975**, *28*, 7-14.

(62) Kamakura, T.; Kobayashi, K.; Tanaka, T.; Yamaguchi, I.; Endo, T. Cloning and expression of a new structural gene for blasticidin S deaminase, a nucleoside aminohydrolase. *Agric. Biol. Chem* **1987**, *51*, 3165.

(63) Yamaguchi, I.; Takagi, K.; Misato, T. The Sites of Degradation of Blasticidin S. *Agric. Biol. Chem.* **1972**, *36*, 1719-1727.

(64) Yamaguchi, I.; Misato, T. Active Center and Mode of Reaction of Blasticidin S Deaminase. *Agric. Biol. Chem.* **1985**, *49*, 3355-3361.

(65) Yamaguchi, I.; Seto, H.; Misato, T. Substrate Binding by Blasticidin S Deaminase, an Aminohydrolase for Novel 4-aminopyrimidine Nucleosides. *Pest. Biochem. Phys.* **1986**, *25*, 54-62.

(66) Kimura, M.; Kamakura, T.; Tao, Q. Z.; Kaneko, I.; Yamaguchi, I. Cloning of the blasticidin S deaminase gene (BSD) from *Aspergillus terreus* and its use as a selectable marker for *Schizosaccharomyces pombe* and *Pyricularia oryzae*. *Mol. Gen. Genet.* **1994**, *242*, 121-129.

(67) Kimura, M.; Yamaguchi, I. Recent Development in the Use of Blasticidin S, a Microbial Fungicide, as a Useful Reagent in Molecular Biology. *Pest. Biochem. Phys.* **1997**, *56*, 243-248.

(68) Endo, T.; Furuta, K.; Kaneko, A.; Katsuki, T.; Kobayashi, K.; Azuma, A.; Watanabe, A.; Shimazu, A. Inactivation of blasticidin S by *Bacillus cereus*. I. Inactivation mechanism. *J. Antibiot. (Tokyo)* **1987**, *40*, 1791-1793.

(69) Endo, T.; Kobayashi, K.; Nakayama, N.; Tanaka, T.; Kamakura, T.; Yamaguchi, I. Inactivation of blasticidin S by *Bacillus cereus*. II. Isolation and characterization of a plasmid, pBSR8, from *Bacillus cereus*. *J. Antibiot. (Tokyo)* **1988**, *41*, 271-273.

(70) Munch-Petersen, A. *Metabolism of Nucleotides, Nucleosides and Nucleobases in Microorganisms*; Academic Press Inc.: San Diego, 1983.

(71) Yang, C.; Carlow, D.; Wolfenden, R.; Short, S. A. Cloning and nucleotide sequence of the *Escherichia coli* cytidine deaminase (ccd) gene. *Biochemistry* **1992**, *31*, 4168-4174.

(72) Ireton, G. C.; Black, M. E.; Stoddard, B. L. The 1.14 A crystal structure of yeast cytosine deaminase: evolution of nucleotide salvage enzymes and implications for genetic chemotherapy. *Structure* **2003**, *11*, 961-972.

(73) Betts, L.; Xiang, S.; Short, S. A.; Wolfenden, R.; Carter, C. W., Jr. Cytidine deaminase. The 2.3 A crystal structure of an enzyme: transition-state analog complex. *J. Mol. Biol.* **1994**, *235*, 635-656.

(74) Brooks, D. R.; McCulloch, R.; Coombs, G. H.; Mottram, J. C. Stable transformation of trypanosomatids through targeted chromosomal integration of the selectable marker gene encoding blasticidin S deaminase. *FEMS Microbiol. Lett.* **2000**, *186*, 287-291.

(75) Mamoun, C. B.; Gluzman, I. Y.; Goyard, S.; Beverley, S. M.; Goldberg, D. E. A set of independent selectable markers for transfection of the human malaria parasite *Plasmodium falciparum*. *Proc. Natl. Acad. Sci. U S A* **1999**, *96*, 8716-8720.

(76) Tamura, K.; Kimura, M.; Yamaguchi, I. Blasticidin S deaminase gene (BSD): a new selection marker gene for transformation of *Arabidopsis thaliana* and *Nicotiana tabacum*. *Biosci. Biotechnol. Biochem.* **1995**, *59*, 2336-2338.

(77) Kimura, M.; Takatsuki, A.; Yamaguchi, I. Blasticidin S deaminase gene from *Aspergillus terreus* (BSD): a new drug resistance gene for transfection of mammalian cells. *Biochim. Biophys. Acta* **1994**, *1219*, 653-659.

(78) Izumi, M.; Miyazawa, H.; Kamakura, T.; Yamaguchi, I.; Endo, T.; Hanaoka, F. Blasticidin S-resistance gene (bsr): a novel selectable marker for mammalian cells. *Exp. Cell. Res.* **1991**, *197*, 229-233.

(79) Yonehara, H. Blasticidin S: Properties, Biosynthesis, and Fermentation. *Biotechnology of Industrial Antibiotics*; Marcel Dekker, Inc.: New York, 1984.

(80) Seto, H.; Yamaguchi, N.; Otake, N.; Yonehara, H. Studies on the biosynthesis of blasticidin S. Part I. Precursors of blasticidin S biosynthesis. *Agric. Biol. Chem.* **1968**, *32*, 1292-1298.

(81) Seto, H.; Yonehara, H. Studies on the biosynthesis of blasticidin S. VII. Isolation of demethylblasticidin S. *J. Antibiot. (Tokyo)* **1977**, *30*, 1022-1024.

(82) Guo, J.; Gould, S. J. Biosynthesis of Blasticidin S from Cytosylglucuronic Acid (CGA). Isolation of Cytosine/UDPglucuronoxyltransferase and Incorporation of CGA by *Streptomyces griseochromogenes*. *J. Am. Chem. Soc.* **1991**, *113*, 5898-5899.

(83) Gould, S. J.; Guo, J. Cytosylglucuronic acid synthase (cytosine: UDP-glucuronosyltransferase) from *Streptomyces griseochromogenes*, the first prokaryotic UDP-glucuronosyltransferase. *J. Bacteriol.* **1994**, *176*, 1282-1286.

(84) Gould, S. J.; Zhang, Q. Cytosinine: pyridoxal phosphate tautomerase, a new enzyme in the blasticidin S biosynthetic pathway. *J. Antibiot. (Tokyo)* **1995**, *48*, 652-656.

(85) Yonehara, H.; Otake, N. Biological and chemical reconstruction of blasticidin S. *Antimicrob. Agents Chemother.* **1965**, 855-857.

(86) Kawashima, A.; Seto, H.; Ishiyama, T.; Kato, M.; Uchida, K.; Otake, N. Fluorinated blasticidin S. *Agric. Biol. Chem.* **1987**, *51*, 1183-1184.

(87) Gould, S. J.; Guo, J. Biosynthesis of blasticidin S. Cell-free demonstration of N-methylation as the last step. *Bioorg. Med. Chem. Lett.* **1991**, *1*, 497-500.

(88) Zhang, Q.; Gould, S. J.; Zabriskie, T. M. A new cytosine glycoside from *Streptomyces griseochromogenes* produced by the use in vivo of enzyme inhibitors. *J. Nat. Prod.* **1998**, *61*, 648-651.

(89) Seno, E. T.; Baltz, R. M. Structural Organization and regulation of antibiotic biosynthesis and resistance genes in actinomycetes. *Regulation of Secondary Metabolism in Actinomycetes*; CRC Press: Boca Raton, 1989; pp 1-48.

(90) Cone, M. C.; Yin, X.; Grochowski, L. L.; Parker, M. R.; Zabriskie, T. M. The blasticidin S biosynthesis gene cluster from *Streptomyces griseochromogenes*:

sequence analysis, organization, and initial characterization. *Chembiochem* **2003**, *4*, 821-828.

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

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

(93) Kim, C. G.; Yu, T. W.; Fryhle, C. B.; Handa, S.; Floss, H. G. 3-Amino-5-hydroxybenzoic acid synthase, the terminal enzyme in the formation of the precursor of mC7N units in rifamycin and related antibiotics. *J Biol Chem* **1998**, *273*, 6030-6040.

(94) Cole, S. T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S. V.; Eiglmeier, K.; Gas, S.; Barry, C. E., 3rd; Tekaia, F.; Badcock, K.; Basham, D.; Brown, D.; Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Hornsby, T.; Jagels, K.; Barrell, B. G.; et al. Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. *Nature* **1998**, *393*, 537-544.

(95) Lauer, B.; Russwurm, R.; Schwarz, W.; Kalmanczhelyi, A.; Bruntner, C.; Rosemeier, A.; Bormann, C. Molecular characterization of co-transcribed genes from *Streptomyces tendae* Tu901 involved in the biosynthesis of the peptidyl moiety and assembly of the peptidyl nucleoside antibiotic nikkomycin. *Mol. Gen. Genet.* **2001**, *264*, 662-673.

(96) Wu, G.; Williams, H. D.; Zamanian, M.; Gibson, F.; Poole, R. K. Isolation and characterization of *Escherichia coli* mutants affected in aerobic respiration: the cloning and nucleotide sequence of ubiG. Identification of an *S*-adenosylmethionine-binding motif in protein, RNA, and small-molecule methyltransferases. *J. Gen. Microbiol.* **1992**, *138 (Pt 10)*, 2101-2112.

(97) Vo, N. V.; Young, K. C.; Lai, M. M. Mutagenic and inhibitory effects of ribavirin on hepatitis C virus RNA polymerase. *Biochemistry* **2003**, *42*, 10462-10471.

(98) Edell, D.; Khoshoo, V.; Ross, G.; Salter, K. Early ribavarin treatment of bronchiolitis: effect on long-term respiratory morbidity. *Chest* **2002**, *122*, 935-939.

(99) Malnick, S. D.; Beergabel, M.; Lurie, Y. Treatment of chronic hepatitis C virus infection. *Ann. Pharmacother.* **2000**, *34*, 1156-1164.

(100) Strohl, W. *Biotechnology of Antibiotics*; Marcel Dekker, Inc.: New York, 1997; 842.

(101) Hanrahan, J. R.; Hutchinson, D. W. The enzymatic synthesis of antiviral agents. *J. Biotech.* **1992**, *23*, 193-210.

(102) Tercero, J. A.; Espinosa, J. C.; Lacalle, R. A.; Jimenez, A. The biosynthetic pathway of the aminonucleoside antibiotic puromycin, as deduced from the molecular analysis of the pur cluster of *Streptomyces alboniger*. *J. Biol. Chem.* **1996**, *271*, 1579-1590.

(103) Saugar, I.; Sanz, E.; Rubio, M. A.; Espinosa, J. C.; Jimenez, A. Identification of a set of genes involved in the biosynthesis of the aminonucleoside moiety of

antibiotic A201A from *Streptomyces capreolus*. *Eur. J. Biochem.* **2002**, *269*, 5527-5535.

(104) Fernandez-Moreno, M. A.; Vallin, C.; Malpartida, F. Streptothricin biosynthesis is catalyzed by enzymes related to nonribosomal peptide bond formation. *J. Bacteriol.* **1997**, *179*, 6929-6936.

(105) Bormann, C.; Mohrle, V.; Bruntner, C. Cloning and heterologous expression of the entire set of structural genes for nikkomycin synthesis from *Streptomyces tendae* Tu901 in *Streptomyces lividans*. *J. Bacteriol.* **1996**, *178*, 1216-1218.

(106) Lacalle, R. A.; Tercero, J. A.; Jimenez, A. Cloning of the complete biosynthetic gene cluster for an aminonucleoside antibiotic, puromycin, and its regulated expression in heterologous hosts. *Embo. J.* **1992**, *11*, 785-792.

(107) Tercero, J. A.; Lacalle, R. A.; Jimenez, A. The pur8 gene from the pur cluster of *Streptomyces alboniger* encodes a highly hydrophobic polypeptide which confers resistance to puromycin. *Eur. J. Biochem.* **1993**, *218*, 963-971.

(108) Rubio, M. A.; Espinosa, J. C.; Tercero, J. A.; Jimenez, A. The Pur10 protein encoded in the gene cluster for puromycin biosynthesis of *Streptomyces alboniger* is an NAD-dependent ATP dehydrogenase. *FEBS Lett.* **1998**, *437*, 197-200.

(109) Espinosa, J. C.; Tercero, J. A.; Rubio, M. A.; Jimenez, A. The pur7 gene from the puromycin biosynthetic pur cluster of *Streptomyces alboniger* encodes a nudix hydrolase. *J. Bacteriol.* **1999**, *181*, 4914-4918.

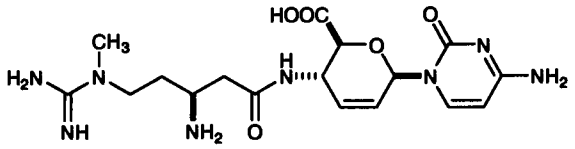
(110) Lauer, B.; Russwurm, R.; Bormann, C. Molecular characterization of two genes from *Streptomyces tendae* Tu901 required for the formation of the 4-formyl-4-imidazolin-2-one-containing nucleoside moiety of the peptidyl nucleoside antibiotic nikkomycin. *Eur. J. Biochem.* **2000**, *267*, 1698-1706.

Chapter 5

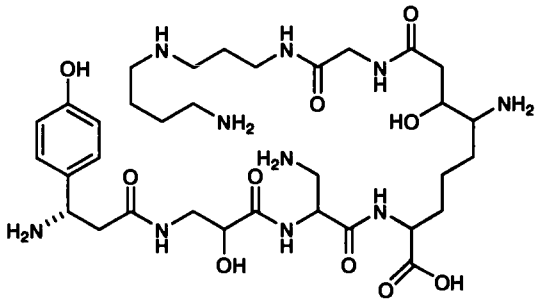
Studies with BlsG, a Putative Arginine 2,3-Aminomutase

Introduction

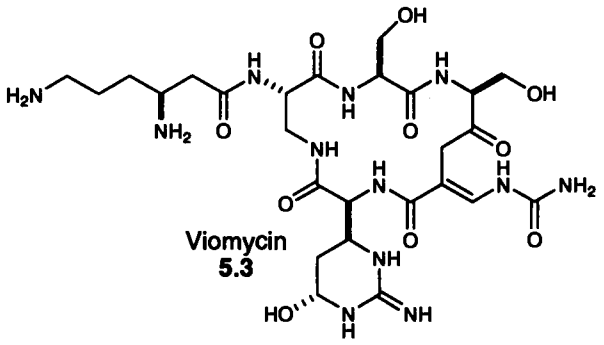
An interesting feature of a number of metabolites, including nucleoside antibiotics, is the presence of β-amino acids. Blasticidin S (**4.1**) is one of these compounds and contains a methylated β-arginine residue.[1] In addition to blasticidin S, β-amino acids are constituents of a number of natural products including jaspamide (**1.1**)[2] and edeine (**5.1**)[3] (β-tyrosine); streptothricins (**5.2**) and viomycin (**5.3**)[4,5] (β-lysine); and pantothenic acid (**5.4**)[6] (β-alanine).



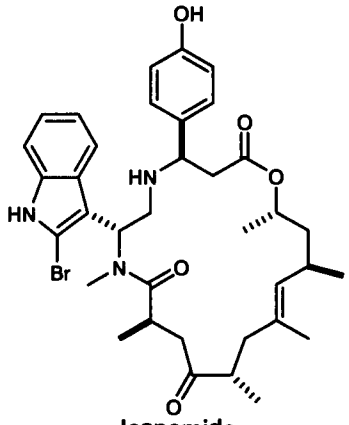
Blasticidin S
4.1



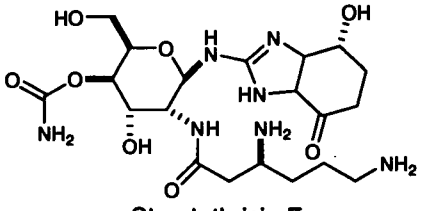
Edeine A
5.1



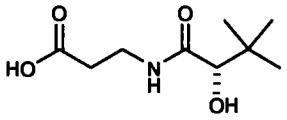
Viomycin
5.3



Jaspamide
1.1



Streptothricin F
5.2



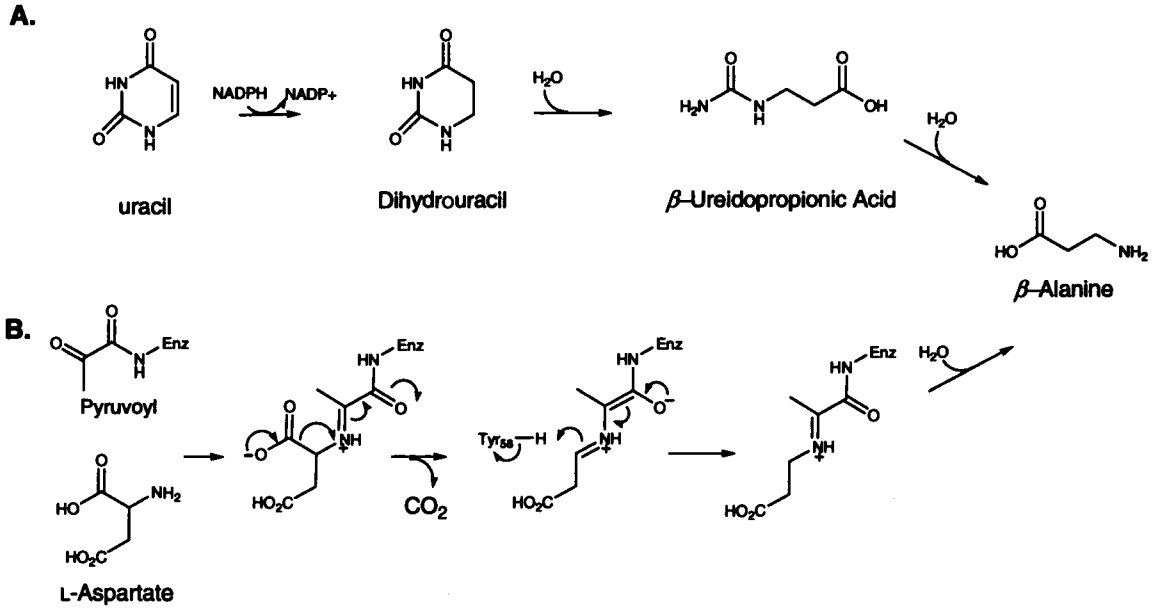
Pantothenic acid
5.4

The modification of the β-arginine residue of blasticidin S, as well as other β-amino acids, presents one avenue to generate new compound analogs that may exhibit altered activity. A good example of the effect of modifying the β-amino acid is found in blasticidin S. Cytomycin, leucylblasticidin S and leucyldemethylblasticidin S are all BS analogs with modified β-arginine moieties. All of these compounds show reduced antibiotic potency, demonstrating the importance of this residue in the biological activity of BS.[7,8] Studying the enzymology involved in the formation of these unusual amino acids may allow for the development of additional BS analogs with altered activity or target specificity. An understanding of the role of β-amino acid forming enzymes in antibiotic biosynthesis may also lead to their use in the design of hybrid gene clusters for the generation of novel antibiotics.

The chemical transformations involved in the formation of β-amino acids have been investigated in a number of biological systems. Despite the common structure motif of β-amino acids, several different mechanisms for their biosynthesis have been elucidated. The formation of β-alanine occurs through the degradation of an appropriate precursor and two mechanisms have been identified. β-alanine can be formed from either the degradation of uracil or through the decarboxylation of aspartic acid. In pyrimidine catabolism, β-alanine is formed from the hydrolysis of uracil which proceeds through a β-ureidopropionic acid intermediate as shown in Scheme 5.1A.[9] An alternate mechanism for the formation of β-alanine is through the decarboxylation of the α-carboxyl group of aspartate by α-aspartate decarboxylase.[6,10] Two forms of α-aspartate decarboxylase have been isolated from *E. coli.* One enzyme has an apparent requirement for pyridoxal phosphate as a cofactor and appears to primarily be a catabolic enzyme.[6,11] The other form of α-aspartate decarboxylase appears to be involved in the biosynthesis of pantothenic acid and requires a covalently bound pyruvate, rather than pyridoxal phosphate as a prosthetic group.[6]

The pyruvoyl α-aspartate decarboxylase has been purified from *E. coli* and the catalytic mechanism has been investigated.[12,13] The proposed mechanism of α-aspartate decarboxylase is shown in Scheme 5.1 B. The enzyme utilizes a pyruvoyl residue that is

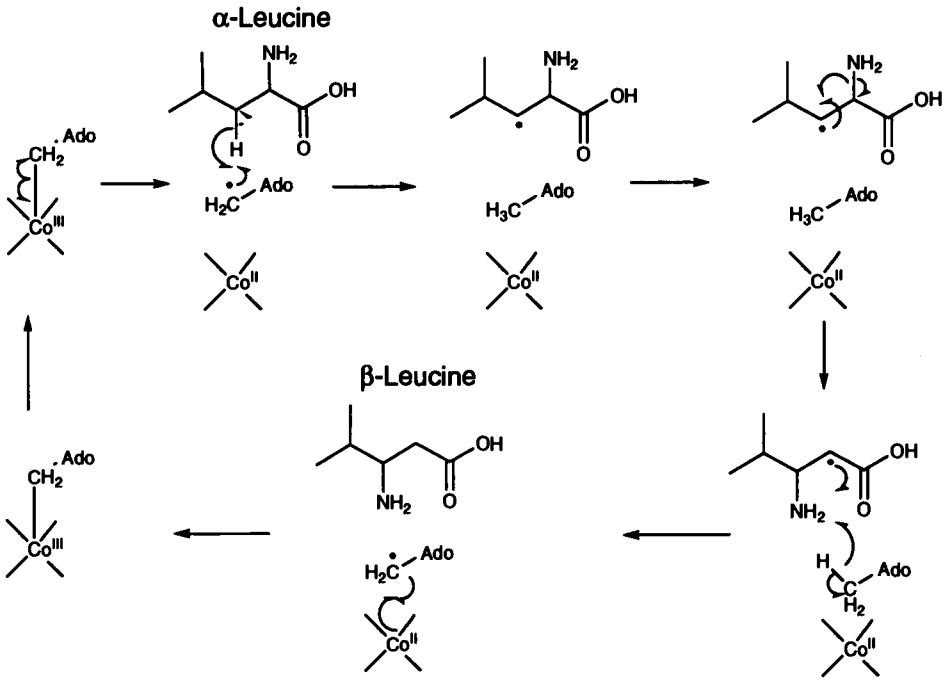
formed from an active site serine. The pyruvoyl residue acts in a manner similar to the pyridoxal phosphate of other amino acid decarboxylases and activates the substrate for decarboxylation through the formation of a Schiff's base. Following the decarboxylation, the resulting enolate is then reprotonated by an active site Tyr residue followed by hydrolysis of the imine to release β-alanine.[12,13] The mechanism for the catabolic α-aspartate decarboxylase utilizes a similar mechanism, however pyridoxal phosphate replaces the pyruvoyl residue as electrophile.



Scheme 5.1. Formation of β-alanine. A. Formation of β-alanine through pyrimidine catabolism. B. Reaction catalyzed by aspartate α-decarboxylase utilizing a modified active site pyruvoyl residue as a Schiff base.

Still another group of aminomutases utilize an adenosylcobalamin (coenzyme B_{12}) dependent mechanism. This group includes leucine 2,3-aminomutase that is involved in leucine catabolism in *Clostridium sporogenes*.[14] Leucine 2,3-aminomutase is perhaps the most wide-spread aminomutase and has been identified in other bacteria, yeast, plants, and animals.[14-18] Studies with human cells determined that the β-leucine pathway also served an anabolic function and was involved in the limited biosynthesis of leucine in humans.[19] B_{12}-dependent aminomutases utilize a radical mechanism that is initiated through a homolytic cleavage of the cobalt-adenosine bond (Scheme 5.2). The resulting 5'-deoxyadenosyl radical subsequently abstracts a hydrogen atom from the

substrate creating a radical at C3 of leucine. Rearrangement of the substrate radical creates an intermediate with the radical centered on C2. Transfer of the hydrogen back from adenosine to the C2 carbon forms the β-amino acid and reforms the 5'-deoxyadenosyl radical which can undergo a single electron transfer back to the cobalt (Scheme 5.2).[20,21]

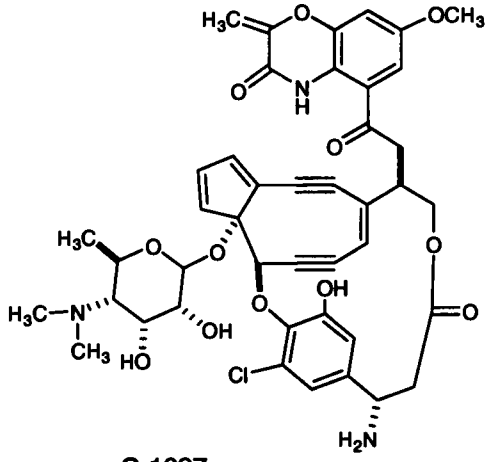


Scheme 5.2. Proposed mechanism of leucine 2,3-aminomutase[20,21]

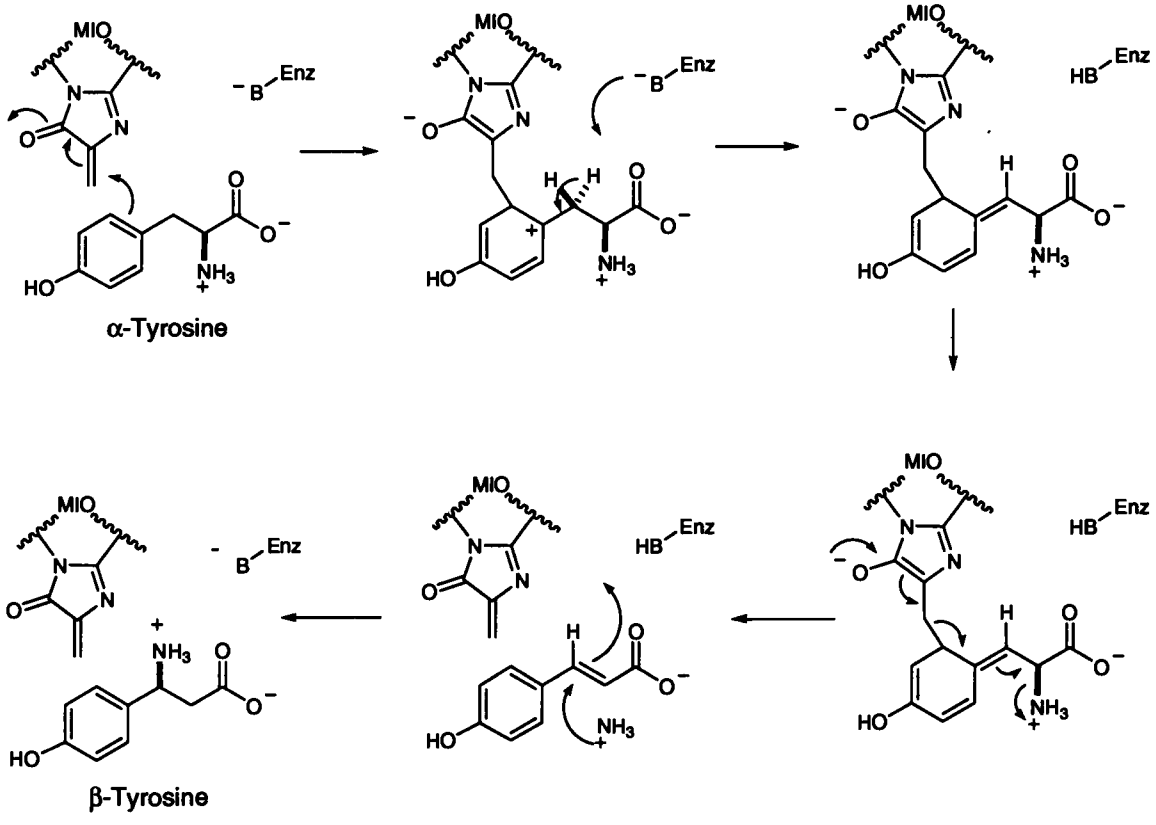
β-alanine and β-leucine are both examples of β-amino acids involved in primary metabolism. An example of a β-amino acid in secondary metabolism is the β-tyrosine residue found in the peptidyl antibiotics, jaspamide and edeine.[3] In *Bacillus brevis* edeine biosynthesis, the formation of β-tyrosine (isotyrosine) from α-tyrosine is catalyzed by tyrosine α,β-mutase and requires ATP, but not pyridoxal phosphate, for catalysis.[3,22] It has been suggested that the reaction mechanism involves the formation of a Schiff's base between the tyrosine amine and an enzyme carbonyl.[3,22] This conversion was found to occur with loss of both the *pro*-3*S* hydrogen and original nitrogen and inversion of configuration at C3 of tyrosine.[22]

A tyrosine aminomutase has been identified and characterized from the enediyne C-1027 (**5.5**) biosynthetic gene cluster. The C-1027 tyrosine aminomutase, SgcC4, uses

an ammonia lyase type mechanism that involves the use of a 4-methylideneimidazole-5-one (MIO). The 2,3 ammonia shift occurs through an initial α,β-elimination reaction followed by Michael addition of the ammonia to the β-carbon (Scheme 5.3).[23]



C-1027
5.5

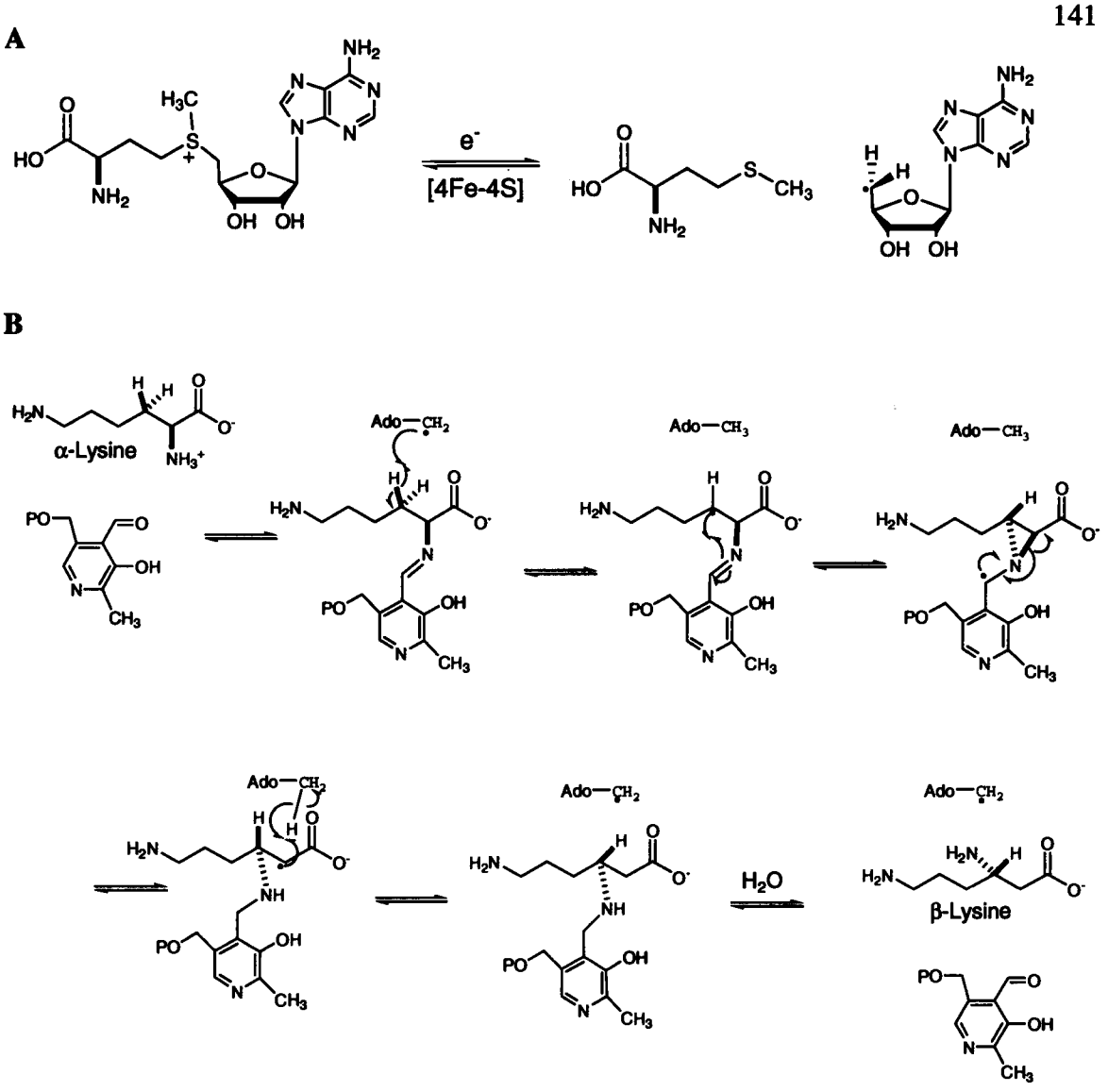


Scheme 5.3. Proposed formation of β-tyrosine by SgcC4.[23]

The last group of aminomutases that will be discussed is also the best studied, and is exemplified by lysine 2,3-aminomutase (LAM). Like leucine 2,3-aminomutase, lysine 2,3-aminomutase was discovered through the investigation of amino acid metabolism.[24,25] Stadtman observed the degradation of lysine to butyric acid, acetic acid and ammonia in *Clostridium* species.[25] Costilow *et al.* later demonstrated that the first step in lysine metabolism was the formation of β-lysine and that this conversion required pyridoxal phosphate as a cofactor.[26] In 1970 LAM was isolated from *C. subterminale* SB4 and shown to catalyze this transformation.[27] Since then, lysine 2,3-aminomutases have been isolated from several bacteria, including *Streptomyces* L-1689-23 and *Bacillus subtilis*.[5,27,28] The genes for several lysine 2,3-aminomutases have been cloned and heterologous expression has permitted thorough characterization of the enzymes.[29,30]

Unlike tyrosine α,β mutase, formation of β-lysine was found to occur through an intramolecular migration of the nitrogen and migration of the *pro*-2*R* hydrogen from C2 with inversion of configuration at both carbons.[5,28] LAM was found to utilize two common enzymatic cofactors, *S*-adenosylmethionine (AdoMet) and pyridoxal phosphate (PLP) in a unique manner.[31]

LAM utilizes AdoMet, PLP, and an iron-sulfur cluster to catalyze the formation of L-β-lysine through a radical mechanism as shown in Scheme 5.4.[21] Prior to substrate binding, PLP is bound to the ε-nitrogen of an active site lysine residue.[32] Upon amine exchange, the substrate is activated for catalysis through the formation of the corresponding substrate-aldimine with PLP. Catalysis is then believed to be initiated by the formation of a 5'deoxyadenosyl radical generated by the [4Fe-4S] iron-sulfur cluster (Scheme 5.4, A).[33] The 5'-deoxyadenosyl radical abstracts the *pro*-3*R* hydrogen from lysine generating the substrate radical (Scheme 5.4,B). Subsequent radical rearrangement forms an aziridine intermediate with the radical located on C4' of PLP.[34] A third radical species results from the homolytic cleavage of the aziridine at the nitrogen-C2 bond generating the β-lysine –PLP intermediate with a radical centered on the substrate C2 carbon. Hydrogen abstraction from 5'-deoxyadenosine reforms the adenosyl radical. Finally, cleavage of the aldimine releases the β-lysine product and reforming the enzyme-PLP aldimine.



Scheme 5.4. Radical catalyzed mechanism of Lysine 2,3-aminomutase.[20,21]
A.Formation of a 5'deoxyadenosyl radical by the [4Fe-4S] iron-sulfur cluster.
B. Radical mediated formation of β-lysine.

Formation of β-arginine in BS Biosynthesis

The formation of the β-arginine moiety of blasticidin S has been of interest from the first biosynthetic studies.[1] The earliest studies on BS biosynthesis by Seto *et al.* demonstrated that L-α-arginine is a precursor to the β-arginine moiety of BS. Seto was able to show 30% incorporation of uniformly labeled L-arginine into BS.[1] Later, studies by Prabhakaran were able to show the direct incorporation of deuterated β-arginine into

BS with 48% incorporation.[35] Additional studies by Prabhakaran *et al.* used labeled precursors to study the mechanism of β-arginine formation.[36] Feeding of DL-[3-^{13}C, 2-^{15}N] arginine demonstrated that the incorporation of L-α-arginine into β-arginine occurs through an intramolecular reaction with a shift of the C2 nitrogen to the C3 position. It was also shown that the amino migration occurred with inversion of configuration at C3 (Scheme 5.5). Additional studies with deuterium labeled arginine precursors were able to show that the *pro*-3*R* hydrogen of α-arginine had migrated to the C2 position during the amino group migration. The results of these incorporation experiments were consistent with a reaction mechanism that is analogous to the formation of β-lysine by lysine 2,3-aminomutase.[36] Despite the evidence for the involvement of an arginine 2,3-aminomutase (RAM) in the biosynthesis of BS, this enzymatic activity could not been detected in crude cell-free extracts of *S. griseochromogenes.*[35]

Scheme 5.5. Summary of arginine 2,3-aminomutase labeling studies.[36]

The preliminary results provided evidence for the presence of an arginine 2,3-aminomutase, much like the *C. subterminale* lysine 2,3-aminomutase, in the biosynthetic pathway of BS. Additional support for the involvement of a LAM type enzyme was the presence of an ORF, *blsG*, in the BS gene cluster with a product predicted to have 48% sequence identity and 66% similarity with lysine 2,3 aminomutase (Figure 5.1).[37] *blsG* codes for a 410 amino acid protein with a calculated molecular weight of 47.2 kDa. The putative ribosomal binding site (GGACCGG) for *blsG* is located seven nucleotides upstream of the translational start site and is in good agreement with known *Streptomyces* ribosomal binding sites.[38,39]

```
BlsG   MSHMSTESDGIRPSLTRREDIPDEQWNDWRWHMRKRITNLDKAREWIRPTPLEEKAIAET 60
KAM    ---------MINRRYELFKDVSDADWNDWRWQVRNRIETVEELKKYIPLTKEEEEGVAQC 51

BlsG   AGKYRWSVTPYYASLMDPDDPGCPVRQQAVPALGELMEFSGAEVDPVGDMYYRRTNRVVH 120
KAM    VKSLRMAITPYYLSLIDPNDPNDPVRKQAIPTALELNKAAADLEDPLHEDTDSPVPGLTH 111

BlsG   KYPDRVIMLITEACPVYCRHCTRKFHTTDVDGTYFERNEGEDFSEDLRYIADHPEIRDVL 180
KAM    RYPDRVLLLITDMCSMYCRHCTRRRFAGQSD---DS-MPMERIDKAIDYIRNTPQVRDVL 167

BlsG   LTGGDPLSYRDGKLEEIIAGLRAIPSVEIIRIGSRFPVLLPQRVTPELCEMLARYHPVWL 240
KAM    LSGGDALLVSDETLEYIIAKLREIPHVEIVRIGSRTPVVLPQRITPELVNMLKKYHPVWL 227

BlsG   NTHFNHPKEITPESERAIDRLLRHGIPVGNQTVLLRGINDDLGTMRRLMTELLRIRVRPY 300
KAM    NTHFNHPNEITEESTRACQLLADAGVPLGNQSVLLRGVNDCVHVMKELVNKLVKIRVRPY 287

BlsG   YLYHCDNVTGVSHFMTSVEKGWEIMEGLQGHITGFGVPQYVLTTRLG---KIPMVRPYYR 357
KAM    YIYQCDLSLGLEHFRTPVSKGIEIIEGLRGHTSGYCVPTFVVDAPGGGGKTPVMPNYVIS 347

BlsG   ETPDGLVLRNYRGEEMLVDDSVCPLTESAAAHAFRNAPDVTENRTTATGEGAR------- 410
KAM    QSHDKVILRNFEGVITTYSEPIN-YTPGCNCDVCTGKKKVHKVGVAGLLNGEGMALEPVG 406

KAM    LERNKRHVQE 416
```

Figure 5.1. Sequence alignment of *S. griseochromogenes* BlsG with *C. subterminale* LAM. Identical residues are highlighted in dark grey, similar residues in light grey

PSI-BLAST analysis confirmed that BlsG is a member of the radical AdoMet superfamily of proteins (CD: pfam04055.6) that includes lysine 2,3-aminomutase, biotin synthase, and lipoic acid synthase, among others.[40] Among the conserved features within this group of proteins, BlsG shares the conserved CxxxCxxC believed to be involved in binding the iron-sulfur cluster. This motif is highly conserved among all of the AdoMet superfamily of proteins.[40] BlsG also contains the glycine rich region TGGDP, that has been proposed to be involved with AdoMet binding as it is similar to the AdoMet binding site in methyltransferases.[40] Another important catalytic residue, corresponding to BlsG lysine 348, has been shown to form a protein-PLP aldimine in LAM.[32]

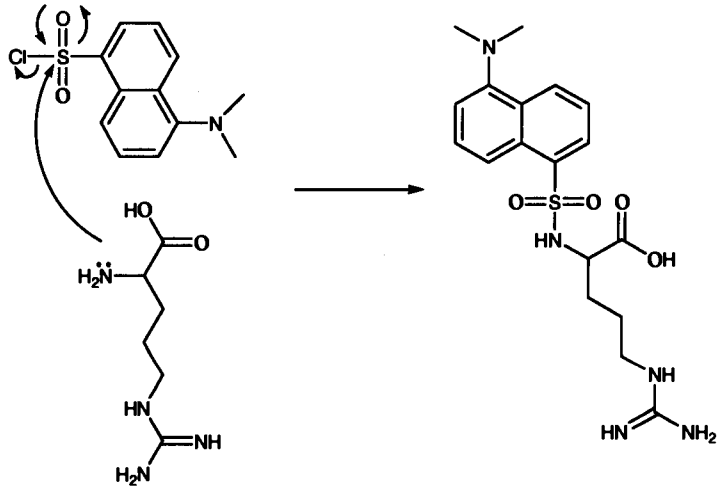
Although there has been a great deal of interest in the formation of β-arginine by *S. griseochromogenes*, early efforts to detect and purify an arginine 2,3-aminomutase from the CFE of *S. griseochromogens* were unsuccessful.[35] The identification of *blsG* within the BS gene cluster presented an opportunity to study the formation of β-arginine as well as gain additional insights into the biosynthesis of blasticidin S. The functional heterologous expression of *blsG* should allow for the production of adequate amounts of

enzyme to permit complete characterization. The objective of the current studies was to characterize the enzyme and enzymatic mechanisms involved in the formation of β-arginine in order to provide insights into biosynthesis of BS as well as the metabolism of β-amino acids.

Results and Discussion

Separation and Detection of L-α-Arginine and β-Arginine

Active BlsG is expected to produce β-arginine when supplied with L-α-arginine. In order to detect BlsG activity both *in vitro* and *in vivo*, a method was developed for the separation and detection of arginine and β-arginine by C_{18}-RP HPLC.[41,42] The arginine from BlsG assays was purified by cation exchange chromatography and dansylated to allow for the detection and separation by HPLC (Scheme 5.6). Baseline separation of α and β-arginine was achieved (Figure 5.2). This method was used to analyze the BlsG activity studies in both *E. coli* and *S. lividans* described below.



Scheme 5.6. Dansylation of arginine for UV detection.

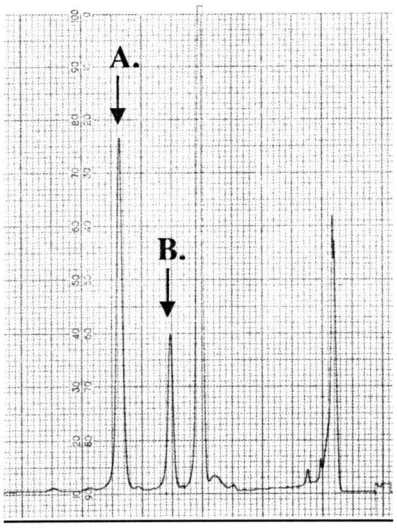


Figure 5.2. HPLC separation and detection of dansylated L-α and β-arginine. A. dansyl-L-α-arginine. B. dansyl-β-arginine.

Heterologous Expression and Purification of BlsG in E. coli

BlsG is the first protein from the BS gene cluster that we attempted to characterize. Our original attempts at characterization involved the heterolgous expression of *blsG* in *E. coli*. A number of expression constructs were generated to produce BlsG in the native form as well as with *N* and *C*-terminal His_6-tag fusions.

The original expression construct for BlsG was in a pET28a+ expression system. At that time, the sequencing of the entire BS cluster had not been completed and there was some uncertainty about the location of the correct start codon. *blsG* contains two methionine residues at the beginning of the open reading frame. The two possible start sites are separated by six nucleotides. Since the correct start site was not known, two expression constructs, one for each proposed start site, were generated, p28*bls*G1 and p28*bls*G2. Heterologous expression of p28*bls*G1 and p28*bls*G2 in *E. coli* BL21 (DE3) resulted in the production of large quantities of soluble protein that was purified by nickel affinity chromatography (purification of BlsG1 is shown in Figure 5.3). Purified enzyme was used to test for RAM activity. When *in vitro* formation of β-arginine was not detected, the *E. coli* culture medium and cellular extracts were analyzed for the formation of β-arginine. β-arginine could not be detected in any component of the *E. coli*/pET28blsG expression system.

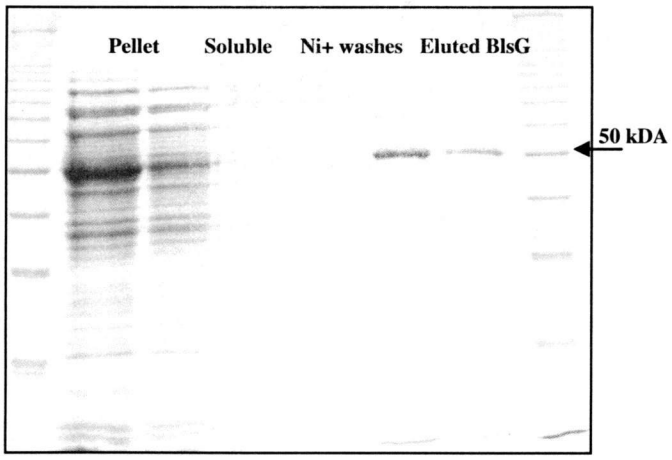


Figure 5.3. SDS-PAGE showing expression and purification of p28*blsG*1 in *E. coli*. BlsG was overexpressed in *E.coli* and soluble BlsG was purified from the cleared cell lysate using Ni^{+2} affinity chromatography. The approximately 50 kD purified BlsG is indicated by an arrow.

One possible reason for apparently inactive BlsG was may have been related to codon usage in *S. griseochromogenes* versus that in *E. coli*. *E. coli* contains low levels of the tRNAs for arginine AGG/AGA codons and the proline CCC codon. For example, in *E. coli* K12 these codons only occur at a rate 1.2 codons /1000 for AGG of and 5.5 codons/1000 for CCC.[43] Unlike *E. coli*, these codons are more common in G+C rich organisms such as *S. griseochromogenes*, and *blsG* contains three AGG codons and thirteen CCC codons. This corresponds to an occurrence rate of 7.3 and 31.6 codons/1000, respectively. The heterologous expression of genes containing these rare codons can be difficult because *E. coli* does not contain high levels of the appropriate tRNAs.[44] Coexpression of the genes coding for these rare tRNAs (*argU* and *proL*) can aid in the production of functional protein in these systems.[45] The ease of obtaining soluble BlsG from *E. coli* suggested that there was not a problem with codon usage, however if there was not adequate amounts of the appropriate tRNAs, an unintentional mutation or amino acid substitution may have occurred. Translational mutation has been reported in the expression of bovine placental hormone (BPL), which contains rare AGG codons, in *E. coli*.[44] A similar situation was found with the lysine aminomutase gene (*kamA*) from *C. subterminale*, which contains the rare arginine codon AGA.[30] In this system, coexpression of the *argU* gene, coding for this rare tRNA, was required for

overexpression and coexpression of *argU* with *kamA* resulted in a six fold increase in the specific activity of the purified enzyme.[30] For this reason, subsequent expression studies in *E. coli* were conducted in *E. coli* BL21-Codon Plus(DE3)-RP cells (Stratagene). This cell line contains extra copies of the *argU* and *proL* genes that code for tRNAs that recognize rare *E. coli* arginine and proline codons.

Another possible reason for not being able to detect arginine 2,3-aminomutase activity was thought to be the presence of an *N*-terminal His_6-tag on BlsG. It is possible that the presence of the additional amino acids may interfere with the correct folding and/or catalytic activity of the protein. For this reason, several additional expression constructs were generated that resulted in the production of either native (p20*blsG*nat) or *C*-terminal His_6-tagged fusions (p20*blsG*CHis). Upon complete sequencing of the BS gene cluster it was determined that the first of the two start codons was the most likely translational start site due to proximity to a putative ribosome binding site. For this reason, only the expression constructs containing the first start site were used in the activity assays. Both the native and the *C*-terminal fusion proteins were solubly expressed in *E. coli* Codon Plus-(DE3)-RP and the cell-free extracts and/or purified protein were evaluated for the ability to convert L-α-arginine to β-arginine.

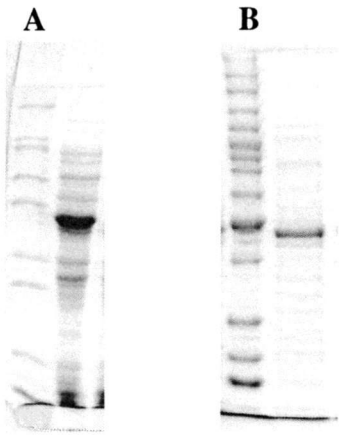


Figure 5.4. Expression of pET20*blsG* constructs in *E. coli*. A. Overexpression of *blsG* with a C-terminal His_6-tag, soluble protein prior to purification. B. Overexpression of BlsG as the native protein, total soluble protein.

Initial assays were conducted in an aerobic environment; however we were concerned that if BlsG does contain an iron-sulfur center, it may be oxidized to an inactive form. Frey and colleagues have identified and characterized several lysine 2,3-

aminomutases from both anaerobic and aerobic bacteria.[29,30] They found that the mutase isolated from an anaerobic bacteria, *C. subterminale*, was extremely sensitive to oxygen and purification under an oxygen free atmosphere was required for optimal enzyme activity.[30] Conversely, it was found that a lysine 2,3-aminomutase isolated from the aerobic bacterium *Bacillus subtilis* was not as sensitive to oxygen and could be isolated under normal laboratory conditions and active enzyme could be generated through an initial reductive preincubation of the enzyme.[29] *blsG* was cloned from an aerobic bacterium and as such was expected to be less sensitive to oxygen. The fermentation of *E. coli* heterologously expressing *blsG* was conducted in an aerobic environment. After our initial attempts to detect aminomutase activity were unsuccessful a reductive preincubation step was introduced into the assay.

In order to prevent oxidation of the iron-sulfur cluster, and to ensure the correct oxidation state of the enzyme, BlsG was incubated under anaerobic, reducing conditions prior to being used in the *in vitro* assays. The preincubation conditions included cysteine, ferrous ammonium sulfate, and pyridoxal phosphate according to the method of Chen *et al.*[29] An anaerobic chamber like the one used in Chen's experiments was not available so the preincublation was conducted in sealed glass ampules under an argon atmosphere. The preincubated enzyme was then used in the *in vitro* assays, also under an argon atmosphere. The reductive preincubation of BlsG did not have an effect on our ability to detect BlsG activity under any of the assay conditions.

Heterologous Expression of BlsG in S. lividans

The inability to detect arginine 2,3-aminomutase activity with protein expressed in *E. coli* lead us to speculate that the protein may not be folding properly in *E. coli*. It was felt that the protein would be more likely to fold properly and be catalytically competent if it was expressed in another *Streptomyces*. *S. lividans* had been previously used to heterologously express the entire BS gene cluster, which necessarily required the functional expression of *blsG*.[46] For this reason, we attempted to express *blsG* in *S. lividans*. *blsG* was cloned into the pXY200 expression vector for expression off of the P_{tipA} thiostrepton inducible promoter. pXY200 is an *E. coli/Streptomyces* shuttle

expression vector developed by Dr. Xihou Yin in our laboratory (unpublished data). The cloning sites used resulted in the addition of an *N*-terminal His_6-tag that was later used for protein purification. pXY2*blsG*, as well as empty vector, pXY200, were introduced into *S. lividans* by protoplast transformation.

The *S. lividans*/pXY2*blsG* expression system does not produce the high levels of protein like those commonly used with *E. coli*, however the production of BlsG was observed and confirmed through detection of the *N*-terminal His_6-tag by Western blot analysis of the purified protein (Figure 5.5). The culture broth and cell lysate of *S. lividans*/pXY2*blsG* were analyzed as described for the *E. coli* system but β-arginine was not detected in either. *In vitro* assays using both cell-free extracts and purified BlsG from *S. lividans* were also unfruitful and no β-arginine production could be detected.

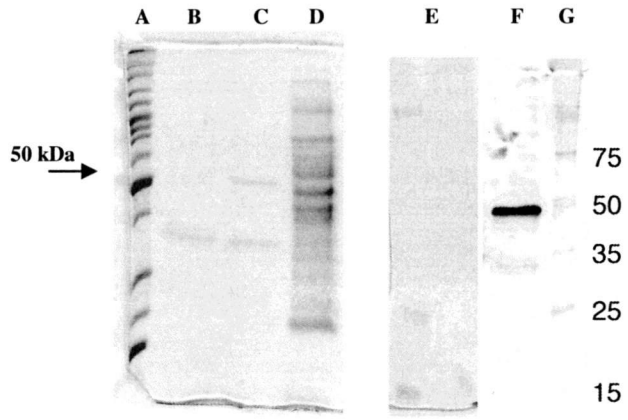


Figure 5.5. Expression of pXY2BlsG in *S. lividans*. All samples were concentrated 10x prior to analysis due to the low level of expression seen in this system. **A-D**. SDS-PAGE showing the purification of BlsG from *S. lividans* under denaturing conditions. **A**. Protein standards, **B**. pXY200 control expression. **C.** Purified BlsG from pXY2*blsG* expression. **D**. Ni^{+2} column washes. The location of the 50kDa BlsG band is indicated. **E-G**. Western blot detection of BlsG from *S. lividans*. **E.** SDS-polyacrylamide gel of protein used for Western blot. **F.** Western detection of BlsG with antibody for the His_6-tag was used to detect BlsG. **G**. MagicMark™Western Protein ladder.

blsG Gene Disruption Experiment.

The inability to detect arginine 2,3-aminomutase activity with our recombinant enzyme lead us to devise an alternate plan to demonstrate the function of BlsG. One common method to demonstrate the role of a biosynthetic enzyme or pathway is to

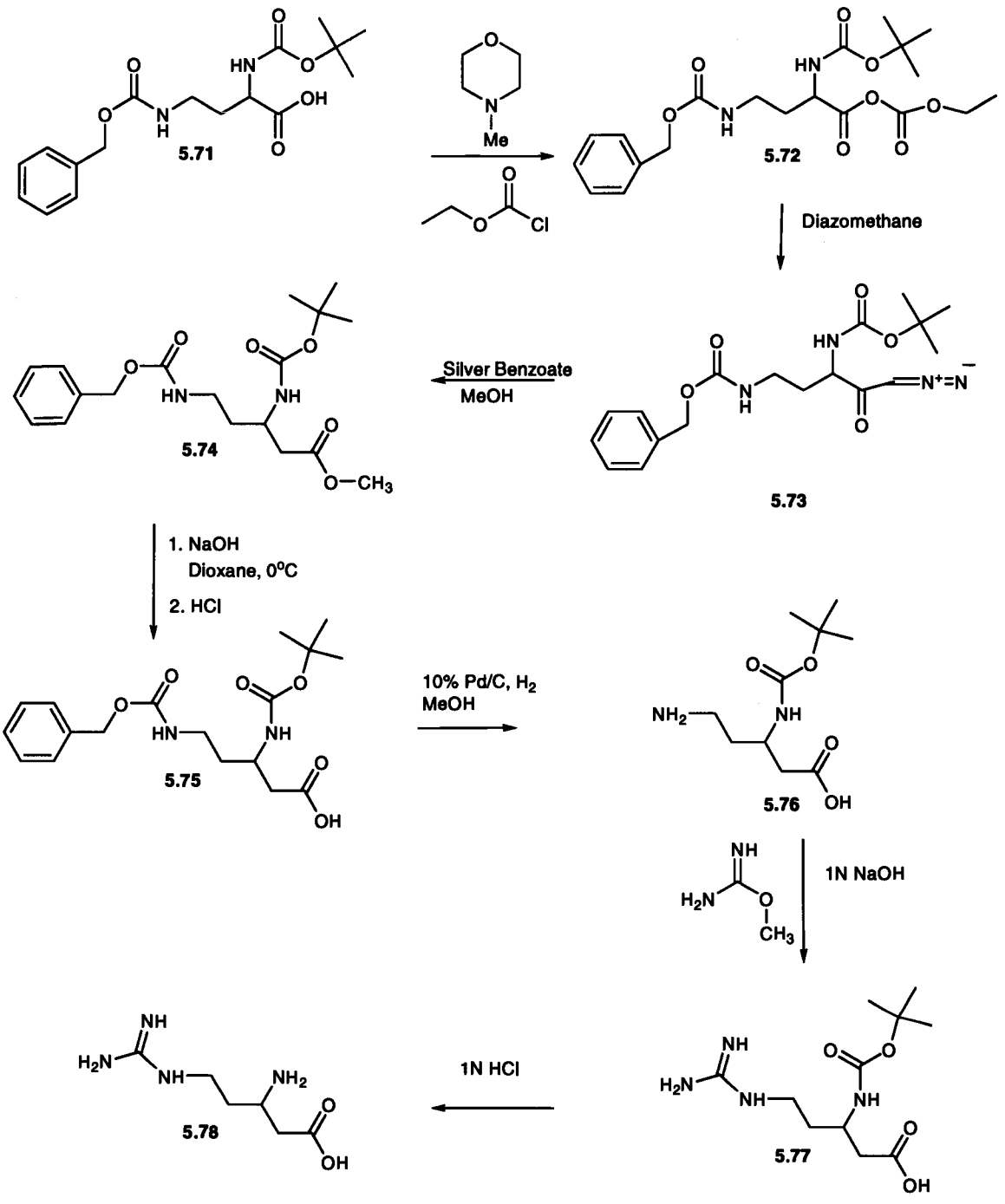
disrupt the gene or gene cluster in the producing organism. Gene disruption will often result in the accumulation of metabolites that occur prior to the step affected by the disruption in the biosynthetic pathway. Analysis of the accumulated intermediate(s) can often provide insight into the function of the disrupted gene. Complementation of the disrupted gene *in trans* can then be used to restore metabolite production and confirm the role of that particular gene product in the biosynthetic pathway and that the observed lack of metabolite production was solely due to the disruption of a single gene. If the function of the disrupted gene product can be predicted, addition of the proposed intermediate can also be used to restore metabolite production and confirm the role of the gene product in biosynthesis.

Synthesis of β-Arginine

The β-arginine standard used in the HPLC analysis of *in vitro* RAM assays was material previously prepared in our lab. While there was an adequate amount for use as a HPLC standard, quantity was not sufficient for the supplementation of culture media as would be required for gene disruption experiments. Although the synthesis of cytosinine was reported in 1972,[47] little attention had been give to the total synthesis of blasticidin S. In recent years two reports on the synthesis of the blastidic acid moiety of BS have been reported.[48,49] Our synthesis of β-arginine was based on the first synthesis of blastidic acid by Nomoto and Shimoyama.[48] The method was modified slightly as the Nomoto synthesis resulted in the formation of blastidic acid (δ-*N*-methyl β-arginine). To prepare β-arginine, the methylation of the δ-nitrogen with methyl iodide was omitted. Other minor modifications were also made, including the method of diazomethane generation and the purification of several intermediates.[50] Our method utilized commercially available *t*-butyloxycarbonyl (BOC) and benzyloxycarbonyl (CBZ) protected diaminobutyric acid (BOC-δ-CBZ-diaminobutyric acid), **5.71**, as a starting material and excluded several otherwise required protection steps. The first step in the synthesis was the formation of *N*-BOC-δ-CBZ-β-ornithine (**5.75**) through the net addition of one carbon to 2,4-diaminobutyric acid (Scheme 5.7). This transformation required four steps through a modified Ardnt-Eistert reaction.[51,52] The first step is the formation of the protected 2,4-

diaminobutyric ethoxyanhydride (**5.72**) through the addition of chloroformate in the presence of an amine base, *N*-methylmorpholine. The resulting product, **5.72,** is immediately treated with diazomethane to form diazoketone **5.73.**[50] A Wolf rearrangement was conducted in methanol to yield the methyl ester of β-ornithine **5.74** in 73% overall yield from **5.71**. Following purification, saponification of the methyl ester resulted in diprotected β-ornithine **5.75** in 92% yield from **5.74**.

The second step in the formation of β-arginine was the guanidination of the β-ornithine derivative.[53] First, the δ-CBZ protecting group was selectively removed from **5.75** by catalytic hydrogenolysis to give **5.76** in an 87% yield. *N*-BOC-β-arginine was formed by reaction of **5.76** with *O*-methylisourea under basic conditions. The acidic work-up of the reaction resulted in the cleavage of the BOC protecting group from **5.77** and resulted in the formation of deprotected β-arginine **5.78** in 82% yield from **5.76**. The formation of **5.78** was confirmed by NMR, mass spectral and HPLC analysis.



Scheme 5.7. Synthesis of β-arginine

Transformation of S. griseochromogenes

Our next approach to demonstrate the function of BlsG was to conduct a gene disruption experiment in the BS producer *S. griseochromogenes*. Gene disruption experiments require the ability to introduce foreign DNA into an organism. Transformation of *S. griseochromogenes* had not previously been demonstrated and so the initial goal was to develop a method for the introduction of DNA into this organism.

Protoplast transformation is a method commonly used to introduce DNA into Gram-positive bacteria such as *Streptomyces*. The method involves digestion of the external cell wall, thus reducing the physical barrier to DNA binding and penetration. The initial transformation method used for *S. griseochromogenes* is commonly used with other *Streptomyces* species in our laboratory and is based on a protocol developed by Hopwood.[54] Many factors affect the ability to transform any given *Streptomyces* species, and extensive optimization of protoplast formation, regeneration and transformation conditions is often required.[55,56] Our initial work with *S. griseochromogenes* involved development of optimal conditions for the formation and regeneration of protoplasts. Protoplasts could easily be formed from *S. griseochromogenes* mycelium, provided it was grown in the presence of glycine and the culture was less than 72 h old. Glycine can interfere with the formation of the cell wall and has been shown to improve the ability to form protoplasts in *Streptomyces*.[57] In several bacteria, the mechanism by which glycine disrupts cell wall formation has been shown to be through the substitution of glycine for D-alanine in peptidoglycian, thus disrupting crosslinking in the cell wall.[58] Other factors shown to affect the ability to form protoplasts in *Streptomyces* include media composition and culture age and conditions must be empirically determined for each species.[56] Optimal protoplast formation was observed when *S. griseochromogenes* was grown in TSB medium containing 0.5% glycine for 30-40 h at 30 °C. Protoplast formation by lysozyme digestion was comparable to the formation of *S. lividans* protoplasts and was generally complete after 30-40 min incubation at 30 °C.

The optimal conditions for protoplast cell wall regeneration is another variable that must be independently determined for each *Streptomyces* species. Once formed, the *S. griseochromogens* protoplasts were quite fragile and excessive care was needed to prevent cell lysis. As is typical for protoplast regeneration, we found that it was

necessary to plate protoplasts on osmotically neutral agar such as R2YE.[54] The ability of protoplasts to regenerate was demonstrated by subjecting them to transformation conditions following plating onto R2YE agar without antibiotic selection. In addition to R2YE, it was found that *S. griseochromogenes* protoplasts could regenerate on HT and YME agar plates that were dried for 2.5 h in the sterile hood according to the preparation method for R2YE plates.[54]

Many *Streptomyces* species contain a restriction system that specifically cleaves methylated DNA.[54] When working with these organisms it is necessary to prepare plasmids from an *E. coli* strain deficient in Dam ($GA^{Me}TC$) and other methylation systems such as *E. coli* INV110 (Invitrogen) or ET12567/pUZ8002.[54] Because it is not known if *S. griseochromogenes* contains such a restriction system, *E. coli* INV110 cells that are deficient in DNA methylation were used to prepare plasmid DNA for the protoplast transformation experiments.

Multiple attempts were made to transform *S. griseochromogenes* through the formation of protoplasts. When the initial attempts at transforming *S. griseochromogenes* failed using the basic *S. lividans* method, attempts were made to optimize the transformation protocol. One possibility for our inability to generate *S. griseochromogenes* transformants was thought to be the use of antibiotic concentrations that were too high. The sensitivity of different *Streptomyces* strains to a given antibiotic varies and must be optimized for individual species.[54] The majority of cloning vectors used in this work contain an apramycin *accC4* aminoglycoside acetyltransferase gene as a resistance marker. The sensitivity of *S. griseochromogenes* to apramycin was not known and had to be determined. Spores of *S. griseochromogenes* were grown on plates containing 2.5-40 μg/mL apramycin. It was found that 20 μg/mL apramycin provided efficient selection and this concentration was used in all subsequent experiments.

Another aspect of the transformation protocol that was varied was the amount of polyethylene glycol (PEG) used to induce transformation. If the concentration of PEG is too high, it may be detrimental to the cells. The standard *S. lividans* protocol uses a 25% PEG solution. Experiments were conducted where PEG levels in the standard transformation protocol were varied from 5-25%. In all cases, protoplasts were able to

regenerate when antibiotic selection was not applied. This indicated that the cells were not being killed by the PEG. However, successful transformation of *S. griseochromogenes* was still not accomplished under any of the experimental conditions.

Transformation of S. griseochromogenes by Conjugation.

We next attempted to transform *S. griseochromogenes* through intergeneric conjugation with *E. coli.* This is another method of transformation that is increasing in usage in *Streptomyces sp.* and was first reported in 1989 by Mazodier.[59] This method requires the use of a vector system that includes an origin of transfer (*oriT)* site that is necessary for the transfer of DNA between bacteria. Intergeneric conjugation requires the initial introduction of the desired DNA into a donor bacterium containing a plasmid mobilization factor, RP4, and is capable of conjugation. Under the proper conditions, the *E.coli* donor can directly transfer the vector DNA to the *Streptomyces* recipient. Use of an appropriate antibiotic selection marker on the transferred plasmid allows for the identification of exconjugates.

The conjugation protocol used in experiments with *S. griseochromogenes* was based on the method of Flett.[54,60] The *E.coli-Streptomyces* shuttle vector pKC1139 was used as a test plasmid for the conjugation experiments and contains the necessary *oriT.* The RP4 mobilization elements were provided by an RP4 plasmid derivative integrated into the donor *E. coli* S17-1 chromosome.[54] Initial conjugation experiments did not produce any *S. griseochromogenes* exconjugates, and it was observed that control cultures, grown in the absence of apramycin selection, had very poor growth on the AS1-agar conjugation medium. In order to improve *S. griseochromogenes* growth, additional conjugation experiments were tried where the matings were plated on either HT or YME agar, as *S. griseochromogenes* was found to grow well on these media. Formation of exconjugates not was observed under any of the conjugation conditions.

Heterologous Disruption of blsG

An alternative to disrupting *blsG* in the producing organism is to disrupt *blsG* on the cosmid containing the blasticidin S gene cluster, cos9, followed by heterologous

expression in *S. lividans*. Cos9 had previously been shown to successfully direct leucylblasticidin S biosynthesis when introduced to *S. lividans*.[46] If BlsG is responsible for the production of β-arginine, gene disruption should abolish the production of leucylblasticidin S. It should then be possible to restore leucylblasticidin S production by adding β-arginine to the production media or by complementing the disrupted gene *in trans* with a functional copy of *blsG*.

A variety of expression constructs were generated with the goal of disrupting the *blsG* gene on a cosmid containing the BS gene cluster and heterologously expressing the construct in *Streptomyces lividans* in parallel to the intact gene cluster. The first constructs for the heterologous disruption of *blsG* included the insertion of a hygromycin resistance cassette into the gene. Although two overlapping *Bcl*I restriction sites were conveniently located in the center of *bls*G, the presence of other *Bcl*I sites in the parent cosmid, cos9, prevented the direct insertion of the resistance cassette. The parent construct for the disruption of BlsG was generated as indicated in Figrue 5.6. First, cos9 was digested with *Eco*RV and *Bsr*GI and the 4.5 kb fragment was purified. End repair of the restricted DNA and ligation into *Eco*RV-digested pBluescript (pBKS) resulted in the regeneration of the original restriction sites. The resulting plasmid was digested with *Bcl*I and ligated to *Bam*HI restricted hygromycin resistance cassette that contains the *hyg* gene for hygromycin phosphotransferase. Following sequence confirmation that the construct was correct, the *blsG* disrupted fragment was excised from pBKS with the *EcoRV* and *BsrGI* and ligated to similarly digested and purified cos9.

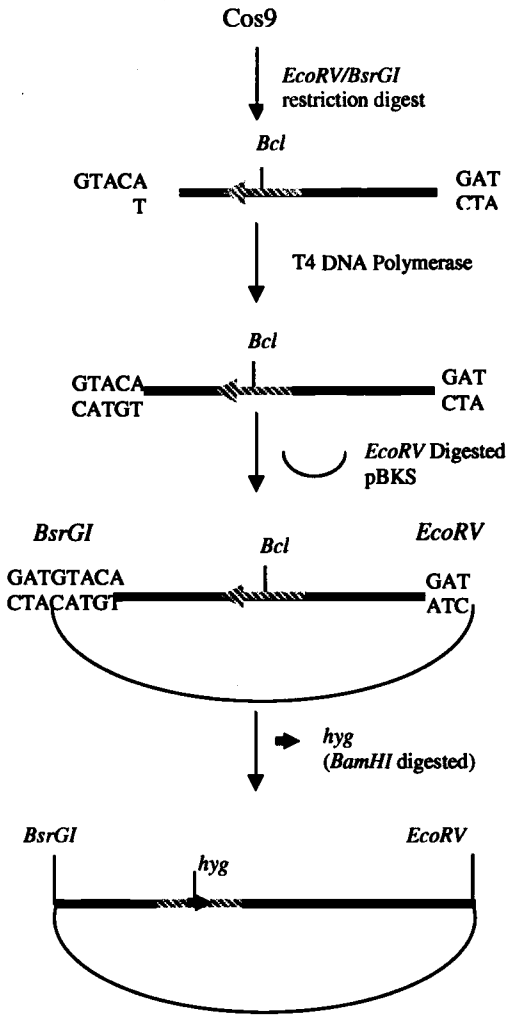


Figure 5.6. Construction of a *blsG* disruption on cos9.

Upon purification and digestion of the resulting cosmid, cos9*blsG*Hyg, it was discovered that three *Bam*HI fragments were missing from the construct, 4.9 kb, 1.7 kb and 2.0 kb (Figure 5.7). Analysis of the new restriction patterns, compared to the location of the altered fragments in the original blasticidin gene cluster, revealed that an approximately 8.5 kb fragment was missing from one side of the cluster. This loss of the cosmid fragment could only be explained by the presence of an additional *Bsr*GI or *Eco*RV restriction site on cos9. The cosmid vector, pOJ446, used to construct cos9, was a gift from Sir David Hopwood (John Innes Centre) and the complete nucleotide sequence of the vector is not available. In order to determine the identity and nature of

the additional restriction site, cos9 was digested independently with both *Eco*RV and *Bsr*GI. Restriction analysis revealed the presence of a single *Bsr*GI site on the cloning vector.

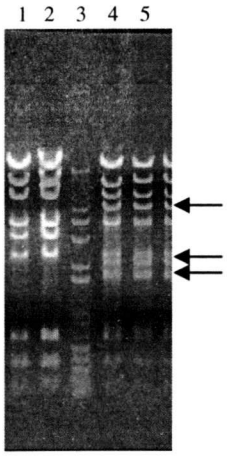


Figure 5.7. Cos9*bls*GHyg constructs showing the alterations in *Bam*HI restriction patterns. Arrows indicate the location of the absent bands corresponding to the missing 8kb fragment Lanes 1,2 *Bam*HI digested Cos9*bls*GHyg, Lane 3: 1kb+ DNA ladder Lane 4: *Bam*HI digested Cos9, Lane 5: *Bam*HI digested pXY4*bls*G.

In order to complete the construction of an insertional deletion of *bls*G in cos9, the missing *Bsr*GI fragment needed to be replaced. The missing 8.5 kb fragment was excised and purified from cos9 and then ligated to cos9*bls*GHyg. The resulting cosmids were initially screened for the presence of the missing fragment by *Bam*HI digestion, (Figure 5.8A). One of the resulting cosmids, cos9RH, was found to contain the missing bands, indicating the successful ligation of the *Bsr*GI fragment (Figure5.8 A). This construct was then screened by *ScaI* restriction analysis in order to determine if the *Bsr*GI fragment had inserted in the proper orientation (Figure 5.8 B). Cos9 only contains two *ScaI* restriction sites, one site is located on the missing *Bsr*GI fragment, while the other site is downstream of the inserted hygromycin cassette. Thus, the size of the *ScaI* fragment can be used to determine if the 8.5 kb fragment was ligated in the proper orientation. Digestion of the parent cosmid, cos9, with *Sca*I results in a 3.3 kb fragment, while digestion of the correctly disrupted cos9RH results in a 5.5 kb fragment (3.3 kb + 2.2 kb hygromycin cassette). On the other hand, if the *Bsr*GI fragment had inserted in the improper direction, the *ScaI* fragment would be over 11 kb. The cosmid indicated,

cos9RH, in Figure 5.8B was found to have the proper *Sca*I restriction fragment. The correct orientation and final construction of cos9RH was confirmed by sequencing.

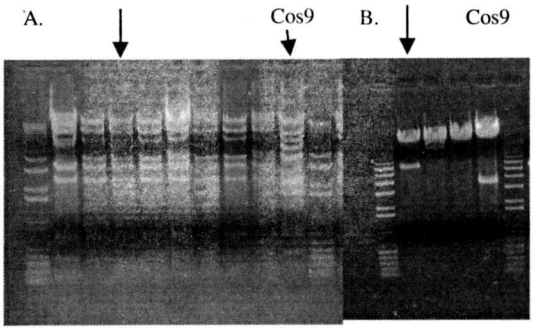


Figure 5.8. Construction of cos9RH. A. *Bam*HI restriction analysis of cosmids resulting from ligation of the missing 8 kb *Bsr*GI fragment with cos9*bls*GHyg. B. *Sca*I digestion of the same cosmids. The cosmid with the correct restriction patterns, cos9RH, is indicated with an arrow. Similarly digested cos9 is also included for comparison.

Following disruption, both cos9 and cos9RH were introduced into *S. lividans* by protoplast transformation. Multiple colonies from the transformations were used to inoculate liquid cultures. After nine days of growth, the length of time previously determined to be optimal for production of blasticidin S,[46] a sample of the culture broth was removed and analyzed for the presence of blasticidin S and related compounds. The culture conditions used were based on a protocol developed by Dr. Martha Cone in her original work on the isolation and characterization of the BS gene cluster.[46]

Although the *S. lividans* transformants grew well and were confirmed to contain the correct cosmid constructs, we were unable to observe blasticidin S production. Genetic instability is a problem that is often encountered with *Streptomyces/E. coli* shuttle vectors.[54,61] The reason for this instability is not know but it has been speculated to involve read through from the *E. coli* segment of the vector as insertion of a transcriptional terminator has been shown to increase plasmid stability.[54] The vector, pOJ446, in particular, has demonstrated instability problems and rearrangement of cosmid and insert DNA has been observed.[46] It was postulated that rearrangement of the cosmid once reintroduced into *S. lividians* may be resulting in other gene disruptions and abolish blasticidin S biosynthesis. Our inability to detect such a rearrangement by *Bam*HI restriction analysis does not preclude the possibility of such a rearrangement. A

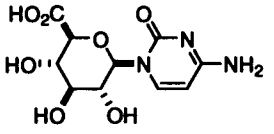
short deletion or point mutation could result in disruption of the gene cluster and would not be detectable through normal restriction analysis. In an attempt to prevent such rearrangements, the BS gene cluster was cloned into two novel *E. coli/ Streptomyces* shuttle vectors, pXY400 and pXY700, developed by Dr. Xihou Yin in our laboratory (Yin and Zabriskie, unpublished results). The first of these two vectors, pXY400, is a derivative of the previously described pXY200. To form pXY400, pXY200 was modified to contain a *cos* site, *oriT*, and the polylinker region from pOJ446. These modifications generated a cosmid vector that is capable of conjugal transfer between bacterial species. The second vector used, pXY700, is an integration vector. pXY700 is a pSET152[60] derivative that includes the polylinker region and *cos* site from pOJ446. Like pSET152, pXY700 is a suicide integration vector that is not capable of replication inside *Streptomyces*. Consequently, chromosomal integration is required for the transformed *Streptomyces* to acquire apramycin resistance carried by the vector. Chromosomal integration has been shown to be a useful approach to the stable heterologous expression of antibiotic gene clusters in *S. lividans* and has been shown to reduce the likelihood of spontaneous mutation.[61]

The cos9 inserts containing the BS gene cluster, with and without the *blsG* disruption, were ligated into pXY400 and pXY700 to form pXY49, pXY49RH, pXY79 and pXY79RH, respectively. The new cosmid constructs were introduced into *S. lividans* TK24 by protoplast transformation. Successful integration of pXY79 and pXY79RH into the *S. lividans* chromosome was confirmed by Southern analysis.

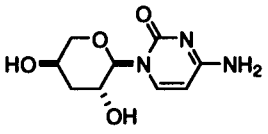
S. lividans/pXY49, *S. lividans*/pXY49RH, *S. lividans*/pXY79 and *S. lividans*/pXY79RH were grown in liquid YEME cultures according to the method of Cone.[46] Aliquots of culture media were removed periodically, starting at nine days and assayed for blasticidin S production. Although differences were observed in the metabolite profiles of *S. lividans* cultures carrying pXY49, pXY49RH, pXY79 or pXY79RH, BS and related metabolites such as CGA and LBS could not be detected. It is unknown if the inability to see BS production is due to the expression system or if a mutation has occurred within the cluster that is preventing the functional expression of the necessary enzymes.

Discussion

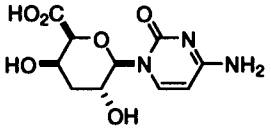
One method used to help define the BS biosynthetic pathway, and identify intermediates, was through the *in vivo* use of enzyme inhibitors.[62,63] The inhibitors were chosen based on their ability to inhibit enzymes that were similar to those predicted to be involved in BS biosynthesis. Among the inhibitors used were a series that were proposed to interfere with the formation of the β-arginine moiety. These included the inhibitors of L-α-arginine biosynthesis; L-arginine hydroxamate[62]and α-methyl aspartate[63] (inhibits argininosuccinate synthesis) as well as transaminase inhibitors[62] (aminooxyacetic acid and 2-methylglutamate). It was found that inclusion of aminooxyacetic acid or L-arginine hydroxamate in the *S. griseochromogenes* culture media resulted in the accumulation of both CGA and pentopyranine C. In addition, Zhang *et al.* found that adding the arginine analog argininic acid to cultures resulted in the formation of a new glycoside by *S. griseochromogenes.*[63] These results demonstrate that inhibition of arginine biosynthesis interfered with BS biosynthesis.



CGA



Pentopyranine C



2-carboxy-pentopyraine C

The results from these inhibitor studies suggested the involvement of a protein complex that is necessary for blasticidin S biosynthesis. In a prokaryotic system where compartmentalization of biosynthetic machinery is not possible, it is important for the bacteria to tightly regulate enzymes involved in secondary metabolism. Without regulation, enzymes such as the aminomutases could deplete the supply of endogenous amino acids, thus hindering protein biosynthesis and killing the cell. One method to control the formation of secondary metabolite intermediates is to assemble the necessary enzymes into protein complexes. The importance of protein-protein interactions for efficient secondary metabolite biosynthesis has been demonstrated in multienzyme polyketide synthases (PKS) where polyketide biosynthesis is dependent on the formation of a 'megasynthase' through specific protein-protein interactions.[64-66] These protein-protein interactions have been found to be involved in directing the movement of

polyketide intermediates between PKS modules.[65] The formation of a similar enzyme complex in BS biosynthesis could provide a direct link between the proteins involved in forming and incorporating β-arginine into BS. In this model, formed β-arginine could be converted into the next biosynthetic product as soon as it is made and may or may not require the formation of a covalently bound enzyme intermediate.

The involvement for such protein-protein interactions may not only be responsible for the results of the enzyme inhibition studies but also for our inability to detect activity from heterologously expressed BlsG. The presence of an enzyme complex in the later steps of BS biosynthesis is also supported by the inability of previous researchers to isolate enzymes involved in the central and late steps of BS biosynthesis. Although enzyme activities involved in several late transformations of BS biosynthesis have been detected, successful purification of these enzymes has been limited. The majority of studies on BS biosynthesis have been conducted in either whole-cell or cell-free systems and only one enzyme involved in BS biosynthesis has been successfully purified.[67] This enzyme, CGA synthase, has been proposed to catalyze the first committed step in the pathway. Enzymes have been characterized for the earliest steps of BS biosynthesis including formation of free cytosine (described in chapter 6 of this thesis) and formation of CGA in cell free systems.[7,67,68] In addition, CFE and inhibitor studies have provided insight into the formation of the aminosugar moiety of cytosinine.[62,69] Similarly, evidence for the later steps in BS biosynthesis have also been detected. These include the formation of LDBS and the final methylation of the δ-nitrogen of the LDBS β-arginine moeity.[7,68] Despite the extensive studies on BS biosynthesis, the key central steps of BS biosynthesis have remained elusive. While the incorporation of β-arginine has been demonstrated *in vivo*, *in vitro* studies have not been successful.

For this reason, gene disruption may present the best opportunity to study BlsG activity. The *in trans* complementation of gene disruptants provides the best opportunity to observe enzyme function in systems where the activity of an enzyme is dependent on the formation of an enzyme complex because all of the necessary biosynthetic enzymes are available for formation of the complex. Alternately, the supplementation of the missing enzyme product may be used to rescue the biosynthesis and demonstrate the

function of a particular enzyme. The metabolic rescue of a given biosynthetic pathway is dependent on the uptake of the exogenous compound and the ability of the compound to be recognized by downstream enzymes for incorporation into the final product. If the downstream enzyme recognizes an enzyme-bound form of the intermediate, it may not be possible to see incorporation of the free intermediate. This is not likely the case in β-arginine incorporation as Zhang has demonstrated the ability of β-arginine to be directly incorporated into BS,[70] however it may account for the inability to see cytosinine incorporation. Regardless, it appears that a mechanism exists in the central steps of BS biosynthesis that may preclude the study of these enzymes from heterologous expression of individual genes.

Additional work to develop a method for the transformation of *S. griseochromogenes* may present the best opportunity to study β-arginine formation. The ability to genetically manipulate the producing organism would allow for the study of BS biosynthesis under native expression and regulatory conditions. One possible reason for the inability to transform *S. griseochromogenes* by conjugation may involve the presence of an effective restriction system in this organism. Similar to protoplast transformation, it has been found that use of a methylation deficient *E. coli* donor strain can increase conjugation efficiency.[60] Typically, one of several *E. coli* strains, including ET12567/pUZ8002 or S17-1, is used in conjugal transfer of DNA to *Streptomyces*. The primary difference in these two strains is the presence of Dam methylases in the latter. As such, S17-1 is not an appropriate conjugal donor for *Streptomyces* species that have methylation-dependent restriction systems, however S17-1 can be used with species that lack this system, such as *S. lividans*.[54] At the time of the *S. griseochromogenes* experiments, the strain of ET12567 cultured in our laboratory did not contain the pUZ8002 RP4 mobilization plasmid, so S17-1 was the only conjugal donor available. As it is not known if *S. griseochromogenes* contains such a methylation dependent restriction system, the use of a methylation deficient conjugal donor such as ET12567/pUZ8002 may result in the successful conjugation of *S. griseochromogenes*.

Molecular genetics of *Streptomyces* is a rapidly expanding field due in large part to the economic importance of this genus.[71] As our understanding of these organisms

grows, new techniques should be developed for their genetic manipulation. The further study and optimization of *S. griseochromogenes* culture and genetic manipulation may allow for the development of an efficient method of transformation. This would present a new and powerful tool for the study of BS biosynthesis through gene disruption and complementation experiments and may allow for the characterization of the elusive central steps in BS biosynthesis.

Materials and Methods

General

Routine molecular biology procedures including DNA manipulation, plasmid construction, transformations, growth and maintenance of *E. coli*, competent cell preparation and SDS-PAGE analysis were conducted according to standard techniques.[72] *S. lividans* and *S. griseochromogenes* were maintained according to standard *Streptomyces* techniques.[54] Restriction enzymes, T4 DNA ligase and *taq* polymerase were purchased from various suppliers and used according to manufacturer's protocol. QIAprep® spin miniprep and QIAquick® gel extraction kits (Qiagen) were used for DNA purification. Sequencing of vector inserts was conducted by the Center for Gene Research and Biotechnology, Oregon State University. Chemicals were purchased from Sigma-Aldrich, Fisher Scientific and VWR International.

Cloning blsG for Heterologous Expression

PCR primers were designed to amplify the *blsG* gene from a cosmid containing the BS gene cluster, cos9.[46] Two primer pairs were initially designed that incorporated an *Nde*I site in forward primer and a *Bgl*II site in the reverse primer. PCR was carried out in a total volume of 50 μL containing 10 ng cos9 template, 1X Promega Thermophilic Buffer, 1.5 mM $MgCl_2$, 0.4 mM each dNTP, 5% DMSO, 50 pmol forward primer UPS#1 5'-GTCCCATATGAGTACGGAATCCG-3' or UPS #2 5'-GAGTTCATATGTCCCACATGAG-3'(*Nde*I sites are underlined), 50 pmol reverse primer DN#1 5'-CGCCTTG TTAGATCTCGTGGCGCAGTTCGC-3' or DN#2 5'GTTAGATCTCGTGG-3' (*Bgl*II sites are underlined), and 5 units Taq polymerase

(Roche). PCR products were gel purified and digested with *Nde*I and *Bgl*II then ligated into *Nde*I and *Bam*HI digested and dephosphorylated pET28a+ or pXY200. The resulting plasmids, pET28*bls*G1, pET28 *bls*G 2, pXY2 *bls*G 1 or pXY2 *bls*G 2, were used to transform *E.coli* DH5α and the correct construction was confirmed by sequencing. Plasmids were then transformed into *E. coli* Bl21 (DE3) (Stratagene) competent cells or *Streptomyces lividans* TK24 protoplasts for expression.

A second set of primers were designed for cloning the *blsG* gene into pET20b+ for expression as the native protein or with a *C*-terminal His_6-tag. PCR was carried out in a total volume of 50 μL containing 10 ng cos9 template, 1X Promega Thermophilic Buffer, 1.5 mM $MgCl_2$, 0.4 mM each dNTP, 5% DMSO, 50 pmol forward primer UPS#1 5'-GTCCCATATGAGTACGGAATCCG-3' or UPS #2 5'GAGTTCATATGTCCCACA TGAG3'(*Nde*I sites are underlined), 50 pmol reverse primer RAM *Native*Rev 5'CACGCCTTGTTCGATCTCGAGGCGCAGTTCGCTCATG 3' or RAM*C-HisRev-* 5'AGGAATTCCTCGAG TCGTGCACCCTCCCCGGTGGC3'(*Xho*I sites are underlined), and 5 units Taq polymerase (Roche).

PCR products were purified, digested with *Nde*I and *Xho*I then ligated into similarly digested and dephosphorylated pET20b+. The resulting plasmids were used to transform *E.coli* JM109 (Promega) and the correct construction was confirmed by sequencing and plasmids were transformed into *E.coli* Bl21 (DE3) or *E.coli* Bl21 Codon Plus (DE3)-RP (Stratagene) competent cells for expression.

Expression and Purification of BlsG in E. coli

Transformants carrying pET28 *bls*G or pET20 *bls*G were grown overnight in Terrific broth supplemented with kanamycin (50 μg/mL) or ampicillin (100 μg/mL), respectively. This seed culture was used to inoculate 1 L of Terrific broth, supplemented with the appropriate antibiotic, grown at 37 °C to an $OD_{600} = 0.5$ and then protein expression was induced with isopropyl-β-D-thiogalactopyranoside (IPTG) at a final concentration of 0.1 mM. The culture was then grown for an additional 8 hours at 37 °C.

Cells were harvested by centrifugation and resuspended in 50 mL lysis buffer (50 mM Tris, 300 mM NaCl, 1 mg/mL lysozyme pH 7.5). Following incubation on ice for

30 min, cells were lysed by sonication in a Microson ultrasonic cell disruptor (six 10 s bursts at 6 watts output with 10 s cooling on ice between bursts). Cellular debris was pelleted by centrifugation at 18000 *x g* for 40 min at 4 °C. Soluble BlsG was purified from the cleared lysate by either nickel affinity chromatography (Ni-NTA spin columns, Qiagen) or cobalt affinity chromatography (TALONspin, BD Bioscience) according to manufacturer's protocol. Eluted protein was concentrated concurrent with the removal of imidazol by buffer exchange in a Centricon 10 (Amicon) centrifugal concentrator. Protein concentration was determined by Bradford analysis using BSA standard.[73]

Expression of blsG in S. lividans

For heterologous expression of pXY2*bls*G in *S. lividans*, a 5 mL seed culture was inoculated from a mycelial stock into YEME broth supplemented with 2.5 μg/mL apramycin. The mycelium harvested from a 20 h seed culture by centrifugation, homogenized, and used to inoculate a 40 mL culture in YEME broth supplemented with 2.5 μg/mL apramycin. 5 mL portions of the inoculated culture broth were transferred to 20 mL glass culture tubes. Protein expression was induced by adding thiostrepton at 5μg/mL to the culture media. Cultures were grown at 30 °C and 300 rpm for 24 or 75 h.

Culture broth was concentrated by lyophilization, redissolved in 3 mL lithium carbonate buffer, and derivatized for HPLC analysis as described below.

In order to confirm heterologous expression of pXY2*bls*G, expressed protein was purified using nickel affinity chromatography (Qiagen NTI spin). The mycelial pellet from the above expression experiments was resuspended in 1 mL denaturing buffer (8 M urea, 100 mM NaH_2PO_4, 10 mM Tris, pH 8.0) and lysed by sonication on ice, using four 10 s bursts on a Microsonix Micorson with 20 s incubation on ice between bursts. The insoluble material was cleared from the cell lysate by centrifugation for 45 min at 3500 rpm on a Beckman GPR centrifuge. Cleared lysate was applied to a Qiagen NTI spin column. Following binding, the resin was washed two times with 700 μL wash buffer (8 M urea, 100 mM NaH_2PO_4, 10 mM Tris, pH 6.3). Protein was eluted with 200 μL elution buffer (8 M urea, 100 mM NaH2PO4, 10 mM Tris pH 4.5) and analyzed by SDS-

PAGE on a 12% gel. Protein expression was confirmed by western blot analysis using Novagen His_6-tag western detection according to manufacturer's protocol.

Soluble BlsM was also purified from the cleared *S. lividans* lysate by cobalt affinity chromatography (TALONspin, BD Bioscience) according to manufacturer's protocols. Eluted protein was concentrated concurrent with the removal of imidazole by buffer exchange in a Centricon 10 (Amicon) centrifugal concentrator. Protein concentration was determined by Bradford analysis using BSA as a standard.[73]

Protoplast Transformation of Streptomyces lividans TK24

Three hundred mL of YEME media, containing 0.5% glycine, was inoculated with approximately $2x10^7$ *Streptomyces lividans TK24* spores in a 2 L baffled flask. Protoplasts were prepared from 36-40 h cultures grown at 30 °C and 300 rpm. Mycelium was harvested by centrifugation at 3000 rpm and 10 °C for 10 min on a Beckman GPR swinging bucket centrifuge. Harvested mycelium was washed once in 10.3% sucrose and twice in Buffer P (10.3% sucrose, 44 mM K_2SO_4, 41 mM $MgCl_2$, 0.2% *Streptomyces* trace element solution,[54] 0.005% KH_2PO_4, 0.368% $CaCl_2$, and 0.573% TES, pH 7.2). Mycelium was then resuspended in 50 mL sterile Buffer P supplemented with 1 mg/mL lysozyme (Sigma) and incubated at 30 °C for 20-40 min. Incubation time varied and was monitored by microscopic examination of protoplast formation. The remaining mycelium was removed from the protoplasts by passing through a sterile protoplast filter[54] on ice. Protoplasts were gently pelleted from the filtrate by centrifuging for 10 min at 10 °C and 2500 rpm on a Beckman GPR swinging bucket centrifuge. Harvested protoplasts were washed twice by gently resuspending in 10 mL 4 °C Buffer P followed by centrifugation and removal of supernatant. Finally, protoplasts were resuspended in 2 mL 4 °C Buffer P and 120 µL aliquots were transferred to sterile 1.5 mL microcentrifuge tubes. Protoplasts were frozen slowly in fresh ice at -20 °C for several hours and then transferred to a -80 °C freezer for storage.

Protoplasts were transformed using the method of Hopwood *et al.*[54] Briefly, 5 µL concentrated plasmid DNA was combined with 120 µL protoplasts in a 1.5 mL microcentrifuge tube at room temperature (RT). 250 µL of 25% PEG in Buffer P was immediately added and the solution was mixed gently. Following incubation at RT for 2

min, 250 μL Buffer P was gently added to the protoplasts and the solution was transferred to a 12x75 mm culture tube. An additional 3.5 mL Buffer P was added and the suspension was mixed gently. Protoplasts were harvested by centrifugation at 10 °C for 7 min at 2500 rpm. The majority of the supernatant was removed and protoplasts were resuspended in approximately 500 μL remaining buffer and plated on two R2YE plates. Following incubation at 30 °C for 18-24 h, soft nutrient agar containing the appropriate selection antibiotic was added. Plates were incubated at 30 °C for up to one week to allow transformants to grow.

Preparation of Cell Free Extracts

Cell free extracts (CFE) of *E. coli* were prepared as follows. Cells from a 100 mL expression culture were thawed on ice and then resuspended in 5 mL 4 °C HEPPS buffer pH 8.0. Cells were lysed by sonication in a Microson ultrasonic cell disruptor (four 30 s bursts at 6 watts output with 60 s cooling on ice between bursts). Cellular debris was pelleted by centrifugation at 18000 *x g* for 30 min at 4 °C. CFE was kept on ice and immediately used in BlsG assays as described below.

Initial assay for BlsG Activity, Aerobic Conditions

Prior to addition of arginine, purified BlsG was preincubated for 1.5 h with 0.1 mM pyridoxal phosphate in 80 mM HEPPS pH 8.0. 50 μL preincubation mix was then added to a final reaction volumn of 100 μL containing 20 mM arginine, 10 μM AdoMet, 50 μM pyridoxal phosphate and 80 mM HEPPS. Assays were incubated at 37 °C for 1.5 h then quenched by adding an equal volume of 100% ethanol and pelleting precipitated proteins by centrifugation. The resulting supernatant was derivatized with dansyl chloride for TLC and HPLC analysis as described below.

Anaerobic Assay for BlsG Activity

Cell free extracts and purified BlsG protein were assayed for activity under anaerobic assay conditions. Assays were based on similar assays with lysine 2,3-aminomutase.[29] Prior to addition of arginine, CFE or purified BlsG was reductively

preincubated with 8 mM cysteine, 1 mM ammonium iron (II) sulfate, 0.5 mM pyridoxal phosphate and 100 mM HEPPS pH 8.0. Reagents were combined under a flow of argon and mixed well. Capillary tubes, 1.6 X 100 mm, were filled with argon before 100 µL of preincubation mix was added. The ends of the capillary tube were immediately sealed by heating briefly in a flame. The sealed preincubation mixtures were incubated at 37 °C for 3 h in the dark.

Following preincubation, BlsG assays were conducted in 1.5 mL microcentrifuge tubes filled with argon. 90 µL preincubation mix was added to 910 µL of a solution containing cofactors and substrate for a final reaction mixture containing 44 mM arginine, 3.5 mM dithionite (sodium hydrosulfite), 80 µM AdoMet, 720 µM cysteine, 90 µM Fe (II) sulfate, 45 µM pyridoxal phosphate and 100 mM HEPPS. Assays were incubated at 30 °C for 1.5 hours then stopped by boiling for 10 min. Following centrifugation to remove protein, arginine was purified from the assays by cation exchange chromatography. Assays were acidified by addition of 8N HCl, applied to the cation exchange resin (Dowex 50 X 8-400, Acros) and washed with three volumes of ddH_2O. Bound material was eluted from the resin with 1N NH_4OH. Ninhydrin positive fractions were combined, concentrated and derivatized for analysis by HPLC as described below.

Purification of Arginine from Culture Media

Basic amino acids were partially purified from *E. coli* and *S. lividans* culture broth by cation exchange chromatography. Culture broth was acidified by addition of 8N HCl and applied to the cation exchange resin (Dowex 50 X 8-400, Acros) and washed with three volumes of ddH_2O. Bound material was eluted from the resin with 1N NH_4OH. Ninhydrin positive fractions were combined and concentrated under vacuum. Fractions were analyzed by TLC on aluminum backed silica plates resolved with 4:2:1 butanol/water/acetic acid and like fractions were pooled and reconcentrated, redissolved in lithium carbonate buffer and derivatized for analysis by HPLC as described below.

In addition to analyzing the culture broth for β-arginine, the mycelial fraction of *S. lividans* pXY2BlsG cultures was also examined. A 0.3-0.5 g mycelial pellet was

suspended in 1 mL sterile H_2O and lysed by sonication as described above. Proteins were precipitated from the cell lysate by boiling for 10 min, followed by incubation on ice for an equal time and centrifugation for 5 min at 18,000 x *g*. The aqueous cellular extract was then extracted three times with an equal volume of ethyl ether, and then three times with ethyl acetate. After each extraction, the aqueous and organic layers were separated by centrifugation at 18,000 *x g*. The aqueous layers were lyophilized and derivatized for analysis by HPLC as described below. Alternately, the aqueous layer was extracted with ethyl acetate and then hexanes to yield cleaner samples with fewer background peaks in the HPLC chromatogram.

Danysl Derivatization of Amino Acids

Arginine, β-arginine, and BlsG assays were derivatized with dansyl chloride to generate the dansyl derivative for UV detection by HPLC. Dansylation reactions were performed by combining 500 µL of 1 mM amino acid standard or BlsG assay mixture in 40 mM lithium carbonate buffer pH 9.5 with 250 µL dansyl chloride solution (1.5 mg/mL in acetonitrile). Reactions were incubated in the dark for thirty minutes at room temperature and quenched by adding 25 µL 2% aqueous ethylamine. Samples were filtered through a 0.2 µm syringe filter before HPLC analysis.

HPLC Analysis of Dansylated Amino Acids

Dansylated α and β-arginine were separated by HPLC (Beckman Ultrasphere C_{18} 5µm, 4.6 x 250 mm column; isocratic mobile phase consisting of 80% 0.1 M potassium phosphate buffer, pH 7.5, 20% acetonitrile, 0.6 mL/min) with UV detection at 265 nm on either a Beckman System Gold or a Beckman Biosys 510 system.

Protoplast Transformation of S. griseochromogenes

Numerous attempts were made to transform *Streptomyces griseochromogenes*. The initial attempts were to transform *S. griseochromogenes* through protoplast transformation. The methods used were based on established protocols for protoplast transformation of *S. lividans*[54] as described above with appropriate modifications for the

growth of *S. griseochromogenes*. Sporulation of *S. griseochromogenes* was conducted as described by Guo.[74] *S. griseochromogenes* was grown from a glycerol spore suspension on YME agar containing 4 g/L yeast extract, 10 g/L malt extract, 4 g/L dextrose, and 20 g/L Bacto agar. Spores were harvested after incubating the plates at 30 °C for nine days. For preparation of protoplasts, Tryptic Soy Broth (TSB) supplemented with 0.5% glycine was used in place of YEME medium. After initial attempts using 25% PEG to induce transformation, a range of PEG concentrations, between 5 and 50% were used to try and transform *S. griseochromogenes.*

Transformation of S. griseochromogenes by Conjugation.

After several attempts to transform *S. griseochromogenes* protoplasts, transformation by conjugation was tried according to the methods of Hopwood.[54] The conjugal plasmid pKC1139 was used to transform *E. coli* S17-1 competent cells. *E. coli* cells were harvested from 20 h, 7.5 mL Terrific Broth cultures by centrifugation and washed three times in fresh culture media. After washing, the cells were resuspended in 2.5 mL Terrific Broth and stored at 4 °C until preparation of spores was complete.

Fresh *S. griseochromogenes* spores were prepared by washing 200 µL spores (harvested from one plate) twice in 2 mL 0.05 M TES pH 8.0. Spores were resuspended in 2 mL 0.05 M TES pH 8.0 and incubated at 50 °C for 10 min before being cooled by placing the tube in tap water. Two mL of 2X pregermination buffer (1% Difco yeast extract, 1% casaminoacids, and 10 mM $CaCl_2$) was added to the spore suspension and incubated for 4 h at 37 °C and 285 rpm. Periodic vortexing was used to keep the spores in solution during incubation. After pregermination, spores were harvested by centrifugation and resuspended in 0.5 mL Terrific Broth.

Spore dilutions (1X, 10X, 100X or 1000X) in Terrific Broth were added to an equal volume of prepared *E. coli* cells and plated on AS-1 plates (containing 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 dextrin, 2.5 g/L NaCl, 10 g/L Na_2SO_4 and 20 g/L agar)[75]. After incubation at 30 °C for 16 to 22 h, soft nutrient agar (5 g/L Bacto agar, 8 g/L nutrient broth) containing nalidixic acid and apramycin was added for a final concentration of 25 µg/mL nalidixic acid and 20 µg/mL

apramycin. Plates were incubated at 30 °C for up to three weeks and monitored for the growth of exconjugates.

Initial experiments determined that *S. griseochromogenes* grows very slowly on AS-1 agar and so conjugation experiments were repeated with the use of HT and YME agar plates. Although control cultures, with no added apramycin, had improved growth of *S. griseochromogenes*, no colonies were formed on the plates with antibiotic selection.

HPLC Detection of Blasticidin S and Leucylblasticidin S.

Detection of blasticidin S and leucylblasticidin S was performed on a Beckman System Gold or a Beckman BioSys 510 HPLC (Varian Microsorb-MV amino, 150 x 4.6 mm column; 5 min at 70% aqueous acetonitrile followed by a linear gradient to 30% aqueous acetonitrile over 20 min and then maintained at 30% acetonitrile for 5 min) with UV detection at 265 nm.

Synthesis of β-Arginine

The preparation of β-arginine was based on the synthesis of blastidic acid by Nomoto and Shimoyama.[48]

Formation of ethoxyanhydride, **5.72**.

N-BOC-δ-CBZ-β-ornithine was formed by a modification of a literature procedure.[51,52] Under an inert argon atmosphere, one equivalent of each *N*-methylmorpholine and ethyl chloroformate was added to a solution of 37 mM BOC-δ-CBZ-diaminobutyric acid **5.71** (Bachem) in dry ethyl acetate. The reaction was cooled to -15 °C and stirred for 4 h, during which a milky white precipitate formed. After 4 h the white precipitant was removed by filtration and washed three times with ethyl acetate. Ethyl acetate washes were combined with filtrate containing **5.72** and used in the next step without purification.

Formation of the diazoketone, **5.73**.

The method used for the formation of the diazoketone **5.73** was altered from that reported by Shiba to utilize a safer method for generating diazomethane.[50,52]

Diazomethane was generated through the dropwise addition of saturated sodium hydroxide to a solution of 2% w/v diazald in ethanol.[50] The diazomethane gas was bubbled through **5.72** in ethyl acetate until the solution became bright yellow, indicating complete saturation with diazomethane. The solution was then covered and stirred in a NaCl/ice bath overnight. After stirring overnight at 0 °C, the reaction was warmed to 50 °C and stirred for an additional 3 h until the color faded to a pale yellow. The ethylacetate was removed under vacuum to give a yellow oil.

Wolf Rearrangement of **5.73** to **5.74.**

The diazoketone **5.73** was redissolved in five mL methanol and 315 μL of a 10% solution of silver benzoate (AlfaAesar) in distilled triethylamine was added. The flask was wrapped in aluminum foil and stirred at room temperature in the dark overnight. The resulting dark brown solution was concentrated under vacuum and the residue redissolved in ethyl acetate. A dark precipitate was removed by filtration and the filtrate was first extracted three times with saturated sodium bicarbonate and then three times with saturated sodium chloride (brine). Following the brine washes the solution had turned from an amber color to a more yellow color. Finally the ethyl acetate layer was washed three times with 1N HCl. After the HCl washes the solution became yellow and a black oil remained on the sides of the separatory funnel. The ethyl acetate fraction was washed with brine until the pH of the brine wash returned to neutral and dried over anhydrous sodium sulfate. The crude **5.74** was further purified by silica gel flash chromatography with a mobile phase of 3:7 pentanes:ethyl ether. The overall yield for the first three steps from **5.71** to **5.74** was 78%. ^{1}H NMR (300 MHz, $CDCl_3$) 7.5 (m, 5H), 5.1 (s, 2H), 4.0 (m, 1H), 3.65 (s, 3H), 3.2 (d, 2H), 2.5 (m, 2H), 1.7 (m, 2H), 1.4 (s, 9H). ^{13}C NMR (300 MHz, $CDCl_3$) 172, 156, 136.7, 128.4, 128.0, 79.6, 66.5, 51.7, 44.6, 39.0, 37.6, 34.8, 28.3. HRMS (FAB) calculated for **5.74** 381.202, observed 381.202.

Saponification of **5.74** to give **5.75**.

The final step in synthesis of N-BOC-δ-CBZ-β-ornithine is the cleavage of the methylester.[52] Compound **5.74** was dissolved in dioxane and stirred in at 0 °C. One

equivalent of sodium hydroxide was added dropwise and the reaction was allowed to stir at room temperature overnight.

The reaction was worked up by first extracting the water/dioxane layer three times with ethyl acetate. The dioxane layer was then cooled in an ice-water bath and slowly acidified to pH 2-3 with the dropwise addition of 1N HCl. Following acidification, the aqueous layer was extracted three times with ethyl acetate. The organic fractions were combined and dried over anhydrous sodium sulfate. The reaction product, **5.75**, was concentrated under vacuum. ^{1}H NMR (300 MHz, $CDCl_3$) 7.4 (s, 5H), 5.1 (s, 2H), 4.0 (m, 1H), 3.2 (d, 2H), 2.5 (m, 2H), 1.7 (m, 2H), 1.4 (s, 9 H). ^{13}C NMR (300 MHz, $CDCl_3$) 175.4, 156.7, 156.0, 136, 128.4, 128.0, 79.7, 66.6, 44.6, 38.9, 37.6, 34.6, 28.2. HRMS (FAB) calculated for **5.75** 367.187, observed 367.186.

Formation of **5.76** from **5.75.**

Before the guanidine group can be added to the β-ornithine derivative, the protecting benzylcarboxy group must be removed from the δ-amine. This is accomplished via a catalytic hydrogenolysis.[48] Twenty mgs of 10% Pd/C (Aldrich) was added with vigorous stirring to 118 mg **5.75** dissolved in 10 mL methanol. The reaction was stirred under a hydrogen atmosphere for 18 hours. Following concentration under vacuum the product was analyzed by NMR. ^{1}H NMR (300 MHz, D_4-MeOH) 4.0 (m, 1H), 3.0 (m, 2H), 2.5 (d, 2H), 1.9 (m, 2H), 1.4 (s, 9H). ^{13}C NMR (300 MHz, MeOH-d_4) 175.0, 157.4, 79.5, 45.7, 40.5, 37.1, 32.9, 27.8. Overall yield for the combined saponification and deprotection was 86%. HRMS (FAB): calculated for **5.76** 233.150, observed 233.150.

Guanidination of **5.76**.

Guanidination of *N*-BOC-β-ornithine was accomplished according to the procedures of Nomoto and Shiba.[53] Compound **5.76** (0.32 mmol) was dissolved in 10 mL 1N sodium hydroxide and *O*-methylisourea (0.42 mmol) was added with vigorous stirring. The reaction stirred at room temperature for 24 h and then the pH of the reaction mixture was adjusted to pH 3 with 4N HCl and concentrated under vacuum. Acidification of the

reaction mixture cleaved the BOC protecting group and the product, β-arginine **5.78** was purified by cation exchange chromatography (Dowex 50 X 8-400, Acros) with an 82% yield. ^{1}H NMR (300 MHz, D_2O) 3.7 (m, 1H), 3.2 (m, 2H), 2.6 (m, 2H), 2.1 (m, 2H). ^{13}C NMR (300 MHz, D_2O) 174.2, 160.5, 44.4, 35.2, 33.7, 27.7. MS (FAB) calculated for **5.78** 173.1, observed 173.1.

References

(1) Seto, H.; Yamaguchi, N.; Otake, N.; Yonehara, H. Studies on the biosynthesis of blasticidin S. Part I. Precursors of blasticidin S biosynthesis. *Agric. Biol. Chem.* **1968**, *32*, 1292-1298.

(2) Zabriskie, T. M.; Klocke, J.; Ireland, C. M.; Marcus, A.; Molinski, T.; Faulkner, J.; Xu, C.; Clardy, J. Jaspamide, a Modified Peptide from a Jaspis Sponge, with Insecticidal and Antifungal Activity. *J. Am. Chem. Soc.* **1986**, *108*, 3123-3124.

(3) Kurylo-Boroska, Z.; Abramsky, T. Biosynthesis of beta-Tyrosine. *Biochim. Biophys. Acta.* **1972**, *264*, 1-10.

(4) Noda, T.; Take, T.; Nagata, A.; Wakamiya, T.; Shiba, T. Chemical studies on tuberactinomycin. 3. The chemical structure of viomycin (tuberactinomycin B). *J. Antibiot. (Tokyo)* **1972**, *25*, 427-428.

(5) Thiruvengadam, T. K.; Gould, S. J.; Aberhart, D. J.; Lin, H. Biosynthesis of Streptothricin F. 5. Formation of β-Lysine by *Streptomyces* L-1689-23. *J. Am. Chem. Soc.* **1983**, *105*, 5470-5476.

(6) Williamson, J. M.; Brown, G. M. Purification and properties of L-Aspartate-alpha-decarboxylase, an enzyme that catalyzes the formation of beta-alanine in *Escherichia coli*. *J. Biol. Chem.* **1979**, *254*, 8074-8082.

(7) Zhang, Q.; Cone, M. C.; Gould, S. J.; Zabriskie, T. M. Reevaluation of the Final Steps in the Biosynthesis of Blasticidin S by *Streptomyces griseochromogenes* and Identification of a Novel Self-Resistance Mechanism. *Tetrahedron* **2000**, *56*, 693-701.

(8) Takana, Y.; Sakagami, Y.; Nishimura, T.; Yamaki, H.; Umezawa, H. Activity of cytomycin and blasticidin S against transplantable animal tumors. *J. Antibiot.* **1961**, *14*, 123-126.

(9) Mathews, C. K.; vanHolde, K. E.; Ahern, K. G. Nucleotide Metabolism. *Biochemistry*; Third ed.; Addison Wesley Longman, Inc.: San Francisco, 2000; pp 806-808.

(10) Walsh, C. T. Reactions at the alpha carbon of amino acids. *Enzymatic Reaction Mechanisms*; W.H. Freeman and Company: San Francisco, 1979; pp 801.

(11) Nakano, Y.; Kitaoka, S. L-aspartate -decarboxylase in a cell-free system from *Escherichia coli*. *J. Biochem. (Tokyo)* **1971**, *70*, 327-334.

(12) Saldanha, S. A.; Birch, L. M.; Webb, M. E.; Nabbs, B. K.; von Delft, F.; Smith, A. G.; Abell, C. Identification of Tyr58 as the proton donor in the aspartate-alpha-decarboxylase reaction. *Chem. Commun.* **2001**, 1760-1761.

(13) Abell, C. Enzymes that exploit imines--one way or the other. *Biochem. Soc. Trans.* **1998**, *26*, 310-315.

(14) Poston, J. M. Leucine 2,3-aminomutase, an enzyme of leucine catabolism. *J. Biol. Chem.* **1976**, *251*, 1859-1863.

(15) Poston, J. M. Leucine 2,3-aminomutase: A Cobalamin-Dependent Enzyme Present in Bean Seedlings. *Science* **1977**, *195*, 301-302.

(16) Poston, J. M. Coenzyme B_{12} Dependent Enzymes in Potatoes: Leucine 2,3-aminomutase and methylmalonyl-CoA mutase. *Phytochemistry* **1978**, *17*, 401-402.

(17) Poston, J. M.; Hemmings, B. A. Cobalamins and cobalamin-dependent enzymes in *Candida utilis*. *J. Bacteriol.* **1979**, *140*, 1013-1016.

(18) Poston, J. M. Distribution of leucine 2,3-aminomutase activity in various organs of the rat and in subcellular organelles of rat liver. *Biochem. Biophys. Res. Commun.* **1980**, *96*, 838-843.

(19) Poston, J. M. Beta-leucine and the beta-keto pathway of leucine metabolism. *Adv. Enzymol. Relat. Areas Mol. Biol.* **1986**, *58*, 173-189.

(20) Frey, P. A. The role of radicals in enzymatic processes. *Chem. Rec.* **2001**, *1*, 277-289.

(21) Frey, P. A. Radical mechanisms of enzymatic catalysis. *Annu. Rev. Biochem.* **2001**, *70*, 121-148.

(22) Parry, R. J.; Kurylo-Boroska, Z. Biosynthesis of Amino Acids. Investigation of the Mechanism of B-Tyrosine Formation. *J. Am. Chem. Soc.* **1980**, *102*, 836-837.

(23) Christenson, S. D.; Liu, W.; Toney, M. D.; Shen, B. A novel 4-methylideneimidazole-5-one-containing tyrosine aminomutase in enediyne antitumor antibiotic C-1027 biosynthesis. *J Am Chem Soc* **2003**, *125*, 6062-6063.

(24) Stadtman, T. Cobamide Coenzyme Requirement for the Anaerobic Degredation of Lysine. *Ann. N. Y. Acad. Sci.* **1964**, *112*, 728-734.

(25) Stadtman, T. Lysine Fermentation to Fatty Acids and Ammonia: A Coamide Coenzyme-dependent Process. *J. Biol. Chem.* **1962**, *237*, PC2409-PC2411.

(26) Costilow, R. N.; Rochovansky, O. M.; Barker, H. A. Isolation and identification of beta-lysine as an intermediate in lysine fermentation. *J. Biol. Chem.* **1966**, *241*, 1573-1580.

(27) Chirpich, T. P.; Zappia, V.; Costilow, R. N.; Barker, H. A. Lysine 2,3-aminomutase. Purification and properties of a pyridoxal phosphate and *S*-adenosylmethionine-activated enzyme. *J. Biol. Chem.* **1970**, *245*, 1778-1789.

(28) Aberhart, D. J.; Gould, S. J.; Lin, H.; Thiruvengadam, T. K.; Weiller, B. H. Stereochemistry of Lysine 2,3-aminomutase Isolated from *Clostridium subterminale* Strain SB4. *J. Am. Chem. Soc.* **1983**, *105*, 5461-5470.

(29) Chen, D.; Ruzicka, F. J.; Frey, P. A. A novel lysine 2,3-aminomutase encoded by the yodO gene of bacillus subtilis: characterization and the observation of organic radical intermediates. *Biochem. J.* **2000**, *348 Pt 3*, 539-549.

(30) Ruzicka, F. J.; Lieder, K. W.; Frey, P. A. Lysine 2,3-aminomutase from Clostridium subterminale SB4: mass spectral characterization of cyanogen bromide-treated peptides and cloning, sequencing, and expression of the gene kamA in *Escherichia coli*. *J. Bacteriol.* **2000**, *182*, 469-476.

(31) Ballinger, M. D.; Frey, P. A.; Reed, G. H. Structure of a substrate radical intermediate in the reaction of lysine 2,3-aminomutase. *Biochemistry* **1992**, *31*, 10782-10789.

(32) Chen, D.; Frey, P. A. Identification of lysine 346 as a functionally important residue for pyridoxal 5'-phosphate binding and catalysis in lysine 2, 3-aminomutase from *Bacillus subtilis*. *Biochemistry* **2001**, *40*, 596-602.

(33) Lieder, K. W.; Booker, S.; Ruzicka, F. J.; Beinert, H.; Reed, G. H.; Frey, P. A. *S*-Adenosylmethionine-dependent reduction of lysine 2,3-aminomutase and observation of the catalytically functional iron-sulfur centers by electron paramagnetic resonance. *Biochemistry* **1998**, *37*, 2578-2585.

(34) Ballinger, M. D.; Frey, P. A.; Reed, G. H.; LoBrutto, R. Pulsed electron paramagnetic resonance studies of the lysine 2,3-aminomutase substrate radical: evidence for participation of pyridoxal 5'-phosphate in a radical rearrangement. *Biochemistry* **1995**, *34*, 10086-10093.

(35) Prabhakaran, P. C. Biosynthesis of Blasticidin S: Mechanistic and Stereochemical Investigations. In *Chemistry*; Oregon State University: Corvallis, 1988; pp 144.

(36) Prabhakaran, P. C.; Woo, N.; Yorgey, P.; Gould, S. J. Biosynthesis of Blasticidin S from L-alpha-arginine. Stereochemistry in the Arginine 2,3-aminomutase Reaction. *J. Am. Chem. Soc.* **1988**, *110*, 5785-5791.

(37) Cone, M. C.; Yin, X.; Grochowski, L. L.; Parker, M. R.; Zabriskie, T. M. The blasticidin S biosynthesis gene cluster from *Streptomyces griseochromogenes*: sequence analysis, organization, and initial characterization. *Chembiochem* **2003**, *4*, 821-828.

(38) Strohl, W. Compilation and analysis of DNA sequences associated with apparent streptomycete promoters. *Nucleic Acids Res.* **1992**, *20*, 961-974.

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

(40) Sofia, H. J.; Chen, G.; Hetzler, B. G.; Reyes-Spindola, J. F.; Miller, N. E. Radical SAM, a novel protein superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: functional characterization using new analysis and information visualization methods. *Nucleic Acids Res.* **2001**, *29*, 1097-1106.

(41) Aberhart, D. J. Separation by high-performance liquid chromatography of alpha- and beta-amino acids: application to assays of lysine 2,3-aminomutase and leucine 2,3-aminomutase. *Anal. Biochem.* **1988**, *169*, 350-355.

(42) Tapuhi, Y.; Schmidt, D. E.; Lindner, W.; Karger, B. L. Dansylation of amino acids for high-performance liquid chromatography analysis. *Anal. Biochem.* **1981**, *115*, 123-129.

(43) Nakamura, Y.; Gojobori, T.; Ikemura, T. Codon usage tabulated from international DNA sequence databases: status for the year 2000. *Nucleic Acids Res.* **2000**, *28*, 292.

(44) Kane, J. F. Effects of rare codon clusters on high-level expression of heterologous proteins in Escherichia coli. *Curr. Opin. Biotechnol.* **1995**, *6*, 494-500.

(45) Brinkmann, U.; Mattes, R. E.; Buckel, P. High-level expression of recombinant genes in *Escherichia coli* is dependent on the availability of the dnaY gene product. *Gene* **1989**, *85*, 109-114.

(46) Cone, M. C.; Petrich, A. K.; Gould, S. J.; Zabriskie, T. M. Cloning and heterologous expression of blasticidin S biosynthetic genes from *Streptomyces griseochromogenes*. *J. Antibiot. (Tokyo)* **1998**, *51*, 570-578.

(47) Kondo, T.; Nakai, H.; Goto, T. Synthesis of Cytosinine, The Nucleoside Component of Antibiotic Blasticidin S. *Tetrahedron Lett.* **1972**, *19*, 1881-1884.

(48) Nomoto, S.; Shimoyama, A. First synthesis of blastidic acid, a component amino acid in an antibiotic, blasticidin S. *Tetrahedron Lett.* **2001**, *42*, 1753-1755.

(49) Ichikawa, Y.; Ohbayashi, M.; Nishizawa, R.; Isobe, M. Synthesis of Blastidic Acid and Cytosinine, Two Components of Blasticidin S. *Synlett* **2001**, *11*, 1763-1766.

(50) Lombardi, P. A rapid, safe, and convenient procedure for the preparation and use of diazomethane. *Chemistry and Industry* **1990**, 708.

(51) Wakamiya, T. U., H.; Teshima, T.; and Shiba, T Synthesis of Acyl Derivatives of B-Lysine for Peptide Synthesis. *Bull. Chem. Soc. Japan* **1975**, *48*, 2401-2402.

(52) Wakamiya, T. T., T.; Sakakibara, H.; Fukukawa, K.; and Shiba, T Chemical Studies on Tuberactinomycin. XI. 1. Semisynthesis of Tuberactinomycin Analogs with Various Amino Acids in the Branched Part. *Bull. Chem. Soc. Japan* **1977**, *50*, 1984-1989.

(53) Nomoto, S.; Shiba, T. Synthesis of N-B-Methyl-L-B-arginine, a component in a new antibiotic LL-BM547B. *Chemistry Lett.* **1978**, 589-590.

(54) Kieser, T. B., M.; Buttner, M.; Chater, K.; and Hopwood, D. *Practical Streptomyces Genetics*; 1*st* ed.; The John Innes Foundation: Nowich, 2000; 613.

(55) Taylor, D. P.; Eckhardt, T.; Fare, L. R. Gene cloning in the actinomycetes. *Ann. N. Y. Acad. Sci.* **1983**, *413*, 47-56.

(56) Hopwood, D. A. Genetic studies with bacterial protoplasts. *Annu. Rev. Microbiol.* **1981**, *35*, 237-272.

(57) Hopwood, D. A.; Wright, H. M.; Bibb, M. J.; Cohen, S. N. Genetic recombination through protoplast fusion in Streptomyces. *Nature* **1977**, *268*, 171-174.

(58) Hammes, W.; Schleifer, K. H.; Kandler, O. Mode of action of glycine on the biosynthesis of peptidoglycan. *J. Bacteriol.* **1973**, *116*, 1029-1053.

(59) Mazodier, P.; Petter, R.; Thompson, C. Intergeneric conjugation between *Escherichia coli* and *Streptomyces* species. *J. Bacteriol.* **1989**, *171*, 3583-3585.

(60) Flett, F.; Mersinias, V.; Smith, C. P. High efficiency intergeneric conjugal transfer of plasmid DNA from *Escherichia coli* to methyl DNA-restricting streptomycetes. *FEMS Microbiol. Lett.* **1997**, *155*, 223-229.

(61) Pugsley, A. P. Genetic instability in streptomycetes. *Microbiol. Sci.* **1987**, *4*, 11-13.

(62) Gould, S. J.; Guo, J.; Geitmann, A.; Dejesus, K. Nucleoside intermediates in blasticidin S biosynthesis identified by the in vivo use of enzyme inhibitors. *Canadian J. Chem.* **1994**, *72*, 6-11.

(63) Zhang, Q.; Gould, S. J.; Zabriskie, T. M. A new cytosine glycoside from *Streptomyces griseochromogenes* produced by the use in vivo of enzyme inhibitors. *J. Nat. Prod.* **1998**, *61*, 648-651.

(64) Wu, N.; Tsuji, S. Y.; Cane, D. E.; Khosla, C. Assessing the balance between protein-protein interactions and enzyme-substrate interactions in the channeling of intermediates between polyketide synthase modules. *J. Am. Chem. Soc.* **2001**, *123*, 6465-6474.

(65) Tsuji, S. Y.; Cane, D. E.; Khosla, C. Selective protein-protein interactions direct channeling of intermediates between polyketide synthase modules. *Biochemistry* **2001**, *40*, 2326-2331.

(66) Tsuji, S. Y.; Wu, N.; Khosla, C. Intermodular communication in polyketide synthases: comparing the role of protein-protein interactions to those in other multidomain proteins. *Biochemistry* **2001**, *40*, 2317-2325.

(67) Guo, J.; Gould, S. J. Biosynthesis of Blasticidin S from Cytosylglucuronic Acid (CGA). Isolation of Cytosine/UDPglucuronoxyltransferase and Incorporation of CGA by *Streptomyces griseochromogenes*. *J. Am. Chem. Soc.* **1991**, *113*, 5898-5899.

(68) Gould, S. J.; Guo, J. Biosynthesis of blasticidin S. Cell-free demonstration of N-methylation as the last step. *Bioorg. Med. Chem. Lett.* **1991**, *1*, 497-500.

(69) Gould, S. J.; Zhang, Q. Cytosinine: pyridoxal phosphate tautomerase, a new enzyme in the blasticidin S biosynthetic pathway. *J. Antibiot. (Tokyo)* **1995**, *48*, 652-656.

(70) Zhang, Q. Advanced Studies of Blasticidin S Biosynthesis. In *Chemistry*; Oregon State University: Corvallis, 1998; pp 159.

(71) Demain, A. L. Small bugs, big business: the economic power of the microbe. *Biotechnol. Adv.* **2000**, *18*, 499-514.

(72) Sambrook, J.; Russell, D. *Molecular Cloning A Laboratory Manual*; 3 ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, 2001.

(73) Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **1976**, *72*, 248-254.

(74) Guo, J. Biosynthesis of Blasticidin S: Pathway and Enzymes for the Nucleoside Formation and Blastidic Acid Assembly. In *Biochem. Biophys.*; Oregon State University: Corvallis, 1992; pp 202.

(75) Baltz, R. Genetic Recombination in *S. fradiae* by protoplast fusion and cell regeneration. *Dev. Ind. Microbiol.* **1980**, *21*, 43-54.

Chapter 6
Characterization of BlsM, a Nucleotide Hydrolase

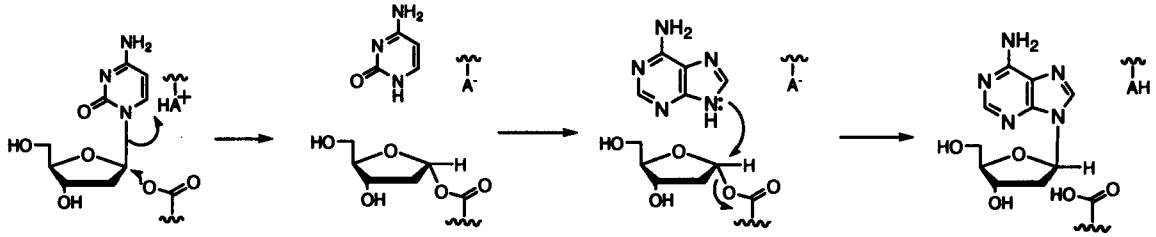
Introduction

Early biosynthetic studies suggested the involvement of free cytosine in blasticidin S (BS) biosynthesis. Radiolabeled precursor feeding studies with [2-^{14}C]cytosine in *S. griseochromogenes* identified cytosine as a primary precursor to BS with incorporation rates as high as 95%.[1] Gould and coworkers were able to show that addition of exogenous cytosine to *S. griseochromogenes* culture media increased BS production 1.6 fold and increased the production of the intermediate cytosylglucuronic acid (CGA) up to 70 fold.[2] These results were not surprising as intracellular concentrations of cytosine are likely quite low because cytosine is not an intermediate in pyrimidine biosynthesis.[3] The stimulation of BS production upon cytosine supplementation indicates that the availability of free cytosine may be a limiting factor in blasticidin biosynthesis.

The most likely source of free cytosine required for BS biosynthesis is the corresponding nucleotides or nucleosides. An example of a specific nucleotide hydrolase with possible involvement in nucleoside antibiotic biosynthesis was reported for the mildiomycin producer, *Streptoverticillium rimofaciens.*[4] A nucleotide hydrolase was partially purified from *S. rimofaciens* and found to catalyze the formation of hydroxymethylcytosine from hydroxymethylcytidine 5'-monophosphate.[4,5] Another pyrimidine nucleotide hydrolase was identified in *Streptomyces virginiae.*[6]

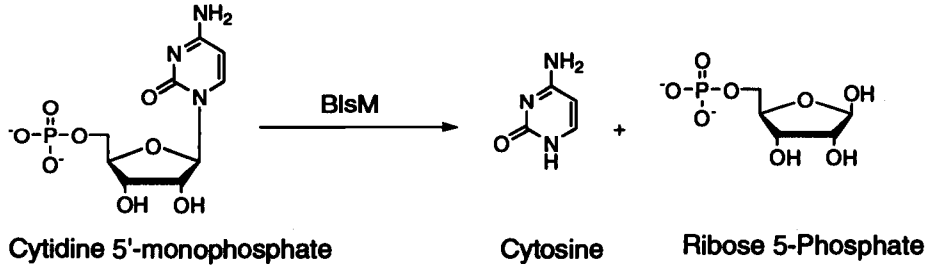
We recently reported the cloning,[7] and the sequencing and partial annotation,[8] of the BS gene cluster from *S. griseochromogenes*. Early experiments supported the presence of a nucleotide hydrolase in the BS gene cluster. Heterologous expression of a fragment of the BS gene cluster in *S. lividans* resulted in a large increase in cytosine levels relative to wild type.[8] Subsequently, sequencing revealed an open reading frame on this fragment, *blsM,* with sequence homology to known nucleoside transferases and

which contains a conserved deoxyribosyltransferases domain (COG 3613, NCBI). The most extensively studied member of this family of proteins is nucleoside 2'-deoxyribosyltransferase (Ndt) from *Lactobacillus leichmannii*.[9] Ndt catalyzes the cleavage of the glycosidic bond in 2'-deoxyribonucleosides resulting in the formation of a ribosylated enzyme intermediate and the free base. The ribose is subsequently transferred from the enzyme to an acceptor base (Scheme 6.1).[9]



Scheme 6.1. Reaction catalyzed by nucleoside 2'-deoxyribosyltransferase

To investigate whether BlsM promoted a similar reaction, we amplified *blsM* from a cosmid containing the blasticidin S gene cluster and expressed it in *E. coli*. In initial characterization, BlsM was found to catalyze the formation of cytosine from cytidine 5'-monophosphate (CMP) in a time and concentration dependent manner (Scheme 6.2).[8] The results of these initial experiments prompted us to fully characterize BlsM.



Scheme 6.2. Reaction catalyzed by BlsM

Results

Sequence Analysis of blsM

blsM codes for a 174 amino acid protein with a calculated molecular weight of 18.9 kDa (Figure 6.1). The probable ribosome binding site (GAGGG) is located six nucleotides upstream from the start codon. A second potential start site is located 42

nucleotides downstream of the proposed translational start site and was used for the initial expression of *blsM*.[8] However the second possible start site does not have a clearly associated ribosome binding site. A definite promoter region has not been determined experimentally, but *bls*M appears to be transcribed on a polycistronic mRNA as no putative promoter regions have been located proximal to *bls*M. The nearest potential promoter region is located upstream of *bls*D with (-10: CATGGT -35: TTGA) and is in good agreement with consensus *Streptomyces* promoters.[10,11] Other notable features of *blsM* include a GTG initiation codon and the use of a TTA codon which makes translation of BlsM dependent on the presence of a specific $tRNA^{Leu}$ coded by the *bld*A gene.[12] The *bldA* gene product has been shown to coordinate antibiotic biosynthesis with the development cycle in many *Streptomyces* species and a similar use is found in the puromycin pathway.[12,13,14]

```
tgaggg aacctcgtgaacgtcatcagcagtgcgtcggagtccgtcctcgacggcgtgcgc
            V  N  V  I  S  S  A  S  E  S  V  L  D  G  V  R
agcgtctttctcgccggtcccttcatgggactggtcaaccccgagaccaacagcatgccg
 S  V  F  L  A  G  P  F  M  G  L  V  N  P  E  T  N  S  M  P
agcgccgaacagcttcccttcctcacactcatcgagcacttcgagaaacagggtctggag
 S  A  E  Q  L  P  F  L  T  L  I  E  H  F  E  K  Q  G  L  E
gtcttcaacgcccaccggcgggaggcctggggtgcgcaggtgctgacaccggaggagtgc
 V  F  N  A  H  R  R  E  A  W  G  A  Q  V  L  T  P  E  E  C
acgcccctggaccagctggagatccgcaaggcggacgtcttcgtcgccatcccgggcatc
 T  P  L  D  Q  L  E  I  R  K  A  D  V  F  V  A  I  P  G  I
ccgccctcccccggtacccatgtcgagatcggctgggcgagcgcgttcgacaaacccatc
 P  P  S  P  G  T  H  V  E  I  G  W  A  S  A  F  D  K  P  I
gtgctcttactggaggagggcagggaggaggagtacggctttctcgtccgcggtctcggt
 V  L  L  L  E  E  G  R  E  E  E  Y  G  F  L  V  R  G  L  G
accgtcgccgcggtcgagttcgtccactacaaggacatcgccctggcgaaaccgcagatc
 T  V  A  A  V  E  F  V  H  Y  K  D  I  A  L  A  K  P  Q  I
gacgcggcgatccgcaaagtggtggaccgagtgaacaacccggccgcaacgccctgaccg
 D  A  A  I  R  K  V  V  D  R  V  N  N  P  A  A  T  P  Stop
```

Figure 6.1. Sequence of BlsM. RBS is underlined.

Expression and Purification of BlsM

In order to determine the function of BlsM, *bls*M was amplified from the cos9 and cloned into a pET41a+ expression vector to form p41*bls*M. Heterologous expression of p41*bls*M in *E. coli* resulted in the fusion of tandem N-terminal glutathione-S-transferase

(GST) and S-tag proteins. Two successive purification steps were used to first remove the GST-fusion protein by thrombin cleavage followed by enterokinase cleavage to yield BlsM (Figure 6.2). The initial characterization of BlsM was conducted with a truncated form of BlsM, BlsM160, which lacks fourteen amino acids from the N-terminus. The translational start site for BlsM was redefined upon reevaluation of the sequence data and the construct reported here is the full length protein.

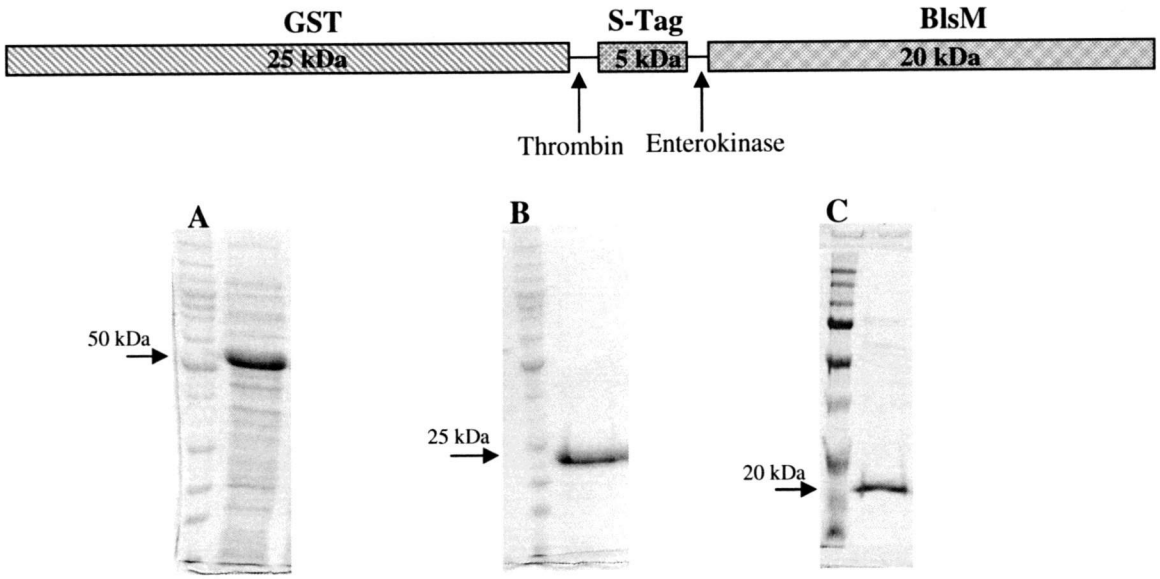


Figure 6.2. Expression and purification of BlsM from *E. coli*. A. Cleared cell lysate containing GST/S-tag/BlsM fusion protein. B. S-tag/BlsM following thrombin cleavage of GST. C. BlsM following cleavage with enterokinase.

Substrate Specificity and Activity

BlsM was assayed for hydrolase and transferase activity using a variety of possible nucleoside and nucleotide substrates. BlsM was found to catalyze the formation of cytosine from both CMP and dCMP (Figure 6.3), but showed no detectable transferase activity when adenine, thymine, uridine and guanine were used as acceptor bases. In addition, BlsM was found to have low levels of activity with CDP and CTP, exhibiting 29% and 7% conversion of substrate to product, respectively, relative to CMP hydrolysis (Table 6.1). Although the CDP and CTP used in the assays were freshly purchased from Sigma, the possible contamination with small amounts of CMP can not be ruled out, and may account for the low level of activity that was observed. No detectable CMP was

formed when BlsM was incubated with cytosine and ribose-5-phosphate at concentrations up to 20 mM, suggesting that CMP hydrolysis is not reversible. Product inhibition was not observed when BlsM was incubated with either cytosine or ribose-5-phosphate at concentrations up to 100 μM. No hydrolytic activity was detected when AMP, TMP, GMP, UMP, cytidine, deoxycytidine, or cytosine β-D-arabinofuranoside 5'-monophosphate were evaluated as substrates.

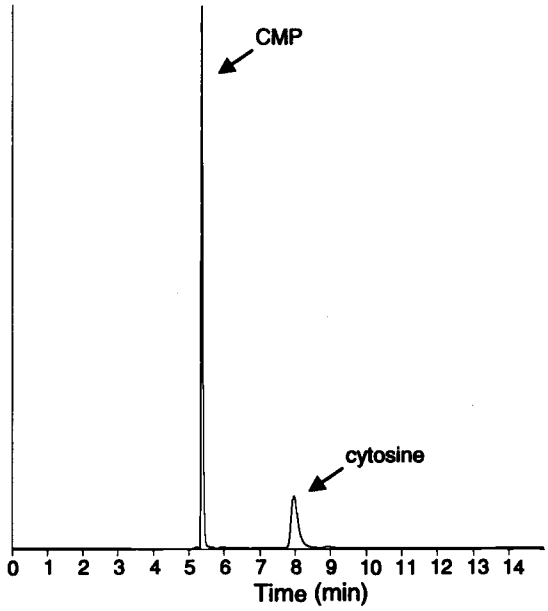


Figure 6.3. HPLC chromatogram showing formation of cytosine from CMP by BlsM.

Table 6.1. Formation of Cytosine by BlsM. Percent conversion of substrate to cytosine, relative to the amount of cytosine formed from CMP

Substrate	Relative % Conversion
CMP	100
dCMP	45
CDP	29
CTP	7

The kinetic parameters for BlsM with both CMP and dCMP substrates were determined (Figure 6.4). BlsM had a significantly higher K_m, and lower k_{cat} and k_{cat}/K_m values with dCMP as a substrate compared to CMP (Table 6.2). As mentioned previously, the initial characterization of BlsM was conducted with a fourteen amino acid

truncation, BlsM160. The kinetic parameters of BlsM160, which contained both a 5 kDa *N*-terminal S-tag fusion protein and a *C*-terminal His_6-tag, were nearly identical to full length BlsM (Table 6.2).

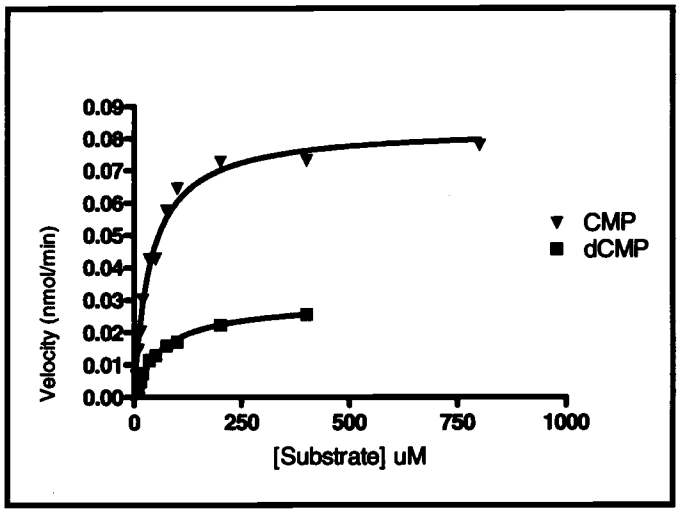


Figure 6.4. Cytosine formation by BlsM. Velocity of cytosine formation by BlsM with CMP (▼) and dCMP (■) as substrates.

Table 6.2. Kinetic parameters observed for BlsM with CMP and dCMP substrates

Substrate	V_{max} (nM/min)	K_m (μM)	k_{cat} (min^{-1})	k_{cat}/K_m [$(\mu mol/L)^{-1}$ min^{-1}]
CMP	0.084±0.003	39±4	$2x10^{-4}$	$4x10^{-6}$
dCMP	0.030±0.001	65±6	$6x10^{-5}$	$9x10^{-7}$
BlsM160 with CMP	0.100±0.005	22±4	$4x10^{-6}$	$1.8x10^{-7}$

Physical Properties of BlsM

The optimal pH for BlsM catalysis was found to be 7.4-7.6. This pH range is consistent with that observed for the pyrimidine 5'-nucleotide hydrolase from *S. virginae*.[6] BlsM was found to have higher activity in Tris buffers than in phosphate buffers. This buffer effect may be due to partial competitive inhibition resulting from binding of the phosphate ion to the BlsM active site (Figure 6.5, A).

The maximum rate of turnover for BlsM was seen at 52°C (Figure 6.5, B). While the optimal temperature for BlsM is relatively high, it is consistent with our initial observations regarding the heat stability of the protein. We found it necessary to boil the enzyme for at least 10 min to completely eliminate hydrolase activity. This heat stability

was also seen with pyrimidine 5'-nucleotide hydrolase from *S. virginiae* which was unaffected by heating to 65°C for 45 min.[6] BlsM was stable and retained equal activity when stored at -80 °C, -20 °C or 4 °C over several months.

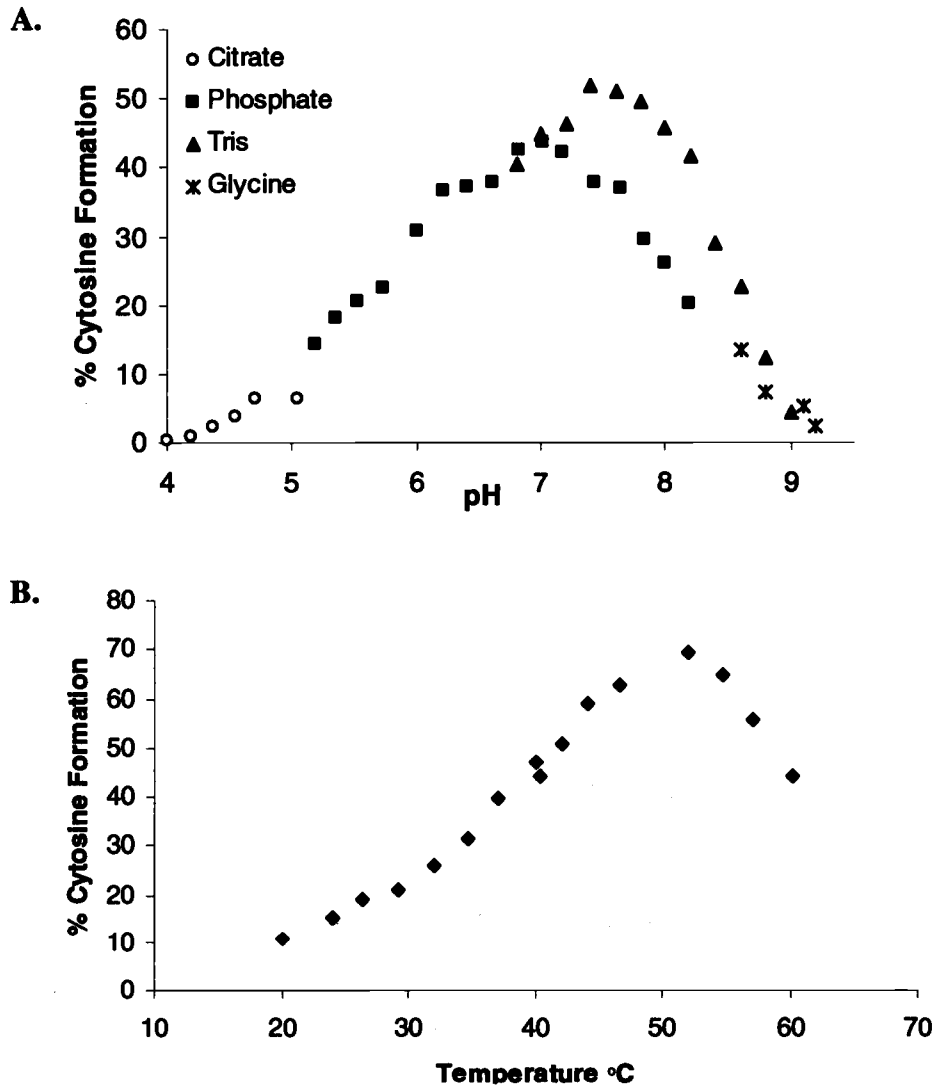


Figure 6.5. Effect of pH and temperature on cytosine formation by BlsM. A. Percent conversion of CMP to cytosine at pH 4-9.5. B. Percent conversion of CMP to cytosine at various reaction temperatures

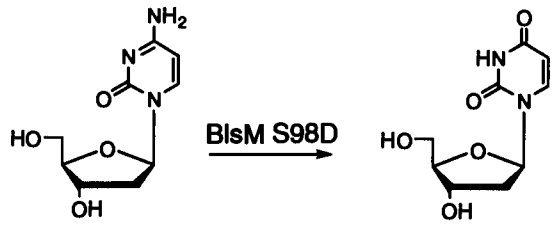
Molecular Weight Determination

Gel filtration showed that recombinant BlsM, as well as the S-tag fusion protein, exists primarily as a trimeric protein with a small proportion of the enzyme (~3%) existing as a hexamer.

Site Directed Mutagenesis

In order to investigate the role of conserved active site residues on substrate specificity and reaction mechanism, three point mutations in the *blsM* were generated through site directed mutagenesis. Two residues were selected for mutation; glutamate 104 and serine 98. Glutamate 104 was chosen because it corresponds to Ndt Glu98, which was shown to be a key catalytic residue.[15] Serine 98 was chosen for mutation because the corresponding site in many nucleotide hydrolases (Ndt Asp92) contains a conserved aspartate residue proposed to act as a general acid/base in catalysis.[16] Three mutants were created, E104A, E104D, and S98D, and all showed complete loss of CMP hydrolysis activity and did not show hydrolase activity with any of the alternate substrates tested.

Interestingly, each of the three mutants was found to have varying levels of an alternate activity when incubated with either cytidine or deoxycytidine. HPLC analysis of assays with each of the three mutants showed the formation of new peaks, depending on substrate, that did not correspond to cytosine. The new peaks were identified as uridine and deoxyuridine by LC-MS (Figure 6.6, Scheme 6.3). The identity of the compounds was further confirmed by comparison of the UV and mass spectral data with authentic samples and by coinjection. The S98D mutant showed the greatest level of deaminase activity while the other mutants and wild type BlsM exhibited 15-28% deaminase activity relative to BlsM S98D. This deamination reaction was found to be both time and enzyme concentration dependent and was abolished by boiling the enzyme for ten minutes (Figure 6.7). EDTA concentrations of up to 1 mM did not affect S98D deaminase activity.



Deoxycytidine Deoxyuridine

Scheme 6.3. Formation of deoxyuridine by BlsM S98D

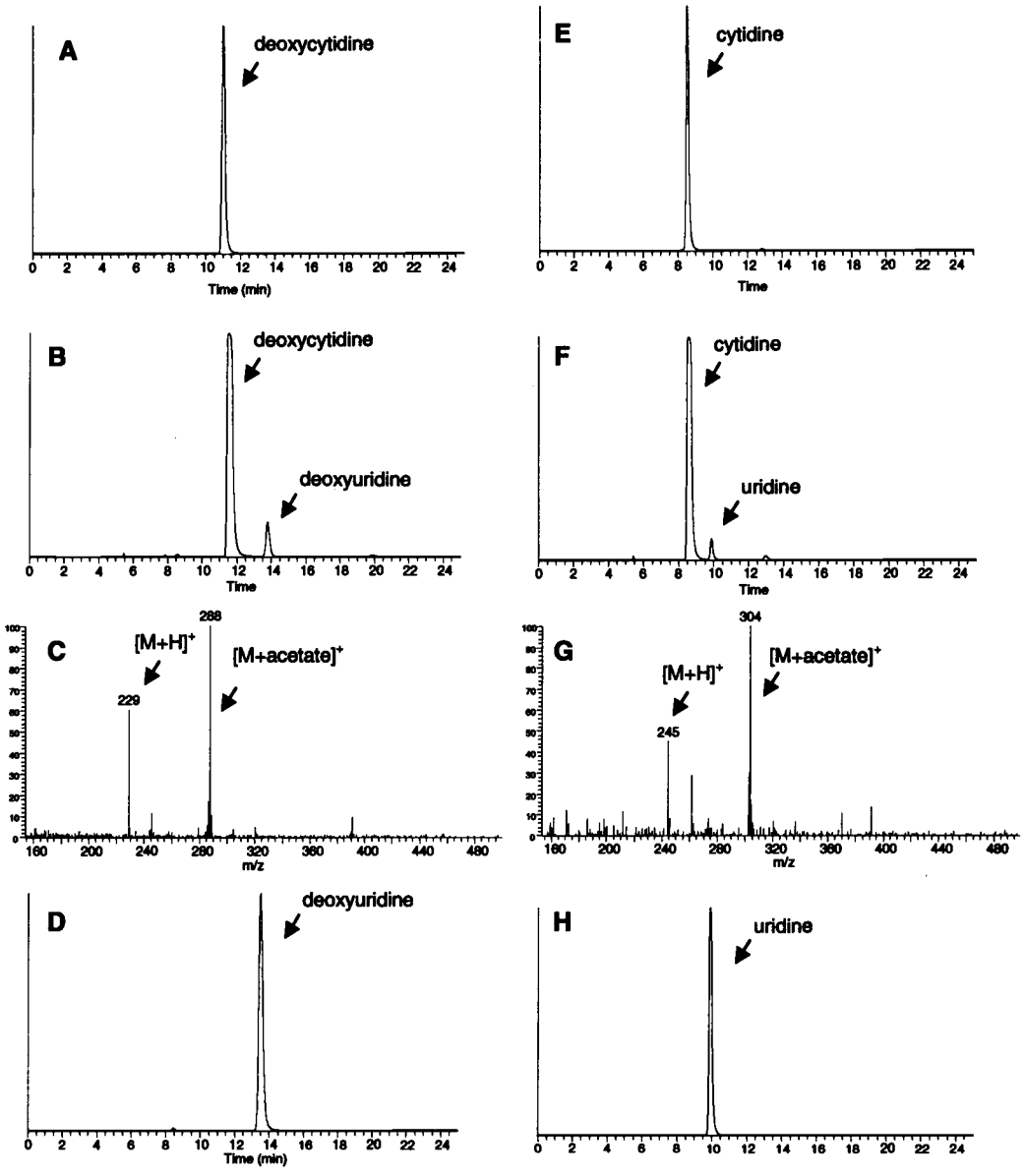


Figure 6.6. Formation of deoxyuridine and uridine from deoxycytidine and cytidine by BlsM S98D. A,E. HPLC chromatogram of deoxycytidine and cytidine standards. B. HPLC chromatogram of BlsM S98D assay with deoxycytidine. C. Mass spectrum of product, corresponding to the LC peak eluting at 13.5 min, from BlsM S98D assay with deoxycytidine. The $[M+H]^+$ peak corresponding to deoxyuridine is indicated. A peak which corresponded to the molecular ion plus acetate (from assay buffer) is also indicated. D. HPLC chromatogram of deoxyuridine standard. F. HPLC chromatogram of BlsM S98D assay with cytidine. G. Mass spectrum of uridine product, corresponding to the peak eluting at 10 min, for BlsM S98D assay with cytidine as substrate. The $[M+H]^+$ and $[M+acetate]^+$ peaks are indicated. H. HPLC chromatogram of uridine standard.

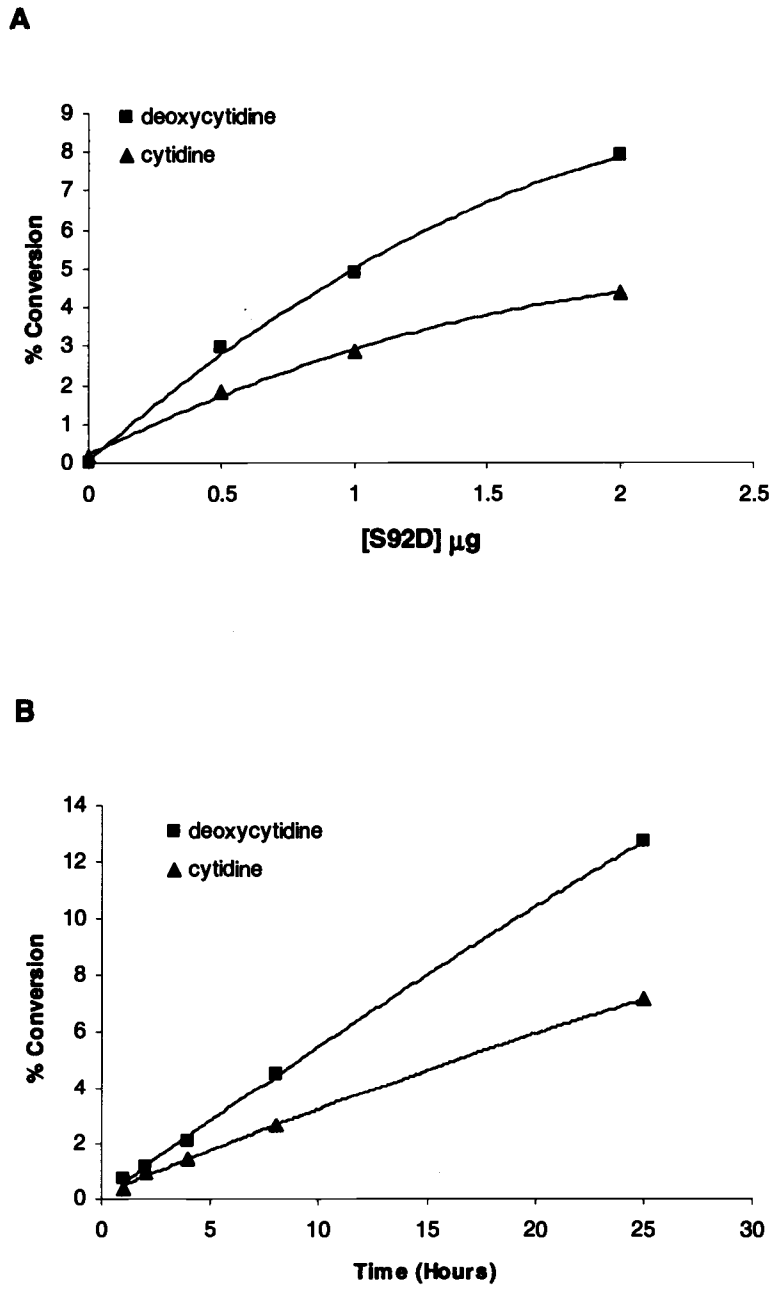


Figure 6.7. Formation of uridine and deoxyuridine by BlsM S98D. A. Increased product formation as a function of BlsM S98D concentration. B. Time dependent formation of uridine and deoxyuridine.

Discussion

In our preliminary analysis of the blasticidin S gene cluster, BlsM was identified as a CMP hydrolase.[8] Initial work with BlsM was conducted with a truncated form of BlsM, BlsM160, that lacked fourteen amino acids from the N-terminus. Reevaluation of the gene sequence identified a better candidate for the translational start 42 nucleotides upstream. A probable ribosomal binding site (GACGG) is located six nucleotides upstream from the start codon.

The protein sequence in the public databases most similar BlsM is to a hypothetical protein with unknown function from *Leuconostoc mesenteroides* (gi:23024599) sharing 40% identity and 59% similarity. The key to elucidating the putative function of BlsM came from analysis of conserved domains. BlsM contains the conserved deoxyribosyltransferase motif (COG 3613, NCBI) found in several nucleoside 2'-deoxyribosyltransferases (Figure 6.8).[17] The most studied protein in this group, nucleoside 2'-deoxyribosyltransferases (Ndt) from *Lactobacillus leichmannii* catalyzes the cleavage of nucleoside *N*-glycosidic bonds with the formation of a covalent deoxyribosyl-enzyme intermediate and subsequent transfer of the deoxyribose to an acceptor base, thus serving to recycle the sugar moiety (Scheme 6.1).[15]

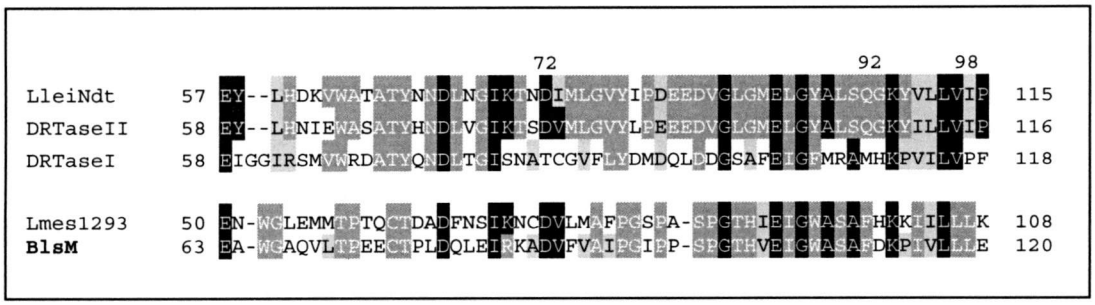


Figure 6.8. Alignment of BlsM with known nucleoside 2'-deoxyribosyltransferase active sites. Residues that are highly conserved are highlighted in black. Residues that are conserved in a subgroup of enzymes are highlighted in dark grey. Similar residues are highlighted in light grey. *L leiNdt* is from *L. leichmannii*; *DRTaseI* and *DRTase*II are from *L. helveticus CNRZ32*; and *Lmes 1293* is from *L. mesenteroides*.

Although Ndt is specific for 2'-deoxynucleoside donor substrates, it exhibits relaxed substrate specificity in its use of acceptor bases and has been used for the

enzymatic synthesis of dideoxy and 2'-deoxynucleosides.[18-21] Armstrong *et al* were unable to explain this specificity based on the crystal structure of Ndt, as the structure suggested that there was adequate room for binding of a ribosyl nucleoside.[22] They proposed that electrostatic interactions between the substrate 2'-OH group and the Glu98 nucleophile may be involved.[22]

Unlike Ndt which recognizes 2'deoxynucleosides, BlsM has rigid substrate specificity for cytidine 5'-monophosphate nucleotides, with a strong preference for CMP over dCMP. BlsM showed only minor activity with the cytidine nucleotides CDP and CTP (Table 6.1). With regard to substrate specificity, BlsM shows the greatest similarity to a nucleotide hydrolase partially purified from *S. rimofaciens* that catalyzes the formation of hydroxymethylcytosine from the ribose nucleotide, 5-hydroxymethylcytidine 5'-monophosphate.[4] The *S. rimofaciens* enzyme is believed to be involved in the biosynthesis of the antibiotic mildiomycin that contains a hydroxymethylcytosine moeity.[5] Much like BlsM, the *S. rimofaciens* enzyme would also utilize the corresponding 2'-deoxynuclrotide as substrate with significantly lower activity (approximately 27%) relative to hydroxymethylCMP.[4] CMP hydrolysis by BlsM is apparently not reversible in as much as no detectable CMP was formed even after incubation with concentrations of cytosine and ribose-5-phosphate up to 20 mM for six hours. The reverse reaction was also not detected for the *S. virginiae* enzyme.[6]

Site-directed mutagenesis was used to generate three point mutations in *blsM*. The targeted residues were chosen based on studies with the *L. leichmannii* Ndt. The reaction mechanism for Ndt was shown to be analogous to retaining glycosidases and to proceed through a covalent ribosylated enzyme intermediate.[15] This was demonstrated through the use of the mechanism based inactivator 2,6-diamino-9-(2'-deoxy-2'fluoro-β-D-arabinofuranosyl)-9H-purine (dFDAP), which irreversibly binds to the active site nucleophile Glu98.[15] The corresponding residue in BlsM, Glu104 (Figure 6.8), was specifically mutated to both an alanine and an aspartate residue to investigate the importance of this group in catalysis. Both the BlsM E104A and E104D mutants showed complete loss of CMP hydrolysis activity. The results were unexpected as similar Ndt mutants retained low levels of catalytic activity.[16]

By analogy to well studied glycosidases, Ndt was proposed to use one or more active site amino acids as general acids/bases. Based on the crystal structure of Ndt, Short and colleagues were able to identify several active site amino acids that may perform this function.[16] Through the characterization of several Ndt mutants, they found that Asp92 is important for catalysis and may be functioning as a general acid/base. Sequence alignment of BlsM with the conserved deoxyribosyl transferase domain, COG3613, of Ndt and other nucleoside transferases showed that BlsM and the nucleoside hydrolase from *L. mesenteroides*, Lmes1293, both contain a serine at this position rather than an aspartate (Figure 6.8). The crystalographic structure of Ndt suggested that Asp92 may be involved with binding the sugar 5'-OH and may help to orient the substrate for optimal catalysis.[22] This hypothesis was supported by the observation that dideoxynucleosides were adequate sugar donors but had decreased rates in the transferase reaction.[22] Because transferase activity was not observed with wild type BlsM we reasoned that changing the corresponding serine 98 in BlsM, corresponding to Ndt position 92, to an aspartate could alter the substrate specificity and/or convert BlsM to a nucleoside transferase.

Similar to the results from the mutation of Glu104, the BlsM S98D mutant did not retain CMP hydrolysis activity. However, a new peak not corresponding to cytosine was observed in the HPLC chromatograms. BlsM S98D was shown to catalyze the formation of uridine and deoxyuridine from cytidine and deoxycytidine, respectively, thus functioning as a deaminase. This finding demonstrates the importance of this residue for BlsM activity.

The native substrate for BlsM is the ribonucleotide rather than the nucleoside (Figure 6.9). The new activity profile of BlsM S98D suggests that substrate selectivity has been altered through this amino acid substitution. The S98D mutant shows no hydrolysis activity with CMP as substrate which may reflect a decrease in the ability of the phosphorylated substrate to bind properly to the active site when an anionic residue occupies position 98. Accordingly, the observed deamination occurs only with nucleoside substrates. This finding is consistent with the suggestion that the corresponding residue in Ndt is involved in substrate selection and binding. Additional

support for the involvement of Ndt Asp92 in substrate binding was recently provided by Anand *et al.* who reported the X-ray crystal structures of purine deoxyribosyltransferase (DRTase I) from *L. helveticus* with bound ligands.[23] The crystal structures of DRTase I with bound deoxynucleoside as well as the ribosylated intermediate show the formation of a key hydrogen bond between the ribose 5'-OH and the corresponding Asp95 (Ndt Asp92). The observed deamination of cytidine and deoxycytidine by BlsM S98D may also be a consequence of altered substrate binding. The substitution of the serine for an aspartate may accommodate the binding of cytidine and deoxycytidine, but it apparently does not allow for the proper substrate orientation to allow for sugar hydrolysis while it does permit an alternate hydrolysis reaction similar to that of cytidine deaminase.[24]

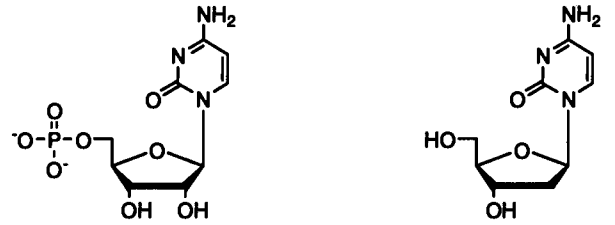
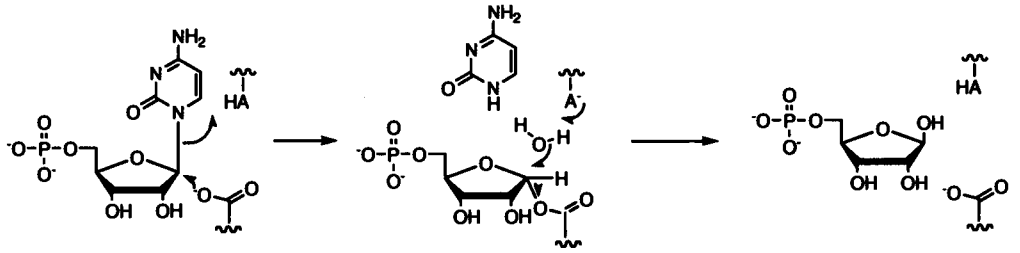


Figure 6.9. Preferred substrate for BlsM hydrolysis, CMP, and preferred substrate for BlsM S98D deamination, deoxycytidine.

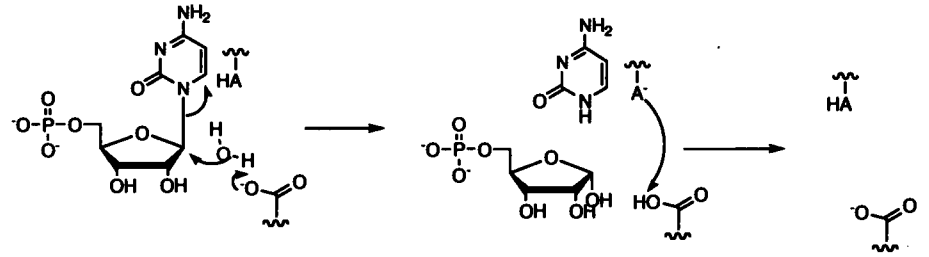
Cytidine deaminase (CDA) is an important enzyme in the pyrimidine salvage pathways of many microorganisms and catalyzes the hydrolytic deamination of cytidine to uridine.[3,24] Genes encoding CDA are being used in gene therapy treatment of prostate cancer currently in phase I clinical trials.[25] Cytidine deaminase is related to nucleoside transferases and BlsM in that it is also a nucleotide binding protein that catalyzes the hydrolysis of its substrate. The fundamental difference between the two types of enzymes is the site of hydrolysis. The catalytic mechanism of CDA has been elucidated and the deamination of cytidine is believed to proceed through a step-wise process that involves an initial addition of hydroxide followed by elimination of ammonia (Scheme 6.4).[26] This mechanism involves the activation of water by zinc to create an hydroxide

nucleophile that is stabilized by an active site glutamate residue. The glutamate residue has been proposed to perform a dual role that involves the concomitant protonation of cytidine N-3 as well as hydroxide stabilization.[26]

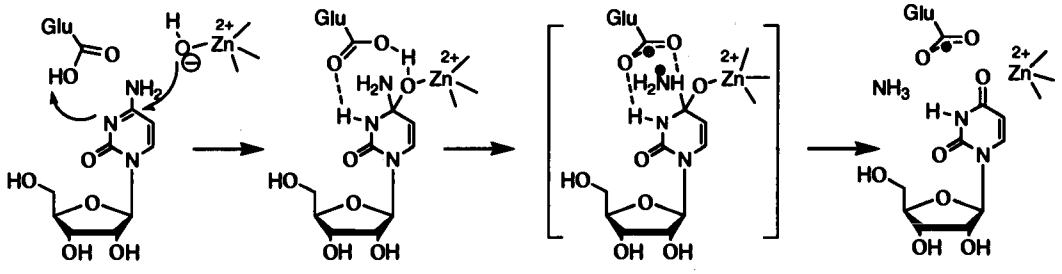
A. Retaining Mechanism



B. Inverting Mechanism



C. Cytidine deaminase



Scheme 6.4. A, B. Possible reaction mechanisms for BlsM . A. Retaining mechanism used by Ndt nucleosidetransferase.[9] B. Inverting mechanism based on glycosidase chemistry. [27] C. Proposed mechanism for cytidine deaminase.[26]

Two possible reaction mechanisms for the hydrolysis of BlsM are shown in Scheme 6.4 and are analogous to well characterized glycosidase mechanisms that result in either the retention or inversion of stereochemistry at the anomeric carbon.[27] Both mechanisms involve the activation of water and subsequent attack of hydroxide ion

analogous to the role performed by the zinc ion in CDA. Similarly, another active site general base, perhaps the introduced aspartate of the BlsM S98D mutant, may be analogous to Glu104 of CDA and serve to protonate N-3 of cytidine as well as activate the attacking water molecule. Alternatively, the S98D mutation may simply serve to orient the nucleoside substrate in a manner that is not aligned for sugar hydrolysis but does lead to the observed low levels of amine hydrolysis.

It is likely that differences exist in either substrate binding or reaction mechanism between BlsM and nucleoside deoxyribosyl transferases. The crystallographic studies reveal the importance of the key catalytic residue, Glu98, in the specificity of Ndt for deoxynucleosides.[23] This is due to the formation of a hydrogen bond between Glu98 and the ribose 2'OH, thus preventing the nucleophilic attack and subsequent formation of the ribosylated enzyme intermediate. The formation of the glycosyl intermediate is important for transferase function as it retains the sugar in a form that can be attacked by an acceptor base. The role of BlsM in blasticidin S biosynthesis is the production of free cytosine, which does not require maintaining an activated sugar. If BlsM substrate binding is similar to Ndt, the results of Anand's study would suggest that CMP hydrolysis by BlsM may not involve the formation of a covalent intermediate.[23]

Additional studies are necessary to provide insight into the catalytic mechanism of CMP hydrolysis as well as the origin of the deaminase activity. Although the primary sequence of BlsM shares significant similarity with active sites of nucleosidetransferases, our results suggest the structure of BlsM differs from that of deoxynucleoside tranferases. Gel filtration showed that recombinant BlsM appears primarily as a trimer with a small proportion of the enzyme (~3%) existing as a hexamer. This finding differs from the characterized deoxyribosyltransferases that are typically hexameric proteins composed of a trimer of catalytic dimers.[28] Studies on the other *Streptomyces* nucleotide hydrolases reported in the literature were conducted with partially purified enzymes and the multimeric forms of these adopted by these proteins is not known. Further studies into the catalytic mechanisms of BlsM should contribute the understanding of both nucleotide hydrolases and deoxynucleoside transferases.

In conclusion, BlsM is a nucleotide hydrolase functioning in a secondary metabolic pathway and is responsible for the formation of free cytosine required for the biosynthesis of blasticidin S. The optimal pH range, heat stability and substrate selectivity of BlsM are more similar to partially characterized nucleotide hydrolases identified from *S. virginiae* and *S. rimofaciens,* than to the well studied deoxynucleosidetransferases involved in primary metabolism. Two active site residues, Glu104 and Ser98, are critical for catalytic activity and Ser98 appears to have a key role in substrate selectivity. Substitution of an aspartate at this position results in the conversion of BlsM to a cytidine deaminase.

Materials and Methods

Preparation of blsM Expression Construct

PCR was carried out in a total volume of 50 μL containing 10 ng cos9 template,[7] 1X Promega Thermophilic Buffer, 1.25 mM $MgCl_2$, 0.4 mM dNTP mix (MBI), 5% dimethylsulfoxide, 50 pmol forward primer BlsMst1-*EcoR1*For 5'ctcgagggaattcggtg aacgtcatcagcagt3'(*Eco*RI site is underlined), 50 pmol reverse primer BlsMstp-*Xho*Rev 5'cggggcacgatgcggtgcacgctcgagtcagggcg3' (*Xho*I site is underlined), and 5 units Platinum Taq DNA polymerase (Invitrogen). Gel purified PCR products were digested with *Eco*RI and *Xho*I then ligated into a similarly restricted pET41a+ (Novagen). The resulting plasmid, pET41BlsM, was used to transform *E.coli* JM109 cells (Promega) and the correct construction was confirmed by sequencing. pET41BlsM was used to transform *E.coli* Rosetta (DE3)pLysS competent cells (Novagen) for expression.

Overexpression and Purification of Recombinant BlsM

Transformants carrying pET41BlsM were grown overnight in Terrific broth supplemented with chloramphenicol and kanamycin (50 μg/mL each). This seed culture was used to inoculate one liter of Terrific broth supplemented with both antibiotics. The cells were grown at 37 °C to an OD_{600}= 0.55 and then protein expression was induced with IPTG at a final concentration of 0.4 mM. The culture was then grown for an additional 8 hours at 37 °C.

Cells were harvested by centrifugation, resuspended in 50 mL phosphate buffered saline (PBS: 140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.0 mM KH_2PO_4, pH 7.3) and lysed by sonication in a Microson ultrasonic cell disruptor (eight 60 s bursts at 6 watts output with 60 s cooling on ice between bursts). Cellular debris was pelleted by centrifugation at 18000 *x g* for 40 min at 4 °C.

Soluble GST-BlsM fusion protein was purified from the cleared lysate by binding to a glutathione affinity column (GST-Bind, Novagen) and the resin-bound complex was treated with thrombin (10 units for 18 h at room temperature) to cleave BlsM from the GST fusion protein. Resulting protein contained a 5 kDa S-tag fused to the N-terminus that was removed by enterokinase digestion (0.05 units/μg for 16 h at room temperature). Cleaved S-tag peptide and undigested fusion protein was removed from the mixture by binding to S-tag agarose (Novagen). Enterokinase was removed with Enkapture agarose (Novagen) according to the manufacturer's protocol. Purified BlsM was stored at 4 °C in PBS (pH 7.3) with 15% glycerol and was stable for several months at 4 °C or -20 °C.

Site Directed Mutagenesis of blsM

Site directed mutagenesis was used to change two putative active site amino acids in BlsM. Mutagenesis was carried out using the Quick Change Mutagenesis Kit (Stratagene) according to the manufacturer's protocol. Primers were designed to alter the codon for the selected amino acid as well as to incorporate a silent mutation that introduced a new restriction site (E104A, E104D) or to remove an existing restriction site (S98D) as indicated in Table 1. PCR was carried out in a total volume of 50 μL containing 25 ng pET41BlsM template, 1X Quick Change reaction buffer, 1 μL dNTP mix, 125 ng each forward and reverse primers, and 2.5 units *Pfu Turbo* DNA polymerase. Following restriction of the parental plasmid strands with *Dpn*I, PCR products were used to transform *E.coli* XLI-Blue competent cells (Stratagene). Resulting plasmids were screened for the correct alteration in restriction patterns with either *Eco47*III or *Sma*I. Plasmids showing the correct restriction patterns were sequenced to confirm mutation and used to transform *E.coli* Rosetta (DE3)pLysS competent cells (Novagen) for expression and purification as described for wild type BlsM.

Table 6.3. Primers used for site directed mutagenesis of BlsM. Codons altered for the generation of point mutations are indicated in bold. Restriction sites that were created or deleted to facilitate identification of mutants are underlined.

Primer	Sequence	Restriction Site
BlsM-E104Dfor	5′ GT ACC CAT GTC **GAT** ATC GGC TGG GCG AGC GCT TTC GAC AAA CC 3′	*Eco47 III* (created
BlsM-E104Drev	5′ GG TTT GTC GAA AGC GCT CGC CCA GCC GAT **ATC** GAC ATG GGT AC 3′	*Eco47 III* (created
BlsM-E104Afor	5′ GT ACC CAT GTC **GCG** ATC GGC TGG GCG AGC GCT TTC GAC AAA CC 3′	*Eco47 III* (created
BlsM-E104Arev	5′ GG TTT GTC GAA AGC GCT CGC CCA GCC GAT **CGC** GAC ATG GGT AC 3′	*Eco47 III* (created
BlsM-S98Dfor	5′ TTC GTC GCC ATC CCT GGC ATC CCG CCC **GAC** CCC GGT ACC C 3′	*Sma*I (removed)
BlsM-S98Drev	5′ G GGT ACC GGG **GTC** GGG CGG GAT GCC AGG GAT GGC GAC GAA 3′	*Sma*I (removed)

Identification of BlsM Substrate

Various nucleotides and nucleosides were evaluated as possible substrates for BlsM. Assays were carried out in a total volume of 100 μL containing 50 mM Tris buffer pH 7.2, 0.1 mM or 1 mM substrate and 1 μg BlsM. Assays were incubated for 1 h at 37 °C then quenched by boiling for 10 min. Following centrifugation (5 min at 18000 *x g*), product and substrate were separated by HPLC (Agilent Zorbax-C_8 5 μm, 4.6 x 250 mm column; isocratic mobile phase 5% MeOH in H_2O, 0.5 mL/min) with UV detection at 265 nm. For each assay, 50 μL was loaded onto a 100 μL injection loop for HPLC analysis. Formation of cytosine was confirmed by comparison of retention time and UV spectra with authentic material.

Nucleotide Transferase Assays

Nucleotide transferase assays included 0.1, 0.3 or 1 mM of various purine or pyrimidine bases added to the standard BlsM assay described above. Assays were incubated for 2 h at 37 °C then quenched by boiling for 10 min. Following centrifugation (5 min at 18000 *x g*), product and substrate were separated by HPLC.

Steady State Kinetic Analysis of BlsM

Assays used to determine BlsM kinetic parameters were carried out in a total volume of 100 μL containing 50 mM Tris buffer pH 7.2, 0.5 μM BlsM, and 5-800 μM CMP. Reactions were incubated for 30 min at 37 °C then quenched by boiling for 10

min. 50 μL of each reaction mixture was analyzed by HPLC. Reaction velocity was determined by measuring cytosine peak area and subsequent conversion to nmol cytosine produced based on a standard curve generated for cytosine. Prism4 software (GraphPad Software, Inc.) utilizing nonlinear regression analysis was used to calculate kinetic parameters.

Determination of the Optimal Temperature and pH for BlsM Activity

BlsM assays were carried out in a total volume of 100 μL containing 50 mM Tris buffer pH 7.2, 0.1 mM CMP and 0.5 μM BlsM. Reactions were incubated for 30 min at temperatures ranging from 20-60°C to determine the optimal temperature. Reaction buffers used to identify the optimal pH included citrate (pH 4-5.2), phosphate (pH 5.2-8.2), bis-Tris (pH 5.8-6.6), Tris-HCl (pH 6.8-9) and glycine (pH 8.6-9.2).

Molecular Weight Analysis

The native molecular weight of BlsM was determined by gel filtration chromatography performed on an Amersham Biosciences ÄKTA FPLC with a Superdex 200 HR10/30 column. The mobile phase consisted of 50 mM Tris buffer pH 7.4 with 0.5 mM dithiothreitol (DTT) at a flow rate of 0.45 mL/min. Protein standards were purchased from Sigma or Calbiochem. BlsM (80 ng) with and without the N-terminal S-tag fusion was analyzed following the generation of a molecular weight standard curve.

Substrate and Product Analysis of BlsM S98D

BlsM S98D was assayed with cytidine or deoxycytidine as substrate in a total volume of 100 μL containing 50 mM Tris buffer pH 7.2, 0.1 mM dCMP and 0.5 μM enzyme. Assays were incubated for 18 h at 37 °C, quenched by boiling for 10 min and then analyzed by LC/MS (Agilent Zorbax-C_8 5 μm, 4.6 x 250 mm column; isocratic mobile phase containing 5% aqueous MeOH at a flow rate of 0.5 mL/min) on a ThermoFinnigan LCQ Advantage system using atmospheric pressure ionization (APCI). For each assay, 50 μL was loaded onto a 100 μL injection loop.

References

(1) Seto, H.; Yamaguchi, N.; Otake, N.; Yonehara, H. Studies on the biosynthesis of blasticidin S. Part I. Precursors of blasticidin S biosynthesis. *Agric. Biol. Chem.* **1968**, *32*, 1292-1298.

(2) Gould, S. J.; Guo, J. Cytosylglucuronic acid synthase (cytosine: UDP-glucuronosyltransferase) from *Streptomyces griseochromogenes*, the first prokaryotic UDP-glucuronosyltransferase. *J. Bacteriol.* **1994**, *176*, 1282-1286.

(3) Munch-Petersen, A. *Metabolism of Nucleotides, Nucleosides and Nucleobases in Microorganisms*; Academic Press Inc.: San Diego, 1983.

(4) Sawada, H. S., T.; Akiyama, SI.; and Nakao, Y Biosynthetic Pathway of 5-Hydroxymethylcytosine by *Streptoverticillium rimofaciens*. *J. Ferment. Technol.* **1985**, *63*, 23-27.

(5) Sawada, H. S., T.; Akiyama, SI.; and Nakao, Y Formation of 5-Hydroxymethylcytosine from Cytidine 5'-Mono-Phosphate by *Streptoverticillium rimofaciens*. *J. Ferment. Technol.* **1985**, *63*, 17-21.

(6) Imada, A. K., M.; and Igarasi, S. Degradation of Pyrimidine Nucleotides by Enzyme Systems of *Streptomyces* II. Pyrimidine 5'-Nucleotide Phosphoribo (Deoxyribo) Hydrolase of *Streptomyces virginiae*. *J. Gen. Appl. Microbiol.* **1967**, *13*, 267-278.

(7) Cone, M. C.; Petrich, A. K.; Gould, S. J.; Zabriskie, T. M. Cloning and heterologous expression of blasticidin S biosynthetic genes from *Streptomyces griseochromogenes*. *J. Antibiot. (Tokyo)* **1998**, *51*, 570-578.

(8) Cone, M. C.; Yin, X.; Grochowski, L. L.; Parker, M. R.; Zabriskie, T. M. The blasticidin S biosynthesis gene cluster from *Streptomyces griseochromogenes*: sequence analysis, organization, and initial characterization. *Chembiochem* **2003**, *4*, 821-828.

(9) Porter, D. J.; Short, S. A. Nucleoside 2-deoxyribosyltransferase. Pre-steady-state kinetic analysis of native enzyme and mutant enzyme with an alanyl residue replacing Glu-98. *J. Biol. Chem.* **1995**, *270*, 15557-15562.

(10) Strohl, W. Compilation and analysis of DNA sequences associated with apparent streptomycete promoters. *Nucleic Acids Res.* **1992**, *20*, 961-974.

(11) Bourn, W. a. B., B. Computer assisted identification and classification of streptomycete promoters. *Nucleic Acids Res.* **1995**, *23*, 3696-3703.

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

(13) Tercero, J. A.; Espinosa, J. C.; Lacalle, R. A.; Jimenez, A. The biosynthetic pathway of the aminonucleoside antibiotic puromycin, as deduced from the molecular analysis of the pur cluster of Streptomyces alboniger. *J Biol Chem* **1996**, *271*, 1579-1590.

(14) Pissowotzki, K.; Mansouri, K.; Piepersberg, W. Genetics of streptomycin production in *Streptomyces griseus*: molecular structure and putative function of genes strELMB2N. *Mol. Gen. Genet.* **1991**, *231*, 113-123.

(15) Porter, D. J.; Merrill, B. M.; Short, S. A. Identification of the active site nucleophile in nucleoside 2-deoxyribosyltransferase as glutamic acid 98. *J. Biol. Chem.* **1995**, *270*, 15551-15556.

(16) Short, S. A.; Armstrong, S. R.; Ealick, S. E.; Porter, D. J. Active site amino acids that participate in the catalytic mechanism of nucleoside 2'-deoxyribosyltransferase. *J. Biol. Chem.* **1996**, *271*, 4978-4987.

(17) Marchler-Bauer A, A. J., DeWeese-Scott C, Fedorova ND, Geer LY, He S, Hurwitz DI, Jackson JD, Jacobs AR, Lanczycki CJ, Liebert CA, Liu C, Madej T, Marchler GH, Mazumder R, Nikolskaya AN, Panchenko AR, Rao BS, Shoemaker BA, Simonyan V, Song JS, Thiessen RA, Vasudevan S, Wang Y, Yamashita RA, Yin JJ, and Bryant SH CDD: a curated Entrez database of conserved domain alignments. *Nucleic Acids Res.* **2003**, *31*, 383-387.

(18) Betbeder, D.; Hutchinson, D. W.; Richards, A. O. The stereoselective enzymatic synthesis of 9-beta-D-2'-deoxyribofuranosyl 1-deazapurine. *Nucleic Acids Res.* **1989**, *17*, 4217-4222.

(19) Betbeder, D.; Heath, C. M.; Hutchinson, D. W. Enzmatic Synthesis of Dideoxyribonucleosides. *Nucleosides and Nucleotides* **1991**, *10*, 465-468.

(20) Okuyama, K.; Shibuya, S.; Hamamoto, T.; Noguchi, T. Enzymatic synthesis of 2'-deoxyguanosine with nucleoside deoxyribosyltransferase-II. *Biosci. Biotechnol. Biochem.* **2003**, *67*, 989-995.

(21) Hanrahan, J. R.; Hutchinson, D. W. The enzymatic synthesis of antiviral agents. *J. Biotech.* **1992**, *23*, 193-210.

(22) Armstrong, S. R.; Cook, W. J.; Short, S. A.; Ealick, S. E. Crystal structures of nucleoside 2-deoxyribosyltransferase in native and ligand-bound forms reveal architecture of the active site. *Structure* **1996**, *4*, 97-107.

(23) Anand, R.; Kaminski, P. A.; Ealick, S. E. Structures of purine 2'-deoxyribosyltransferase, substrate complexes, and the ribosylated enzyme intermediate at 2.0 A resolution. *Biochemistry* **2004**, *43*, 2384-2393.

(24) Yang, C.; Carlow, D.; Wolfenden, R.; Short, S. A. Cloning and nucleotide sequence of the *Escherichia coli* cytidine deaminase (ccd) gene. *Biochemistry* **1992**, *31*, 4168-4174.

(25) Freytag, S. O.; Stricker, H.; Pegg, J.; Paielli, D.; Pradhan, D. G.; Peabody, J.; DePeralta-Venturina, M.; Xia, X.; Brown, S.; Lu, M.; Kim, J. H. Phase I study of replication-competent adenovirus-mediated double-suicide gene therapy in combination with conventional-dose three-dimensional conformal radiation therapy for the treatment of newly diagnosed, intermediate- to high-risk prostate cancer. *Cancer Res.* **2003**, *63*, 7497-7506.

(26) Snider, M. R., L.; Wolfenden, R.; and Cleland, W ^{15}N Kinetic Isotope Effects on Uncatalyzed and Enzymatic Deamination of Cytidine. *Biochemistry* **2002**, *41*, 415-421.

(27) McCarter, J. D.; Withers, S. G. Mechanisms of enzymatic glycoside hydrolysis. *Curr. Opin. Struct. Biol.* **1994**, *4*, 885-892.

(28) Cook, W. J.; Short, S. A.; Ealick, S. E. Crystallization and preliminary X-ray investigation of recombinant *Lactobacillus leichmannii* nucleoside deoxyribosyltransferase. *J. Biol. Chem.* **1990**, *265*, 2682-2683.

Chapter 7
Conclusions

The cloning of biosynthetic gene clusters has traditionally been motivated by an exciting activity or structural feature of a natural product. The development of molecular genetic approaches to natural product chemistry, however, may allow for the discovery of new natural products that may not have otherwise been detected.

Collection issues often result in the search for natural products from just a single specimen or collection. When a natural product chemist analyzes a tissue sample for the presence of a bioactive metabolite, he/she is only looking at a small window in the lifetime of the organism and exciting new compounds may be missed simply because they were not being produced or stored at the time of collection. The genomic analysis of a number of bacteria has often revealed the presence of multiple secondary metabolic gene clusters.[40,41,43] In many cases, the metabolites produced by these gene clusters are not known. An advantage to utilizing a molecular approach to studying marine natural products is that the biosynthetic genes are present even if it the metabolite is not being produced at the time of specimen collection. The development of methods to analyze an organism's metabolic potential at the genetic level could allow for the identification of interesting natural products that may otherwise be missed. Screening of metagenomic libraries for specific types of enzymes, such as mixed NRPS/PKS clusters, could be used to identify new gene clusters. If a biosynthetic gene cluster is located that has interesting or unusual features, structural predictions based on sequence analysis can be used to direct the isolation of the corresponding compound. Alternately, heterologous expression of the gene cluster could be used to produce the new natural product. This approach was recently demonstrated in the identification of a new mixed peptide-polyketide natural product from *Stigmatella aurantiaca*.[43] In a scan for NRPS and PKS genes in the *S. aurantiaca* genome one NRPS, three PKS and four mixed NPRS/PKS gene clusters were

located. A structural prediction for the product of one of the mixed pathways was then used to identify a new secondary metabolite.

The successful culture of bacteria that contain the *Jaspis* PKS/NRPS gene cluster presents an opportunity to use this 'reverse natural product chemistry' approach to try and identify the associated metabolite. If the rules of colinearity are obeyed, analysis of the *Jaspis* mixed PKS/NRPS gene cluster and preliminary biochemical characterization would predict that this gene cluster is involved in the biosynthesis of a mixed tripeptide-polyketide product.

Even if the compound produced by the *Jaspis* gene cluster can not be found, the continuing development of combinatorial biosynthesis could still find an application for the gene cluster. Ultimately, the success of combinatorial biosynthesis will rely on the availability of unique biosynthetic gene clusters to provide building blocks for hybrid biosynthetic pathways. In order for maximum diversity to be created through combinatorial biosynthesis, enzymes that are capable of unique chemical transformations must be found. Herein lies the advantage of characterizing new biosynthetic gene clusters even if the associated natural product is not known. Heterologous expression and *in vitro* studies with the genes could provide information about the chemistry involved in the gene cluster while combinatorial methods could be used to introduce these features into engineered systems. Additionally, site directed mutagenesis can be used to alter substrate selectivity or even catalytic activity, as was observed for the blasticidin S CMP hydrolase.

The extraordinary diversity in both structure and biological activity of natural products being isolated from marine organisms is generating increased interest in this field. As the demand for new drugs with unique mechanisms of action continues to grow, the molecular approaches to natural product discovery and drug development are likely to become more prevalent. The work presented in these chapters demonstrates both the difficulties and the advantages of developing molecular biology approaches for these applications. Although the jaspamide biosynthetic gene cluster has not been identified, the results of these experiments are helping to guide the design of new approaches to cloning the jaspamide gene cluster. More significantly, this work has

demonstrated the ability to clone gene clusters from uncultured organisms and has helped to define several hurdles that must be overcome for the success of this approach. The methods developed for the cloning of gene clusters from symbiotic organisms can also be applied to study the metabolic potential of complex marine assemblages and used to expand our knowledge of marine natural products and antibiotic biosynthesis. Once these gene clusters are cloned, enzymatic characterization of the systems can provide insight into the chemistry and provide opportunities to generate natural product analogs through engineered biosynthesis.

Bibliography

Abell, C. 1998 Enzymes that exploit imines--one way or the other. *Biochem. Soc. Trans.* **26**, 310-5.

Aberhart, D. J. 1988 Separation by high-performance liquid chromatography of alpha- and beta-amino acids: application to assays of lysine 2,3-aminomutase and leucine 2,3-aminomutase. *Anal. Biochem.* **169**, 350-5.

Aberhart, D. J., Gould, S. J., Lin, H., Thiruvengadam, T. K. & Weiller, B. H. 1983 Stereochemistry of Lysine 2,3-aminomutase Isolated from *Clostridium subterminale* Strain SB4. *J. Am. Chem. Soc.* **105**, 5461-5470.

Admiraal, S. J., Khosla, C. & Walsh, C. T. 2002 The loading and initial elongation modules of rifamycin synthetase collaborate to produce mixed aryl ketide products. *Biochemistry* **41**, 5313-24.

Akee, R., Cartner, L., McCloud, T., Colin, P. & Newman, D. 1997 Distribution and Preparative Isolation of Jaspamide, a Marine Anticancer Cyclic Peptide. In *American Society of Pharmacognosy- 38th Annual Meeting-July 26-30.* Iowa City, IA.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. 1990 Basic local alignment search tool. *J. Mol. Biol.* **215**, 403-10.

Anand, R., Kaminski, P. A. & Ealick, S. E. 2004 Structures of purine 2'-deoxyribosyltransferase, substrate complexes, and the ribosylated enzyme intermediate at 2.0 A resolution. *Biochemistry* **43**, 2384-93.

Arakawa, K., Muller, R., Mahmud, T., Yu, T. W. & Floss, H. G. 2002 Characterization of the early stage aminoshikimate pathway in the formation of 3-amino-5-hydroxybenzoic acid: the RifN protein specifically converts kanosamine into kanosamine 6-phosphate. *J. Am. Chem. Soc.* **124**, 10644-5.

Argoudelis, A. D., Baczynskyj, L., Kuo, M. T., Laborde, A. L., Sebek, O. K., Truesdell, S. E. & Shilliday, F. B. 1987 Arginomycin: production, isolation, characterization and structure. *J. Antibiot. (Tokyo)* **40**, 750-60.

Armstrong, S. R., Cook, W. J., Short, S. A. & Ealick, S. E. 1996 Crystal structures of nucleoside 2-deoxyribosyltransferase in native and ligand-bound forms reveal architecture of the active site. *Structure* **4**, 97-107.

August, P. R., Tang, L., Yoon, Y. J., Ning, S., Muller, R., Yu, T. W., Taylor, M., Hoffmann, D., Kim, C. G., Zhang, X., Hutchinson, C. R. & Floss, H. G. 1998 Biosynthesis of the ansamycin antibiotic rifamycin: deductions from the molecular analysis of the rif biosynthetic gene cluster of *Amycolatopsis mediterranei* S699. *Chem. Biol.* **5**, 69-79.

Ballinger, M. D., Frey, P. A. & Reed, G. H. 1992 Structure of a substrate radical intermediate in the reaction of lysine 2,3-aminomutase. *Biochemistry* **31**, 10782-9.

Ballinger, M. D., Frey, P. A., Reed, G. H. & LoBrutto, R. 1995 Pulsed electron paramagnetic resonance studies of the lysine 2,3-aminomutase substrate radical: evidence for participation of pyridoxal 5'-phosphate in a radical rearrangement. *Biochemistry* **34**, 10086-93.

Baltz, R. 1980 Genetic Recombination in *S. fradiae* by protoplast fusion and cell regeneration. *Dev. Ind. Microbiol.* **21**, 43-54.

Bamas-Jacques, N., Lorenzon, S., Lacroix, P., De Swetschin, C. & Crouzet, J. 1999 Cluster organization of the genes of *Streptomyces pristinaespiralis* involved in pristinamycin biosynthesis and resistance elucidated by pulsed-field gel electrophoresis. *J. Appl. Microbiol.* **87**, 939-948.

Becker, J. E., Moore, R. E. & Moore, B. S. 2004 Cloning, sequencing, and biochemical characterization of the nostocyclopeptide biosynthetic gene cluster: molecular basis for imine macrocyclization. *Gene* **325**, 35-42.

Belshaw, P. J., Walsh, C. T. & Stachelhaus, T. 1999 Aminoacyl-CoAs as probes of condensation domain selectivity in nonribosomal peptide synthesis. *Science* **284**, 486-9.

Bentley, S. D., Chater, K. F., Cerdeno-Tarraga, A. M., Challis, G. L., Thomson, N. R., James, K. D., Harris, D. E., Quail, M. A., Kieser, H., Harper, D., Bateman, A., Brown, S., Chandra, G., Chen, C. W., Collins, M., Cronin, A., Fraser, A., Goble, A., Hidalgo, J., Hornsby, T., Howarth, S., Huang, C. H., Kieser, T., Larke, L., Murphy, L., Oliver, K., O'Neil, S., Rabbinowitsch, E., Rajandream, M. A., Rutherford, K., Rutter, S., Seeger, K., Saunders, D., Sharp, S., Squares, R., Squares, S., Taylor, K., Warren, T., Wietzorrek, A., Woodward, J., Barrell, B. G., Parkhill, J. & Hopwood, D. A. 2002 Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* **417**, 141-7.

Beppu, T. 1995 Signal transduction and secondary metabolism: prospects for controlling productivity. *Trends. Biotechnol.* **13**, 264-9.

Berger, S. L. & Kimmel, A. R. 1987 Guide to molecular cloning techniques. *Methods enzymol.* **152**, 410-413.

Betbeder, D., Heath, C. M. & Hutchinson, D. W. 1991 Enzmatic Synthesis of Dideoxyribonucleosides. *Nucleosides and Nucleotides* **10**, 465-468.

Betbeder, D., Hutchinson, D. W. & Richards, A. O. 1989 The stereoselective enzymatic synthesis of 9-beta-D-2'-deoxyribofuranosyl 1-deazapurine. *Nucleic Acids Res.* **17**, 4217-22.

Betlach, M. C., Kealey, J. T., Ashley, G. W. & McDaniel, R. 1998 Characterization of the macrolide P-450 hydroxylase from *Streptomyces venezuelae* which converts narbomycin to picromycin. *Biochemistry* **37**, 14937-42.

Betts, L., Xiang, S., Short, S. A., Wolfenden, R. & Carter, C. W., Jr. 1994 Cytidine deaminase. The 2.3 A crystal structure of an enzyme: transition-state analog complex. *J. Mol. Biol.* **235**, 635-56.

Blunt, J. W., Copp, B. R., Munro, M. H., Northcote, P. T. & Prinsep, M. R. 2004 Marine natural products. *Nat. Prod. Rep.* **21**, 1-49.

Borchert, S., Patil, S. S. & Marahiel, M. A. 1992 Identification of putative multifunctional peptide synthetase genes using highly conserved oligonucleotide sequences derived from known synthetases. *FEMS Microbiol. Lett.* **71**, 175-80.

Bormann, C., Mohrle, V. & Bruntner, C. 1996 Cloning and heterologous expression of the entire set of structural genes for nikkomycin synthesis from *Streptomyces tendae* Tu901 in *Streptomyces lividans*. *J. Bacteriol.* **178**, 1216-8.

Bourn, W. a. B., B. 1995 Computer assisted identification and classification of streptomycete promoters. *Nucleic Acids Res.* **23**, 3696-3703.

Bradford, M. M. 1976 A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248-54.

Braekman, J. C., Daloze, D., Moussiaux, B. & Riccio, R. 1987 Jaspamide from the Marine Sponge *Jaspis johnstoni*. *J. Natural Products* **50**, 994-995.

Brinkmann, U., Mattes, R. E. & Buckel, P. 1989 High-level expression of recombinant genes in *Escherichia coli* is dependent on the availability of the dnaY gene product. *Gene* **85**, 109-14.

Brooks, D. R., McCulloch, R., Coombs, G. H. & Mottram, J. C. 2000 Stable transformation of trypanosomatids through targeted chromosomal integration of the selectable marker gene encoding blasticidin S deaminase. *FEMS Microbiol. Lett.* **186**, 287-91.

Bubb, M. R., Senderowicz, A. M., Sausville, E. A., Duncan, K. L. & Korn, E. D. 1994 Jasplakinolide, a cytotoxic natural product, induces actin polymerization and competitively inhibits the binding of phalloidin to F-actin. *J. Biol. Chem.* **269**, 14869-71.

Bubb, M. R., Spector, I., Beyer, B. B. & Fosen, K. M. 2000 Effects of jasplakinolide on the kinetics of actin polymerization. An explanation for certain *in vivo* observations. *J. Biol. Chem.* **275**, 5163-70.

Cane, D. E. & Walsh, C. T. 1999 The parallel and convergent universes of polyketide synthases and nonribosomal peptide synthetases. *Chem. Biol.* **6**, R319-25.

Cane, D. E., Walsh, C. T. & Khosla, C. 1998 Harnessing the biosynthetic code: combinations, permutations, and mutations. *Science* **282**, 63-8.

Cassell, G. H. & Mekalanos, J. 2001 Development of antimicrobial agents in the era of new and reemerging infectious diseases and increasing antibiotic resistance. *J.A.M.A.* **285**, 601-5.

Challis, G. L., Ravel, J. & Townsend, C. A. 2000 Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains. *Chem. Biol.* **7**, 211-24.

Chan, W., Tinto, W. F., Manchand, P. S. & Todaro, L. J. 1987 Stereostructrues of Geodiamolides A and B, Novel Cyclodepsipeptides from the Marine Sponge *Geodia sp. J. Org. Chem.* **52**, 3091-3097.

Chang, Z., Flatt, P., Gerwick, W. H., Nguyen, V. A., Willis, C. L. & Sherman, D. H. 2002 The barbamide biosynthetic gene cluster: a novel marine cyanobacterial system of mixed polyketide synthase (PKS)-non-ribosomal peptide synthetase (NRPS) origin involving an unusual trichloroleucyl starter unit. *Gene* **296**, 235-47.

Chater, K. F. 1992 Genetic regulation of secondary metabolic pathways in Streptomyces. *Ciba. Found. Symp.* **171**, 144-56; discussion 156-62.

Chen, D. & Frey, P. A. 2001 Identification of lysine 346 as a functionally important residue for pyridoxal 5'-phosphate binding and catalysis in lysine 2, 3-aminomutase from *Bacillus subtilis*. *Biochemistry* **40**, 596-602.

Chen, D., Ruzicka, F. J. & Frey, P. A. 2000 A novel lysine 2,3-aminomutase encoded by the yodO gene of bacillus subtilis: characterization and the observation of organic radical intermediates. *Biochem. J.* **348 Pt 3**, 539-49.

Chen, H., Hubbard, B. K., O'Connor, S. E. & Walsh, C. T. 2002 Formation of beta-hydroxy histidine in the biosynthesis of nikkomycin antibiotics. *Chem. Biol.* **9**, 103-12.

Cheng, Y. Q., Tang, G. L. & Shen, B. 2002 Identification and localization of the gene cluster encoding biosynthesis of the antitumor macrolactam leinamycin in *Streptomyces atroolivaceus* S-140. *J. Bacteriol.* **184**, 7013-24.

Chevallier, C., Richardson, A. D., Edler, M. C., Hamel, E., Harper, M. K. & Ireland, C. M. 2003 A new cytotoxic and tubulin-interactive milnamide derivative from a marine sponge *Cymbastela sp. Org. Lett.* **5**, 3737-9.

Chirpich, T. P., Zappia, V., Costilow, R. N. & Barker, H. A. 1970 Lysine 2,3-aminomutase. Purification and properties of a pyridoxal phosphate and *S*-adenosylmethionine-activated enzyme. *J. Biol. Chem.* **245**, 1778-89.

Christiansen, G., Fastner, J., Erhard, M., Borner, T. & Dittmann, E. 2003 Microcystin biosynthesis in planktothrix: genes, evolution, and manipulation. *J. Bacteriol.* **185**, 564-72.

Clugston, S. L., Sieber, S. A., Marahiel, M. A. & Walsh, C. T. 2003 Chirality of peptide bond-forming condensation domains in nonribosomal peptide synthetases: the C5 domain of tyrocidine synthetase is a (D)C(L) catalyst. *Biochemistry* **42**, 12095-104.

Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. V., Eiglmeier, K., Gas, S., Barry, C. E., 3rd, Tekaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Barrell, B. G. & et al. 1998 Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. *Nature* **393**, 537-44.

Coleman, J. E., Van Soest, R. & Andersen, R. J. 1999 New Geodiamolides from the Sponge *Cymbastela sp.* Collected in Papau New Guinea. *J. Nat. Prod.* **62**, 1137-1141.

Cone, M. C., Petrich, A. K., Gould, S. J. & Zabriskie, T. M. 1998 Cloning and heterologous expression of blasticidin S biosynthetic genes from *Streptomyces griseochromogenes. J. Antibiot. (Tokyo)* **51**, 570-8.

Cone, M. C., Yin, X., Grochowski, L. L., Parker, M. R. & Zabriskie, T. M. 2003 The blasticidin S biosynthesis gene cluster from *Streptomyces griseochromogenes*: sequence analysis, organization, and initial characterization. *Chembiochem* **4**, 821-8.

Conti, E., Stachelhaus, T., Marahiel, M. A. & Brick, P. 1997 Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S. *Embo. J.* **16**, 4174-83.

Cook, W. J., Short, S. A. & Ealick, S. E. 1990 Crystallization and preliminary X-ray investigation of recombinant *Lactobacillus leichmannii* nucleoside deoxyribosyltransferase. *J. Biol. Chem.* **265**, 2682-3.

Cooper, R., Conover, M. & Patel, M. 1988 Sch 36605, structure of a novel nucleoside. *J. Antibiot. (Tokyo)* **41**, 123-5.

Cortes, J., Wiesmann, K. E., Roberts, G. A., Brown, M. J., Staunton, J. & Leadlay, P. F. 1995 Repositioning of a domain in a modular polyketide synthase to promote specific chain cleavage. *Science* **268**, 1487-9.

Costilow, R. N., Rochovansky, O. M. & Barker, H. A. 1966 Isolation and identification of beta-lysine as an intermediate in lysine fermentation. *J. Biol. Chem.* **241**, 1573-80.

Crews, P., Farias, J., Emrich, R. & Keifer, P. 1994 Milnamide A, an Unusual Cytotoxic Tripeptide from the Marine Sponge *Auletta cf. constricta. J. Org. Chem.* **59**, 2932-2934.

Crews, P., Manes, L. & Boehler, M. 1986 Jasplakinolide, A Cyclopepsipeptide from the Marine Sponge, *Jaspis* sp. *Tetrahedron Lett.* **27**, 2797-2800.

Cropp, T. A., Kim, B. S., Beck, B. J., Yoon, Y. J., Sherman, D. H. & Reynolds, K. A. 2002 Recent developments in the production of novel polyketides by combinatorial biosynthesis. *Biotechnol. Genet. Eng. Rev.* **19**, 159-72.

Cundliffe, E. 1978 Mechanism of resistance to thiostrepton in the producing-organism *Streptomyces azureus. Nature* **272**, 792-5.

Cundliffe, E. 1984 Self defence in antibiotic-producing organisms. *Br. Med. Bull.* **40**, 61-7.

Cundliffe, E. 1989 How antibiotic-producing organisms avoid suicide. *Annu. Rev. Microbiol.* **43**, 207-33.

Cundliffe, E. 1992 Self-protection mechanisms in antibiotic producers. *Ciba. Found. Symp.* **171**, 199-208; discussion 208-14.

Cundliffe, E. & Thompson, J. 1981 The mode of action of nosiheptide (multhiomycin) and the mechanism of resistance in the producing organism. *J.Gen. Microbiol.* **126**, 185-92.

D'Auria, M. V., Gomez Paloma, L., Minale, L., Zampella, A., Debitus, C. & Perez, J. 1995 Neosiphoniamolide A, a novel cyclodepsipeptide, with antifungal activity from the marine sponge *Neosiphonia superstes. J. Nat. Prod.* **58**, 121-3.

de Crecy-Lagard, V., Blanc, V., Gil, P., Naudin, L., Lorenzon, S., Famechon, A., Bamas-Jacques, N., Crouzet, J. & Thibaut, D. 1997 Pristinamycin I biosynthesis in *Streptomyces pristinaespiralis*: molecular characterization of the first two structural peptide synthetase genes. *J. Bacteriol.* **179**, 705-13.

de Ferra, F., Rodriguez, F., Tortora, O., Tosi, C. & Grandi, G. 1997 Engineering of peptide synthetases. Key role of the thioesterase-like domain for efficient production of recombinant peptides. *J. Biol. Chem.* **272**, 25304-9.

Del Vecchio, F., Petkovic, H., Kendrew, S. G., Low, L., Wilkinson, B., Lill, R., Cortes, J., Rudd, B. A., Staunton, J. & Leadlay, P. F. 2003 Active-site residue, domain and module swaps in modular polyketide synthases. *J. Ind. Microbiol. Biotechnol.* **30**, 489-94.

Demain, A. L. 2000 Small bugs, big business: the economic power of the microbe. *Biotechnol. Adv.* **18**, 499-514.

Denoya, C. D., Fedechko, R. W., Hafner, E. W., McArthur, H. A., Morgenstern, M. R., Skinner, D. D., Stutzman-Engwall, K., Wax, R. G. & Wernau, W. C. 1995 A second branched-chain alpha-keto acid dehydrogenase gene cluster (bkdFGH) from *Streptomyces avermitilis*: its relationship to avermectin biosynthesis and the

construction of a bkdF mutant suitable for the production of novel antiparasitic avermectins. *J. Bacteriol.* **177**, 3504-11.

Dieckmann, R., Lee, Y. O., van Liempt, H., von Dohren, H. & Kleinkauf, H. 1995 Expression of an active adenylate-forming domain of peptide synthetases corresponding to acyl-CoA-synthetases. *FEBS Lett.* **357**, 212-6.

Dilip de Silva, E., Andersen, R. J. & Allen, T. 1990 Geodiamolides C to F, New Cytotoxic Cyclodepsipeptides from the Marine Sponge *Pseudaxinyssa sp. Tetrahedron Lett.* **31**, 489-492.

Doekel, S. & Marahiel, M. A. 2000 Dipeptide formation on engineered hybrid peptide synthetases. *Chem. Biol.* **7**, 373-84.

Donadio, S., McAlpine, J. B., Sheldon, P. J., Jackson, M. & Katz, L. 1993 An erythromycin analog produced by reprogramming of polyketide synthesis. *Proc. Natl. Acad. Sci. U S A* **90**, 7119-23.

Donadio, S., Staver, M. J., McAlpine, J. B., Swanson, S. J. & Katz, L. 1991 Modular organization of genes required for complex polyketide biosynthesis. *Science* **252**, 675-9.

Donadio, S., Staver, M. J., McAlpine, J. B., Swanson, S. J. & Katz, L. 1992 Biosynthesis of the erythromycin macrolactone and a rational approach for producing hybrid macrolides. *Gene* **115**, 97-103.

Doumith, M., Legrand, R., Lang, C., Salas, J. A. & Raynal, M. C. 1999 Interspecies complementation in *Saccharopolyspora erythraea* : elucidation of the function of oleP1, oleG1 and oleG2 from the oleandomycin biosynthetic gene cluster of *Streptomyces antibioticus* and generation of new erythromycin derivatives. *Mol. Microbiol.* **34**, 1039-48.

Du, L., Sanchez, C., Chen, M., Edwards, D. J. & Shen, B. 2000 The biosynthetic gene cluster for the antitumor drug bleomycin from *Streptomyces verticillus* ATCC15003 supporting functional interactions between nonribosomal peptide synthetases and a polyketide synthase. *Chem. Biol.* **7**, 623-42.

Duerfahrt, T., Doekel, S., Sonke, T., Quaedflieg, P. J. & Marahiel, M. A. 2003 Construction of hybrid peptide synthetases for the production of alpha-l-aspartyl-l-phenylalanine, a precursor for the high-intensity sweetener aspartame. *Eur. J. Biochem.* **270**, 4555-63.

Duitman, E. H., Hamoen, L. W., Rembold, M., Venema, G., Seitz, H., Saenger, W., Bernhard, F., Reinhardt, R., Schmidt, M., Ullrich, C., Stein, T., Leenders, F. & Vater, J. 1999 The mycosubtilin synthetase of *Bacillus subtilis* ATCC6633: a multifunctional hybrid between a peptide synthetase, an amino transferase, and a fatty acid synthase. *Proc. Natl. Acad. Sci. U S A* **96**, 13294-9.

Dutton, C. J., Gibson, S. P., Goudie, A. C., Holdom, K. S., Pacey, M. S., Ruddock, J. C., Bu'Lock, J. D. & Richards, M. K. 1991 Novel avermectins produced by mutational biosynthesis. *J. Antibiot. (Tokyo)* **44**, 357-65.

Edell, D., Khoshoo, V., Ross, G. & Salter, K. 2002 Early ribavarin treatment of bronchiolitis: effect on long-term respiratory morbidity. *Chest* **122**, 935-9.

Endo, T., Furuta, K., Kaneko, A., Katsuki, T., Kobayashi, K., Azuma, A., Watanabe, A. & Shimazu, A. 1987 Inactivation of blasticidin S by *Bacillus cereus*. I. Inactivation mechanism. *J. Antibiot. (Tokyo)* **40**, 1791-3.

Endo, T., Kobayashi, K., Nakayama, N., Tanaka, T., Kamakura, T. & Yamaguchi, I. 1988 Inactivation of blasticidin S by *Bacillus cereus*. II. Isolation and characterization of a plasmid, pBSR8, from *Bacillus cereus*. *J. Antibiot. (Tokyo)* **41**, 271-3.

Endo, T., Otake, N., Takeuchi, S. & Yonehara, H. 1964 The Structure of Blastidone, a Degradative Component of Blasticidin S. *J. Antibiot. (Tokyo)* **17**, 172-3.

Engel, S., Jensen, P. R. & Fenical, W. 2002 Chemical ecology of marine microbial defense. *J. Chem. Ecol.* **28**, 1971-85.

Eppelmann, K., Stachelhaus, T. & Marahiel, M. A. 2002 Exploitation of the selectivity-conferring code of nonribosomal peptide synthetases for the rational design of novel peptide antibiotics. *Biochemistry* **41**, 9718-26.

Espinosa, J. C., Tercero, J. A., Rubio, M. A. & Jimenez, A. 1999 The pur7 gene from the puromycin biosynthetic pur cluster of *Streptomyces alboniger* encodes a nudix hydrolase. *J. Bacteriol.* **181**, 4914-8.

Feduchi, E., Cosin, M. & Carrasco, L. 1985 Mildiomycin: a nucleoside antibiotic that inhibits protein synthesis. *J. Antibiot. (Tokyo)* **38**, 415-9.

Fernandez-Moreno, M. A., Vallin, C. & Malpartida, F. 1997 Streptothricin biosynthesis is catalyzed by enzymes related to nonribosomal peptide bond formation. *J. Bacteriol.* **179**, 6929-36.

Flett, F., Mersinias, V. & Smith, C. P. 1997 High efficiency intergeneric conjugal transfer of plasmid DNA from *Escherichia coli* to methyl DNA-restricting streptomycetes. *FEMS Microbiol. Lett.* **155**, 223-9.

Frey, P. A. 1993 Lysine 2,3-aminomutase: is adenosylmethionine a poor man's adenosylcobalamin? *Faseb. J.* **7**, 662-70.

Frey, P. A. 1997 Radicals in enzymatic reactions. *Curr. Opin. Chem. Biol.* **1**, 347-56.

Frey, P. A. 2001a Radical mechanisms of enzymatic catalysis. *Annu. Rev. Biochem.* **70**, 121-48.

Frey, P. A. 2001b The role of radicals in enzymatic processes. *Chem. Rec.* **1**, 277-89.

Frey, P. A., Ballinger, M. D. & Reed, G. H. 1998 S-adenosylmethionine: a 'poor man's coenzyme B12' in the reaction of lysine 2,3-aminomutase. *Biochem. Soc. Trans.* **26**, 304-10.

Frey, P. A. & Booker, S. J. 2001 Radical mechanisms of *S*-adenosylmethionine-dependent enzymes. *Adv. Protein. Chem.* **58**, 1-45.

Frey, P. A., Chang, C. H., Ballinger, M. D. & Reed, G. H. 2002 Kinetic characterization of transient free radical intermediates in reaction of lysine 2,3-aminomutase by EPR lineshape analysis. *Methods Enzymol.* **354**, 426-35.

Frey, P. A., Moss, M., Petrovich, R. & Baraniak, J. 1990 The roles of *S*-adenosylmethionine and pyridoxal phosphate in the lysine 2,3-aminomutase reaction. *Ann. N. Y. Acad. Sci.* **585**, 368-78.

Frey, P. A. & Moss, M. L. 1987 *S*-adenosylmethionine and the mechanism of hydrogen transfer in the lysine 2,3-aminomutase reaction. *Cold Spring Harb. Symp. Quant. Biol.* **52**, 571-7.

Frey, P. A. & Reed, G. H. 1993 Lysine 2,3-aminomutase and the mechanism of the interconversion of lysine and beta-lysine. *Adv. Enzymol. Relat. Areas Mol. Biol.* **66**, 1-39.

Frey, P. A. & Reed, G. H. 2000 Radical mechanisms in adenosylmethionine- and adenosylcobalamin-dependent enzymatic reactions. *Arch. Biochem. Biophys.* **382**, 6-14.

Freytag, S. O., Khil, M., Stricker, H., Peabody, J., Menon, M., DePeralta-Venturina, M., Nafziger, D., Pegg, J., Paielli, D., Brown, S., Barton, K., Lu, M., Aguilar-Cordova, E. & Kim, J. H. 2002 Phase I study of replication-competent adenovirus-mediated double suicide gene therapy for the treatment of locally recurrent prostate cancer. *Cancer Res.* **62**, 4968-76.

Freytag, S. O., Stricker, H., Pegg, J., Paielli, D., Pradhan, D. G., Peabody, J., DePeralta-Venturina, M., Xia, X., Brown, S., Lu, M. & Kim, J. H. 2003 Phase I study of replication-competent adenovirus-mediated double-suicide gene therapy in combination with conventional-dose three-dimensional conformal radiation therapy for the treatment of newly diagnosed, intermediate- to high-risk prostate cancer. *Cancer Res.* **63**, 7497-506.

Gaisser, S., Lill, R., Staunton, J., Mendez, C., Salas, J. & Leadlay, P. F. 2002 Parallel pathways for oxidation of 14-membered polyketide macrolactones in *Saccharopolyspora erythraea*. *Mol. Microbiol.* **44**, 771-81.

Gaisser, S., Reather, J., Wirtz, G., Kellenberger, L., Staunton, J. & Leadlay, P. F. 2000 A defined system for hybrid macrolide biosynthesis in *Saccharopolyspora erythraea*. *Mol. Microbiol.* **36**, 391-401.

Gaitatzis, N., Hans, A., Muller, R. & Beyer, S. 2001 The mtaA gene of the myxothiazol biosynthetic gene cluster from *Stigmatella aurantiaca* DW4/3-1 encodes a phosphopantetheinyl transferase that activates polyketide synthases and polypeptide synthetases. *J. Biochem. (Tokyo)* **129**, 119-24.

Gillespie, D. E., Brady, S. F., Bettermann, A. D., Cianciotto, N. P., Liles, M. R., Rondon, M. R., Clardy, J., Goodman, R. M. & Handelsman, J. 2002 Isolation of antibiotics turbomycin a and B from a metagenomic library of soil microbial DNA. *Appl. Environ. Microbiol.* **68**, 4301-6.

Gokhale, R. S., Lau, J., Cane, D. E. & Khosla, C. 1998 Functional orientation of the acyltransferase domain in a module of the erythromycin polyketide synthase. *Biochemistry* **37**, 2524-8.

Gold, H. S. & Moellering, R. C., Jr. 1996 Antimicrobial-drug resistance. *N. Engl. J. Med.* **335**, 1445-53.

Gould, S. J. 1997 Blasticidin S and Related Peptidyl Nucleoside Antibiotics. In *Biotechnology of Antibiotics*, vol. 82 (ed. W. Strohl), pp. 703-731. New York: Marcel Dekker, Inc.

Gould, S. J. & Guo, J. 1991 Biosynthesis of blasticidin S. Cell-free demonstration of N-methylation as the last step. *Bioorg. Med. Chem. Lett.* **1**, 497-500.

Gould, S. J. & Guo, J. 1994 Cytosylglucuronic acid synthase (cytosine: UDP-glucuronosyltransferase) from *Streptomyces griseochromogenes*, the first prokaryotic UDP-glucuronosyltransferase. *J. Bacteriol.* **176**, 1282-6.

Gould, S. J., Guo, J., Geitmann, A. & Dejesus, K. 1994 Nucleoside intermediates in blasticidin S biosynthesis identified by the in vivo use of enzyme inhibitors. *Canadian J. Chem.* **72**, 6-11.

Gould, S. J. & Zhang, Q. 1995 Cytosinine: pyridoxal phosphate tautomerase, a new enzyme in the blasticidin S biosynthetic pathway. *J. Antibiot. (Tokyo)* **48**, 652-6.

Gourley, D. G., Shrive, A. K., Polikarpov, I., Krell, T., Coggins, J. R., Hawkins, A. R., Isaacs, N. W. & Sawyer, L. 1999 The two types of 3-dehydroquinase have distinct structures but catalyze the same overall reaction. *Nat. Struct. Biol.* **6**, 521-5.

Gullo, V., Conover, M., Cooper, R., Federbush, C., Horan, A. C., Kung, T., Marquez, J., Patel, M. & Watnick, A. 1988 Sch 36605, a novel anti-inflammatory compound. Taxonomy, fermentation, isolation and biological properties. *J. Antibiot. (Tokyo)* **41**, 20-4.

Guo, J. 1992 Biosynthesis of Blasticidin S: Pathway and Enzymes for the Nucleoside Formation and Blastidic Acid Assembly. In *Biochem. Biophys.*, pp. 202. Corvallis: Oregon State University.

Guo, J. & Gould, S. J. 1991a Biosynthesis of Blasticidin S from Cytosylglucuronic Acid (CGA). Isolation of Cytosine/UDPglucuronoxyltransferase and Incorporation of CGA by *Streptomyces griseochromogenes*. *J. Am. Chem. Soc.* **113**, 5898-5899.

Guo, J. & Gould, S. J. 1991b Biosynthesis of Blasticidin S. Cell-Free Demonstration of *N*-methylation as the Last Step. *Bioorg. Med. Chem. Lett.* **1**, 497-500.

Haese, A., Schubert, M., Herrmann, M. & Zocher, R. 1993 Molecular characterization of the enniatin synthetase gene encoding a multifunctional enzyme catalysing *N*-methyldepsipeptide formation in *Fusarium scirpi*. *Mol. Microbiol.* **7**, 905-14.

Hammes, W., Schleifer, K. H. & Kandler, O. 1973 Mode of action of glycine on the biosynthesis of peptidoglycan. *J. Bacteriol.* **116**, 1029-53.

Handelsman, J., Rondon, M. R., Brady, S. F., Clardy, J. & Goodman, R. M. 1998 Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chem. Biol.* **5**, R245-9.

Hanrahan, J. R. & Hutchinson, D. W. 1992 The enzymatic synthesis of antiviral agents. *J. Biotech.* **23**, 193-210.

Hansen, J. L., Moore, P. B. & Steitz, T. A. 2003 Structures of five antibiotics bound at the peptidyl transferase center of the large ribosomal subunit. *J. Mol. Biol.* **330**, 1061-75.

Haydock, S. F., Aparicio, J. F., Molnar, I., Schwecke, T., Khaw, L. E., Konig, A., Marsden, A. F., Galloway, I. S., Staunton, J. & Leadlay, P. F. 1995 Divergent sequence motifs correlated with the substrate specificity of (methyl)malonyl-CoA:acyl carrier protein transacylase domains in modular polyketide synthases. *FEBS Lett.* **374**, 246-8.

Haygood, M. G., Schmidt, E. W., Davidson, S. K. & Faulkner, D. J. 1999 Microbial symbionts of marine invertebrates: opportunities for microbial biotechnology. *J. Mol. Microbiol. Biotechnol.* **1**, 33-43.

Hildebrand, M., Waggoner, L. E., Lim, G. E., Sharp, K. H., Ridley, C. P. & Haygood, M. G. 2004 Approaches to identify, clone, and express symbiont bioactive metabolite genes. *Nat. Prod. Rep.* **21**, 122-42.

Hirai, A., Wildman, S. G. & Hirai, T. 1968 Specific Inhibition of TMV-RNA Synthesis by Blasticidin S. *Virology* **36**, 646-651.

Hirai, T. & Shimomura, T. 1965 Blasticidin S, An Effective Antibiotic Against Plant Virus Multiplication. *Phytopathology* **55**, 291-295.

Hoffmann, D., Hevel, J. M., Moore, R. E. & Moore, B. S. 2003 Sequence analysis and biochemical characterization of the nostopeptolide A biosynthetic gene cluster from Nostoc sp. GSV224. *Gene* **311**, 171-80.

Hoffmeister, D., Ichinose, K., Domann, S., Faust, B., Trefzer, A., Drager, G., Kirschning, A., Fischer, C., Kunzel, E., Bearden, D., Rohr, J. & Bechthold, A. 2000 The NDP-sugar co-substrate concentration and the enzyme expression level influence the substrate specificity of glycosyltransferases: cloning and characterization of deoxysugar biosynthetic genes of the urdamycin biosynthetic gene cluster. *Chem. Biol.* **7**, 821-31.

Holzbaur, I. E., Ranganathan, A., Thomas, I. P., Kearney, D. J., Reather, J. A., Rudd, B. A., Staunton, J. & Leadlay, P. F. 2001 Molecular basis of Celmer's rules: role of the ketosynthase domain in epimerisation and demonstration that ketoreductase domains can have altered product specificity with unnatural substrates. *Chem. Biol.* **8**, 329-40.

Holzinger, A. 2001 Jasplakinolide. An actin-specific reagent that promotes actin polymerization. *Methods Mol. Biol.* **161**, 109-20.

Holzinger, A. & Meindl, U. 1997 Jasplakinolide, a novel actin targeting peptide, inhibits cell growth and induces actin filament polymerization in the green alga Micrasterias. *Cell. Motil. Cytoskeleton* **38**, 365-72.

Hopwood, D. A. 1981 Genetic studies with bacterial protoplasts. *Annu. Rev. Microbiol.* **35**, 237-72.

Hopwood, D. A., Malpartida, F., Kieser, H. M., Ikeda, H., Duncan, J., Fujii, I., Rudd, B. A., Floss, H. G. & Omura, S. 1985 Production of 'hybrid' antibiotics by genetic engineering. *Nature* **314**, 642-4.

Hopwood, D. A., Wright, H. M., Bibb, M. J. & Cohen, S. N. 1977 Genetic recombination through protoplast fusion in Streptomyces. *Nature* **268**, 171-4.

Horinouchi, S., Furuya, K., Nishiyama, M., Suzuki, H. & Beppu, T. 1987 Nucleotide sequence of the streptothricin acetyltransferase gene from *Streptomyces lavendulae* and its expression in heterologous hosts. *J. Bacteriol.* **169**, 1929-37.

Huang, K. T., Misato, T. & Asuyama, H. 1964a Effect of Blasticidin S on Protein Synthesis of *Piricularia Oryzae*. *J. Antibiot. (Tokyo)* **17**, 65-70.

Huang, K. T., Misato, T. & Asuyama, H. 1964b Selective Toxicity of Blasticidin S to *Piricularia Oryzae* and *Pellicularia Sasakii*. *J. Antibiot. (Tokyo)* **17**, 71-4.

Ichikawa, Y., Ohbayashi, M., Nishizawa, R. & Isobe, M. 2001 Synthesis of Blastidic Acid and Cytosinine, Two Components of Blasticidin S. *Synlett* **11**, 1763-1766.

Imada, A. K., M.; and Igarasi, S. 1967 Degradation of Pyrimidine Nucleotides by Enzyme Systems of *Streptomyces* II. Pyrimidine 5'-Nucleotide Phosphoribo (Deoxyribo) Hydrolase of *Streptomyces virginiae*. *J. Gen. Appl. Microbiol.* **13**, 267-278.

Ireton, G. C., Black, M. E. & Stoddard, B. L. 2003 The 1.14 A crystal structure of yeast cytosine deaminase: evolution of nucleotide salvage enzymes and implications for genetic chemotherapy. *Structure* **11**, 961-72.

Ishiguro, J. & Miyazaki, M. 1985 Characterization of blasticidin S-resistant mutants of *Saccharomyces cerevisiae*. *Curr. Genet.* **9**, 179-81.

Ishikawa, J. & Hotta, K. 1999 FramePlot: a new implementation of the frame analysis for predicting protein-coding regions in bacterial DNA with a high G + C content. *FEMS Microbiol. Lett.* **174**, 251-3.

Isono, K. 1988 Nucleoside antibiotics: structure, biological activity, and biosynthesis. *J. Antibiot. (Tokyo)* **41**, 1711-39.

Isono, K. 1991 Current progress on nucleoside antibiotics. *Pharmacol. Ther.* **52**, 269-86.

Izumi, M., Miyazawa, H., Kamakura, T., Yamaguchi, I., Endo, T. & Hanaoka, F. 1991 Blasticidin S-resistance gene (bsr): a novel selectable marker for mammalian cells. *Exp. Cell. Res.* **197**, 229-33.

Jackson, M. D., Gould, S. J. & Zabriskie, T. M. 2002 Studies on the formation and incorporation of streptolidine in the biosynthesis of the peptidyl nucleoside antibiotic streptothricin F. *J. Org. Chem.* **67**, 2934-41.

Jacobsen, J. R., Hutchinson, C. R., Cane, D. E. & Khosla, C. 1997 Precursor-directed biosynthesis of erythromycin analogs by an engineered polyketide synthase. *Science* **277**, 367-9.

Jacobsen, J. R., Keatinge-Clay, A. T., Cane, D. E. & Khosla, C. 1998 Precursor-directed biosynthesis of 12-ethyl erythromycin. *Bioorg. Med. Chem. Lett.* **6**, 1171-7.

Jansen, R., Kunze, B., Reichenbach, H. & Hofle, G. 1996 Chondramides A-D, New Cytostatic and Antifungal Cyclodepsipeptides from *Chondromyces crocatus* (Myxobacteria): Isolation and Structure Elucidation. *Liebigs. Ann.*, 285-290.

Jez, J. M., Ferrer, J. L., Bowman, M. E., Dixon, R. A. & Noel, J. P. 2000 Dissection of malonyl-coenzyme A decarboxylation from polyketide formation in the reaction mechanism of a plant polyketide synthase. *Biochemistry* **39**, 890-902.

Jez, J. M. & Noel, J. P. 2000 Mechanism of chalcone synthase. pKa of the catalytic cysteine and the role of the conserved histidine in a plant polyketide synthase. *J. Biol. Chem.* **275**, 39640-6.

Kaeberlein, T., Lewis, K. & Epstein, S. S. 2002 Isolating "uncultivable" microorganisms in pure culture in a simulated natural environment. *Science* **296**, 1127-9.

Kamakura, T., Kobayashi, K., Tanaka, T., Yamaguchi, I. & Endo, T. 1987 Cloning and expression of a new structural gene for blasticidin S deaminase, a nucleoside aminohydrolase. *Agric. Biol. Chem* **51**, 3165.

Kane, J. F. 1995 Effects of rare codon clusters on high-level expression of heterologous proteins in Escherichia coli. *Curr. Opin. Biotechnol.* **6**, 494-500.

Kao, C. M., Luo, G., Katz, L., Cane, D. E. & Khosla, C. 1995 Manipulation of Macrolide Ring Size by Directed Mutagenesis of a Modular Polyketide Synthase. *J. Am. Chem. Soc.* **117**, 9105-9106.

Kao, C. M., Pieper, R., Cane, D. E. & Khosla, C. 1996 Evidence for two catalytically independent clusters of active sites in a functional modular polyketide synthase. *Biochemistry* **35**, 12363-8.

Katz, L. 1997 Manipulation of Modular Polyketide Synthases. *Chem. Rev.* **97**, 2557-2576.

Katz, L. & Donadio, S. 1993 Polyketide synthesis: prospects for hybrid antibiotics. *Annu. Rev. Microbiol.* **47**, 875-912.

Katz, L. & McDaniel, R. 1999 Novel macrolides through genetic engineering. *Med. Res. Rev.* **19**, 543-58.

Kawai, S., Mori, S., Mukai, T., Hashimoto, W. & Murata, K. 2001 Molecular characterization of *Escherichia coli* NAD kinase. *Eur. J. Biochem.* **268**, 4359-65.

Kawashima, A., Seto, H., Ishiyama, T., Kato, M., Uchida, K. & Otake, N. 1987 Fluorinated blasticidin S. *Agric. Biol. Chem.* **51**, 1183-1184.

Kealey, J. T., Liu, L., Santi, D. V., Betlach, M. C. & Barr, P. J. 1998 Production of a polyketide natural product in nonpolyketide-producing prokaryotic and eukaryotic hosts. *Proc. Natl. Acad. Sci. U S A* **95**, 505-9.

Keating, T. A., Ehmann, D. E., Kohli, R. M., Marshall, C. G., Trauger, J. W. & Walsh, C. T. 2001 Chain termination steps in nonribosomal peptide synthetase assembly lines: directed acyl-S-enzyme breakdown in antibiotic and siderophore biosynthesis. *Chembiochem* **2**, 99-107.

Keating, T. A., Suo, Z., Ehmann, D. E. & Walsh, C. T. 2000 Selectivity of the yersiniabactin synthetase adenylation domain in the two-step process of amino acid activation and transfer to a holo-carrier protein domain. *Biochemistry* **39**, 2297-306.

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

Keen, N. T. & Tamaki, S. 1986 Structure of two pectate lyase genes from *Erwinia chrysanthemi* EC16 and their high-level expression in *Escherichia coli*. *J. Bacteriol.* **168**, 595-606.

Kelecom, A. 2002 Secondary metabolites from marine microorganisms. *An. Acad. Bras. Cienc.* **74**, 151-70.

Khosla, C. 1997 Harnessing the Biosynthetic Potential of Modular Polyketide Synthases. *Chem. Rev.* **97**, 2577-2590.

Kieser, T. B., M.; Buttner, M.; Chater, K.; and Hopwood, D. 2000 *Practical Streptomyces Genetics*. Nowich: The John Innes Foundation.

Kim, C. G., Kirschning, A., Bergon, P., Zhou, P., Su, E., Sauerbrei, B., Ning, S., Ahn, Y., Breuer, M., Leistner, E. & Floss, H. G. 1996 Biosynthesis of 3-Amino-5-hydroxybenzoic Acid, the Precursor of mC_7N Units in Ansamycin Antibiotics. *J. Am. Chem. Soc.* **118**, 7486-7491.

Kim, C. G., Yu, T. W., Fryhle, C. B., Handa, S. & Floss, H. G. 1998 3-Amino-5-hydroxybenzoic acid synthase, the terminal enzyme in the formation of the precursor of mC7N units in rifamycin and related antibiotics. *J Biol Chem* **273**, 6030-40.

Kimura, M., Kamakura, T., Tao, Q. Z., Kaneko, I. & Yamaguchi, I. 1994 Cloning of the blasticidin S deaminase gene (BSD) from *Aspergillus terreus* and its use as a selectable marker for *Schizosaccharomyces pombe* and *Pyricularia oryzae*. *Mol. Gen. Genet.* **242**, 121-9.

Kimura, M., Takatsuki, A. & Yamaguchi, I. 1994 Blasticidin S deaminase gene from *Aspergillus terreus* (BSD): a new drug resistance gene for transfection of mammalian cells. *Biochim. Biophys. Acta* **1219**, 653-9.

Kimura, M. & Yamaguchi, I. 1997 Recent Development in the Use of Blasticidin S, a Microbial Fungicide, as a Useful Reagent in Molecular Biology. *Pest. Biochem. Phys.* **56**, 243-248.

Kobayashi, T., Horinouchi, S., Uozumi, T. & Beppu, T. 1987 Purification and biochemical characterization of streptothricin acetyltransferase coded by the cloned streptothricin-resistance gene of *Streptomyces lavendulae*. *J. Antibiot. (Tokyo)* **40**, 1016-22.

Kohli, R. M., Trauger, J. W., Schwarzer, D., Marahiel, M. A. & Walsh, C. T. 2001 Generality of peptide cyclization catalyzed by isolated thioesterase domains of nonribosomal peptide synthetases. *Biochemistry* **40**, 7099-108.

Kohli, R. M., Walsh, C. T. & Burkart, M. D. 2002 Biomimetic synthesis and optimization of cyclic peptide antibiotics. *Nature* **418**, 658-61.

Kondo, T., Nakai, H. & Goto, T. 1972 Synthesis of Cytosinine, The Nucleoside Component of Antibiotic Blasticidin S. *Tetrahedron Lett.* **19**, 1881-1884.

Konz, D. & Marahiel, M. A. 1999 How do peptide synthetases generate structural diversity? *Chem. Biol.* **6**, R39-48.

Kruger, N., Oppermann, F. B., Lorenzl, H. & Steinbuchel, A. 1994 Biochemical and molecular characterization of the *Clostridium magnum* acetoin dehydrogenase enzyme system. *J. Bacteriol.* **176**, 3614-30.

Kuhstoss, S., Huber, M., Turner, J. R., Paschal, J. W. & Rao, R. N. 1996 Production of a novel polyketide through the construction of a hybrid polyketide synthase. *Gene* **183**, 231-6.

Kurylo-Boroska, Z. & Abramsky, T. 1972 Biosynthesis of beta-Tyrosine. *Biochim. Biophys. Acta.* **264**, 1-10.

Lacalle, R. A., Tercero, J. A. & Jimenez, A. 1992 Cloning of the complete biosynthetic gene cluster for an aminonucleoside antibiotic, puromycin, and its regulated expression in heterologous hosts. *Embo. J.* **11**, 785-92.

Lacalle, R. A., Tercero, J. A., Vara, J. & Jimenez, A. 1993 Identification of the gene encoding an N-acetylpuromycin N-acetylhydrolase in the puromycin biosynthetic gene cluster from *Streptomyces alboniger*. *J. Bacteriol.* **175**, 7474-8.

Lambalot, R. H., Gehring, A. M., Flugel, R. S., Zuber, P., LaCelle, M., Marahiel, M. A., Reid, R., Khosla, C. & Walsh, C. T. 1996 A new enzyme superfamily - the phosphopantetheinyl transferases. *Chem. Biol.* **3**, 923-36.

Larsen, S. H., Berry, D. M., Paschal, J. W. & Gilliam, J. M. 1989 5-Hydroxymethylblasticidin S and blasticidin S from *Streptomyces setonii* culture A83094. *J. Antibiot. (Tokyo)* **42**, 470-1.

Lauer, B., Russwurm, R. & Bormann, C. 2000 Molecular characterization of two genes from *Streptomyces tendae* Tu901 required for the formation of the 4-formyl-4-imidazolin-2-one-containing nucleoside moiety of the peptidyl nucleoside antibiotic nikkomycin. *Eur. J. Biochem.* **267**, 1698-706.

Lauer, B., Russwurm, R., Schwarz, W., Kalmanczhelyi, A., Bruntner, C., Rosemeier, A. & Bormann, C. 2001 Molecular characterization of co-transcribed genes from *Streptomyces tendae* Tu901 involved in the biosynthesis of the peptidyl moiety and assembly of the peptidyl nucleoside antibiotic nikkomycin. *Mol. Gen. Genet.* **264**, 662-73.

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

Lee, E., Shelden, E. A. & Knecht, D. A. 1998 Formation of F-actin aggregates in cells treated with actin stabilizing drugs. *Cell. Motil. Cytoskeleton* **39**, 122-33.

Lee, S. G. & Lipmann, F. 1975 Tyrocidine synthetase system. *Methods Enzymol.* **43**, 585-602.

Lieder, K. W., Booker, S., Ruzicka, F. J., Beinert, H., Reed, G. H. & Frey, P. A. 1998 *S*-Adenosylmethionine-dependent reduction of lysine 2,3-aminomutase and observation of the catalytically functional iron-sulfur centers by electron paramagnetic resonance. *Biochemistry* **37**, 2578-85.

Linne, U., Doekel, S. & Marahiel, M. A. 2001 Portability of epimerization domain and role of peptidyl carrier protein on epimerization activity in nonribosomal peptide synthetases. *Biochemistry* **40**, 15824-34.

Linne, U. & Marahiel, M. A. 2000 Control of directionality in nonribosomal peptide synthesis: role of the condensation domain in preventing misinitiation and timing of epimerization. *Biochemistry* **39**, 10439-47.

Lombardi, P. 1990 A rapid, safe, and convenient procedure for the preparation and use of diazomethane. *Chemistry and Industry*, 708.

Lopez, J. V. 2003 Naturally mosaic operons for secondary metabolite biosynthesis: variability and putative horizontal transfer of discrete catalytic domains of the epothilone polyketide synthase locus. *Mol. Genet. Genomics* **270**, 420-31.

Luo, L., Burkart, M. D., Stachelhaus, T. & Walsh, C. T. 2001 Substrate recognition and selection by the initiation module PheATE of gramicidin S synthetase. *J. Am. Chem. Soc.* **123**, 11208-18.

Luo, L., Kohli, R. M., Onishi, M., Linne, U., Marahiel, M. A. & Walsh, C. T. 2002 Timing of epimerization and condensation reactions in nonribosomal peptide assembly lines: kinetic analysis of phenylalanine activating elongation modules of tyrocidine synthetase B. *Biochemistry* **41**, 9184-96.

Maehr, H. 1997 Combinatorial chemistry in drug research from a new vantage point. *Bioorg. Med. Chem. Lett.* **5**, 473-91.

Malik, V. S. & Vining, L. C. 1972 Chloramphenicol resistance in a chloramphenicol-producing *Streptomyces*. *Can. J. Microbiology* **18**, 583-590.

Malnick, S. D., Beergabel, M. & Lurie, Y. 2000 Treatment of chronic hepatitis C virus infection. *Ann. Pharmacother.* **34**, 1156-64.

Malpartida, F., Hallam, S. E., Kieser, H. M., Motamedi, H., Hutchinson, C. R., Butler, M. J., Sugden, D. A., Warren, M., McKillop, C. & Bailey, C. R. 1987 Homology between Streptomyces genes coding for synthesis of different polyketides used to clone antibiotic biosynthetic genes. *Nature* **325**, 818-21.

Malpartida, F. & Hopwood, D. A. 1984 Molecular cloning of the whole biosynthetic pathway of a Streptomyces antibiotic and its expression in a heterologous host. *Nature* **309**, 462-4.

Mamoun, C. B., Gluzman, I. Y., Goyard, S., Beverley, S. M. & Goldberg, D. E. 1999 A set of independent selectable markers for transfection of the human malaria parasite *Plasmodium falciparum*. *Proc. Natl. Acad. Sci. U S A* **96**, 8716-20.

Mansouri, K. & Piepersberg, W. 1991 Genetics of streptomycin production in Streptomyces griseus: nucleotide sequence of five genes, strFGHIK, including a phosphatase gene. *Mol. Gen. Genet.* **228**, 459-69.

Marahiel, M. A. 1997 Protein templates for the biosynthesis of peptide antibiotics. *Chem. Biol.* **4**, 561-7.

Marchler-Bauer A, A. J., DeWeese-Scott C, Fedorova ND, Geer LY, He S, Hurwitz DI, Jackson JD, Jacobs AR, Lanczycki CJ, Liebert CA, Liu C, Madej T, Marchler GH, Mazumder R, Nikolskaya AN, Panchenko AR, Rao BS, Shoemaker BA, Simonyan V, Song JS, Thiessen RA, Vasudevan S, Wang Y, Yamashita RA, Yin JJ, and Bryant SH. 2003 CDD: a curated Entrez database of conserved domain alignments. *Nucleic Acids Res.* **31**, 383-387.

Mathews, C. K., vanHolde, K. E. & Ahern, K. G. 2000 Nucleotide Metabolism. In *Biochemistry*, pp. 806-808. San Francisco: Addison Wesley Longman, Inc.

Matthey, B., Engert, A., Klimka, A., Diehl, V. & Barth, S. 1999 A new series of pET-derived vectors for high efficiency expression of *Pseudomonas* exotoxin-based fusion proteins. *Gene* **229**, 145-53.

Mazodier, P., Petter, R. & Thompson, C. 1989 Intergeneric conjugation between *Escherichia coli* and *Streptomyces* species. *J. Bacteriol.* **171**, 3583-5.

McCarter, J. D. & Withers, S. G. 1994 Mechanisms of enzymatic glycoside hydrolysis. *Curr. Opin. Struct. Biol.* **4**, 885-92.

McDaniel, R., Kao, C. M., Hwang, S. J. & Khosla, C. 1997 Engineered intermodular and intramodular polyketide synthase fusions. *Chem. Biol.* **4**, 667-74.

McDaniel, R., Thamchaipenet, A., Gustafsson, C., Fu, H., Betlach, M. & Ashley, G. 1999 Multiple genetic modifications of the erythromycin polyketide synthase to produce a library of novel "unnatural" natural products. *Proc. Natl. Acad. Sci. U S A* **96**, 1846-51.

Miller, D. A., Luo, L., Hillson, N., Keating, T. A. & Walsh, C. T. 2002 Yersiniabactin synthetase: a four-protein assembly line producing the nonribosomal peptide/polyketide hybrid siderophore of *Yersinia pestis*. *Chem. Biol.* **9**, 333-44.

Moore, B. S. & Hertweck, C. 2002 Biosynthesis and attachment of novel bacterial polyketide synthase starter units. *Nat. Prod. Rep.* **19**, 70-99.

Mootz, H. D. & Marahiel, M. A. 1997 The tyrocidine biosynthesis operon of Bacillus brevis: complete nucleotide sequence and biochemical characterization of functional internal adenylation domains. *J. Bacteriol.* **179**, 6843-50.

Mootz, H. D., Schwarzer, D. & Marahiel, M. A. 2000 Construction of hybrid peptide synthetases by module and domain fusions. *Proc. Natl. Acad. Sci. U S A* **97**, 5848-53.

Mosher, R. H., Camp, D. J., Yang, K., Brown, M. P., Shaw, W. V. & Vining, L. C. 1995 Inactivation of chloramphenicol by O-phosphorylation. A novel resistance mechanism in *Streptomyces venezuelae* ISP5230, a chloramphenicol producer. *J. Biol. Chem.* **270**, 27000-6.

Munch-Petersen, A. 1983 *Metabolism of Nucleotides, Nucleosides and Nucleobases in Microorganisms*. San Diego: Academic Press Inc.

Munro, M. H., Blunt, J. W., Dumdei, E. J., Hickford, S. J., Lill, R. E., Li, S., Battershill, C. N. & Duckworth, A. R. 1999 The discovery and development of marine compounds with pharmaceutical potential. *J. Biotechnol.* **70**, 15-25.

Nakamura, Y., Gojobori, T. & Ikemura, T. 2000 Codon usage tabulated from international DNA sequence databases: status for the year 2000. *Nucleic Acids Res.* **28**, 292.

Nakano, M. M., Corbell, N., Besson, J. & Zuber, P. 1992 Isolation and characterization of sfp: a gene that functions in the production of the lipopeptide biosurfactant, surfactin, in *Bacillus subtilis*. *Mol. Gen. Genet.* **232**, 313-21.

Nakano, Y. & Kitaoka, S. 1971 L-aspartate -decarboxylase in a cell-free system from *Escherichia coli*. *J. Biochem. (Tokyo)* **70**, 327-34.

Neilan, B. A., Dittmann, E., Rouhiainen, L., Bass, R. A., Schaub, V., Sivonen, K. & Borner, T. 1999 Nonribosomal peptide synthesis and toxigenicity of cyanobacteria. *J. Bacteriol.* **181**, 4089-97.

Newman, D. J., Cragg, G. M. & Snader, K. M. 2003 Natural products as sources of new drugs over the period 1981-2002. *J. Nat. Prod.* **66**, 1022-37.

Noda, T., Take, T., Nagata, A., Wakamiya, T. & Shiba, T. 1972 Chemical studies on tuberactinomycin. 3. The chemical structure of viomycin (tuberactinomycin B). *J. Antibiot. (Tokyo)* **25**, 427-8.

Nomoto, S. & Shiba, T. 1978 Synthesis of N-B-Methyl-L-B-arginine, a component in a new antibiotic LL-BM547B. *Chemistry Lett.*, 589-590.

Nomoto, S. & Shimoyama, A. 2001 First synthesis of blastidic acid, a component amino acid in an antibiotic, blasticidin S. *Tetrahedron Lett.* **42**, 1753-1755.

Nomura, H., Paik, S. Y., Izuhara, M. & Sakagami, Y. 1991 Resistant Mechanism to Blasticidin S in *Streptomyces lividans*. *Korean Biochem. J.* **24**, 327-330.

Odaka, C., Sanders, M. L. & Crews, P. 2000 Jasplakinolide induces apoptosis in various transformed cell lines by a caspase-3-like protease-dependent pathway. *Clin. Diagn. Lab. Immunol.* **7**, 947-52.

Okuyama, K., Shibuya, S., Hamamoto, T. & Noguchi, T. 2003 Enzymatic synthesis of 2'-deoxyguanosine with nucleoside deoxyribosyltransferase-II. *Biosci. Biotechnol. Biochem.* **67**, 989-95.

Oliynyk, M., Brown, M. J., Cortes, J., Staunton, J. & Leadlay, P. F. 1996 A hybrid modular polyketide synthase obtained by domain swapping. *Chem. Biol.* **3**, 833-9.

Omura, S., Ikeda, H., Ishikawa, J., Hanamoto, A., Takahashi, C., Shinose, M., Takahashi, Y., Horikawa, H., Nakazawa, H., Osonoe, T., Kikuchi, H., Shiba, T., Sakaki, Y. & Hattori, M. 2001 Genome sequence of an industrial microorganism *Streptomyces avermitilis*: deducing the ability of producing secondary metabolites. *Proc. Natl. Acad. Sci. U S A* **98**, 12215-20.

Otake, N., Takeuchi, S., Endo, T. & Yonehara, H. 1965a The structure of blasticidin S. *Tetrahedron Lett.* **19**, 1411-9.

Otake, N., Takeuchi, S., Endo, T. & Yonehara, H. 1965b Structure of cytosinine, a structural component of blasticidin S. *Tetrahedron Lett.* **19**, 1405-9.

Parry, R. J. & Kurylo-Boroska, Z. 1980 Biosynthesis of Amino Acids. Investigation of the Mechanism of B-Tyrosine Formation. *J. Am. Chem. Soc.* **102**, 836-837.

Perez-Gonzalez, J. A., Ruiz, D., Esteban, J. A. & Jimenez, A. 1990 Cloning and characterization of the gene encoding a blasticidin S acetyltransferase from *Streptoverticillum sp. Gene* **86**, 129-34.

Piel, J., Hertweck, C., Shipley, P. R., Hunt, D. M., Newman, M. S. & Moore, B. S. 2000 Cloning, sequencing and analysis of the enterocin biosynthesis gene cluster from the marine isolate '*Streptomyces maritimus*': evidence for the derailment of an aromatic polyketide synthase. *Chem. Biol.* **7**, 943-55.

Pissowotzki, K., Mansouri, K. & Piepersberg, W. 1991 Genetics of streptomycin production in *Streptomyces griseus*: molecular structure and putative function of genes strELMB2N. *Mol. Gen. Genet.* **231**, 113-23.

Porse, B. T., Rodriguez-Fonseca, C., Leviev, I. & Garrett, R. A. 1995 Antibiotic inhibition of the movement of tRNA substrates through a peptidyl transferase cavity. *Biochem. Cell. Biol.* **73**, 877-85.

Porter, D. J., Merrill, B. M. & Short, S. A. 1995 Identification of the active site nucleophile in nucleoside 2-deoxyribosyltransferase as glutamic acid 98. *J. Biol. Chem.* **270**, 15551-6.

Porter, D. J. & Short, S. A. 1995 Nucleoside 2-deoxyribosyltransferase. Pre-steady-state kinetic analysis of native enzyme and mutant enzyme with an alanyl residue replacing Glu-98. *J. Biol. Chem.* **270**, 15557-62.

Pospiech, A., Bietenhader, J. & Schupp, T. 1996 Two multifunctional peptide synthetases and an O-methyltransferase are involved in the biosynthesis of the DNA-binding antibiotic and antitumour agent saframycin Mx1 from *Myxococcus xanthus*. *Microbiology* **142 (Pt 4)**, 741-6.

Poston, J. M. 1976 Leucine 2,3-aminomutase, an enzyme of leucine catabolism. *J. Biol. Chem.* **251**, 1859-63.

Poston, J. M. 1977 Leucine 2,3-aminomutase: A Cobalamin-Dependent Enzyme Present in Bean Seedlings. *Science* **195**, 301-302.

Poston, J. M. 1978 Coenzyme B_{12} Dependent Enzymes in Potatoes: Leucine 2,3-aminomutase and methylmalonyl-CoA mutase. *Phytochemistry* **17**, 401-402.

Poston, J. M. 1980 Distribution of leucine 2,3-aminomutase activity in various organs of the rat and in subcellular organelles of rat liver. *Biochem. Biophys. Res. Commun.* **96**, 838-43.

Poston, J. M. 1986 Beta-leucine and the beta-keto pathway of leucine metabolism. *Adv. Enzymol. Relat. Areas Mol. Biol.* **58**, 173-89.

Poston, J. M. & Hemmings, B. A. 1979 Cobalamins and cobalamin-dependent enzymes in *Candida utilis*. *J. Bacteriol.* **140**, 1013-6.

Prabhakaran, P. C. 1988 Biosynthesis of Blasticidin S: Mechanistic and Stereochemical Investigations. In *Chemistry*, pp. 144. Corvallis: Oregon State University.

Prabhakaran, P. C., Woo, N., Yorgey, P. & Gould, S. J. 1988 Biosynthesis of Blasticidin S from L-alpha-arginine. Stereochemistry in the Arginine 2,3-aminomutase Reaction. *J. Am. Chem. Soc.* **110**, 5785-5791.

Pradella, S., Hans, A., Sproer, C., Reichenbach, H., Gerth, K. & Beyer, S. 2002 Characterisation, genome size and genetic manipulation of the myxobacterium *Sorangium cellulosum* So ce56. *Arch. Microbiol.* **178**, 484-92.

Priefert, H., Hein, S., Kruger, N., Zeh, K., Schmidt, B. & Steinbuchel, A. 1991 Identification and molecular characterization of the *Alcaligenes eutrophus* H16 aco operon genes involved in acetoin catabolism. *J. Bacteriol.* **173**, 4056-71.

Proksch, P., Edrada, R. A. & Ebel, R. 2002 Drugs from the seas - current status and microbiological implications. *Appl. Microbiol. Biotechnol.* **59**, 125-34.

Pugsley, A. P. 1987 Genetic instability in streptomycetes. *Microbiol. Sci.* **4**, 11-3.

Rajendran, N. 1999 Identification and cloning of a gene locus encoding peptide synthetase of *Pseudomonas fluorescens* by two sets of PCR primers. *Z. Naturforsch [C]* **54**, 105-9.

Recktenwald, J., Shawky, R., Puk, O., Pfennig, F., Keller, U., Wohlleben, W. & Pelzer, S. 2002 Nonribosomal biosynthesis of vancomycin-type antibiotics: a heptapeptide backbone and eight peptide synthetase modules. *Microbiology* **148**, 1105-18.

Reynolds, K. A. 1998 Combinatorial biosynthesis: lesson learned from nature. *Proc. Natl. Acad. Sci. U S A* **95**, 12744-6.

Rodriguez-Fonseca, C., Amils, R. & Garrett, R. A. 1995 Fine structure of the peptidyl transferase centre on 23 S-like rRNAs deduced from chemical probing of antibiotic-ribosome complexes. *J. Mol. Biol.* **247**, 224-35.

Rondon, M. R., August, P. R., Bettermann, A. D., Brady, S. F., Grossman, T. H., Liles, M. R., Loiacono, K. A., Lynch, B. A., MacNeil, I. A., Minor, C., Tiong, C. L., Gilman, M., Osburne, M. S., Clardy, J., Handelsman, J. & Goodman, R. M. 2000 Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. *Appl. Environ. Microbiol.* **66**, 2541-7.

Rotsch, C. & Radmacher, M. 2000 Drug-induced changes of cytoskeletal structure and mechanics in fibroblasts: an atomic force microscopy study. *Biophys. J.* **78**, 520-35.

Ruan, X., Pereda, A., Stassi, D. L., Zeidner, D., Summers, R. G., Jackson, M., Shivakumar, A., Kakavas, S., Staver, M. J., Donadio, S. & Katz, L. 1997 Acyltransferase domain substitutions in erythromycin polyketide synthase yield novel erythromycin derivatives. *J. Bacteriol.* **179**, 6416-25.

Rubio, M. A., Espinosa, J. C., Tercero, J. A. & Jimenez, A. 1998 The Pur10 protein encoded in the gene cluster for puromycin biosynthesis of *Streptomyces alboniger* is an NAD-dependent ATP dehydrogenase. *FEBS Lett.* **437**, 197-200.

Rudd, B. A. & Hopwood, D. A. 1979 Genetics of actinorhodin biosynthesis by *Streptomyces coelicolor* A3(2). *J. Gen. Microbiol.* **114**, 35-43.

Ruzicka, F. J., Lieder, K. W. & Frey, P. A. 2000 Lysine 2,3-aminomutase from Clostridium subterminale SB4: mass spectral characterization of cyanogen bromide-treated peptides and cloning, sequencing, and expression of the gene kamA in *Escherichia coli*. *J. Bacteriol.* **182**, 469-76.

Saldanha, S. A., Birch, L. M., Webb, M. E., Nabbs, B. K., von Delft, F., Smith, A. G. & Abell, C. 2001 Identification of Tyr58 as the proton donor in the aspartate-alpha-decarboxylase reaction. *Chem. Commun.*, 1760-1.

Salomon, C. E., Magarvey, N. A. & Sherman, D. H. 2004 Merging the potential of microbial genetics with biological and chemical diversity: an even brighter future for marine natural product drug discovery. *Nat. Prod. Rep.* **21**, 105-21.

Sambrook, J. & Russell, D. 2001 *Molecular Cloning A Laboratory Manual.* Cold Spring Harbor: Cold Spring Harbor Laboratory Press.

Saugar, I., Sanz, E., Rubio, M. A., Espinosa, J. C. & Jimenez, A. 2002 Identification of a set of genes involved in the biosynthesis of the aminonucleoside moiety of antibiotic A201A from *Streptomyces capreolus*. *Eur. J. Biochem.* **269**, 5527-35.

Sawada, H. S., T.; Akiyama, SI.; and Nakao, Y. 1985a Biosynthetic Pathway of 5-Hydroxymethylcytosine by *Streptoverticillium rimofaciens*. *J. Ferment. Technol.* **63**, 23-27.

Sawada, H. S., T.; Akiyama, SI.; and Nakao, Y. 1985b Formation of 5-Hydroxymethylcytosine from Cytidine 5'-Mono-Phosphate by *Streptoverticillium rimofaciens*. *J. Ferment. Technol.* **63**, 17-21.

Schauwecker, F., Pfennig, F., Grammel, N. & Keller, U. 2000 Construction and in vitro analysis of a new bi-modular polypeptide synthetase for synthesis of *N*-methylated acyl peptides. *Chem. Biol.* **7**, 287-97.

Schmidt, E. W., Obraztsova, A. Y., Davidson, S. K., Faulkner, D. J. & Haygood, M. G. 2000 Identification of the antifungal peptide-containing symbiont of the marine sponge *Theonella swinhoei* as a novel δ-proteobacterium, "*Candidatus Entotheonella palauensis*". *Marine Biology* **136**, 969-977.

Schwarzer, D., Finking, R. & Marahiel, M. A. 2003 Nonribosomal peptides: from genes to products. *Nat. Prod. Rep.* **20**, 275-87.

Schwarzer, D. & Marahiel, M. A. 2001 Multimodular biocatalysts for natural product assembly. *Naturwissenschaften* **88**, 93-101.

Senderowicz, A. M., Kaur, G., Sainz, E., Laing, C., Inman, W. D., Rodriguez, J., Crews, P., Malspeis, L., Grever, M. R., Sausville, E. A. & et al. 1995 Jasplakinolide's inhibition of the growth of prostate carcinoma cells in vitro with disruption of the actin cytoskeleton. *J. Natl. Cancer Inst.* **87**, 46-51.

Seno, E. T. & Baltz, R. M. 1989 Structural Organization and regulation of antibiotic biosynthesis and resistance genes in actinomycetes. In *Regulation of Secondary Metabolism in Actinomycetes* (ed. S. Shapiro), pp. 1-48. Boca Raton: CRC Press.

Seto, H., Furihata, K. & Yonehara, H. 1976 Studies on the biosynthesis of blasticidin S. V. Isolation and structure of pentopyranic acid. *J. Antibiot. (Tokyo)* **29**, 595-6.

Seto, H., Otake, N. & Yonehara, H. 1968 Biosynthesis of blasticidin S. II. Leucylblasticidin S, a metabolic intermediate of blasticidin S biosynthesis. *Agric. Biol Chem.* **32**, 1299-305.

Seto, H., Yamaguchi, N., Otake, N. & Yonehara, H. 1968 Studies on the biosynthesis of blasticidin S. Part I. Precursors of blasticidin S biosynthesis. *Agric. Biol. Chem.* **32**, 1292-1298.

Seto, H. & Yonehara, H. 1977a Studies on the biosynthesis of blasticidin S. VI. The isolation and structure of blasticidin H. *J. Antibiot. (Tokyo)* **30**, 1019-21.

Seto, H. & Yonehara, H. 1977b Studies on the biosynthesis of blasticidin S. VII. Isolation of demethylblasticidin S. *J. Antibiot. (Tokyo)* **30**, 1022-4.

Sheikh, S., Gratzer, W. B., Pinder, J. C. & Nash, G. B. 1997 Actin polymerisation regulates integrin-mediated adhesion as well as rigidity of neutrophils. *Biochem. Biophys. Res. Commun.* **238**, 910-5.

Shen, Z. & Byers, D. M. 1996 Isolation of *Vibrio harveyi* acyl carrier protein and the fabG, acpP, and fabF genes involved in fatty acid biosynthesis. *J. Bacteriol.* **178**, 571-3.

Shomura, T., Inoue, M., Niida, T. & Hara, T. 1964 The Improved Assay Method of Blasticidin S. *J. Antibiot. (Tokyo)* **17**, 253-61.

Short, S. A., Armstrong, S. R., Ealick, S. E. & Porter, D. J. 1996 Active site amino acids that participate in the catalytic mechanism of nucleoside 2'-deoxyribosyltransferase. *J. Biol. Chem.* **271**, 4978-87.

Sieber, S. A., Linne, U., Hillson, N. J., Roche, E., Walsh, C. T. & Marahiel, M. A. 2002 Evidence for a monomeric structure of nonribosomal Peptide synthetases. *Chem. Biol.* **9**, 997-1008.

Silakowski, B., Kunze, B. & Muller, R. 2001 Multiple hybrid polyketide synthase/non-ribosomal peptide synthetase gene clusters in the myxobacterium *Stigmatella aurantiaca*. *Gene* **275**, 233-40.

Silakowski, B., Nordsiek, G., Kunze, B., Blocker, H. & Muller, R. 2001 Novel features in a combined polyketide synthase/non-ribosomal peptide synthetase: the . myxalamid biosynthetic gene cluster of the myxobacterium *Stigmatella aurantiaca* Sga15. *Chem. Biol.* **8**, 59-69.

Silakowski, B., Schairer, H. U., Ehret, H., Kunze, B., Weinig, S., Nordsiek, G., Brandt, P., Blocker, H., Hofle, G., Beyer, S. & Muller, R. 1999 New lessons for combinatorial biosynthesis from myxobacteria. The myxothiazol biosynthetic gene cluster of *Stigmatella aurantiaca* DW4/3-1. *J. Biol. Chem.* **274**, 37391-9.

Skinner, R., Cundliffe, E. & Schmidt, F. J. 1983 Site of action of a ribosomal RNA methylase responsible for resistance to erythromycin and other antibiotics. *J. Biol. Chem.* **258**, 12702-6.

Skinner, R. H. & Cundliffe, E. 1980 Resistance to the antibiotics viomycin and capreomycin in the *Streptomyces* species which produce them. *J. Gen. Microbiol.* **120**, 95-104.

Smyth, D. R., Mrozkiewicz, M. K., McGrath, W. J., Listwan, P. & Kobe, B. 2003 Crystal structures of fusion proteins with large-affinity tags. *Protein Sci.* **12**, 1313-22.

Snider, M. R., L.; Wolfenden, R.; and Cleland, W. 2002 ^{15}N Kinetic Isotope Effects on Uncatalyzed and Enzymatic Deamination of Cytidine. *Biochemistry* **41**, 415-421.

Sofia, H. J., Chen, G., Hetzler, B. G., Reyes-Spindola, J. F. & Miller, N. E. 2001 Radical SAM, a novel protein superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: functional characterization using new analysis and information visualization methods. *Nucleic Acids Res.* **29**, 1097-106.

Spector, I., Braet, F., Shochet, N. R. & Bubb, M. R. 1999 New anti-actin drugs in the study of the organization and function of the actin cytoskeleton. *Microsc. Res. Tech.* **47**, 18-37.

Stachelhaus, T. & Marahiel, M. A. 1995a Modular structure of genes encoding multifunctional peptide synthetases required for non-ribosomal peptide synthesis. *FEMS Microbiol. Lett.* **125**, 3-14.

Stachelhaus, T. & Marahiel, M. A. 1995b Modular structure of peptide synthetases revealed by dissection of the multifunctional enzyme GrsA. *J. Biol. Chem.* **270**, 6163-9.

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

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

Stadtman, T. 1962 Lysine Fermentation to Fatty Acids and Ammonia: A Coamide Coenzyme-dependent Process. *J. Biol. Chem.* **237**, PC2409-PC2411.

Stadtman, T. 1964 Cobamide Coenzyme Requirement for the Anaerobic Degredation of Lysine. *Ann. N. Y. Acad. Sci.* **112**, 728-734.

Staunton, J., Caffrey, P., Aparicio, J. F., Roberts, G. A., Bethell, S. S. & Leadlay, P. F. 1996 Evidence for a double-helical structure for modular polyketide synthases. *Nat. Struct. Biol.* **3**, 188-92.

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

Stein, T., Vater, J., Kruft, V., Otto, A., Wittmann-Liebold, B., Franke, P., Panico, M., McDowell, R. & Morris, H. R. 1996 The multiple carrier model of nonribosomal peptide biosynthesis at modular multienzymatic templates. *J. Biol. Chem.* **271**, 15428-35.

Strohl, W. 1992 Compilation and analysis of DNA sequences associated with apparent streptomycete promoters. *Nucleic Acids Res.* **20**, 961-974.

Strohl, W. 1997 *Biotechnology of Antibiotics*. New York: Marcel Dekker, Inc.

Sugiyama, M., Mochizuki, H., Nimi, O. & Nomi, R. 1981 Mechanism of protection of protein synthesis against streptomycin inhibition in a producing strain. *J. Antibiot. (Tokyo)* **34**, 1183-8.

Sugiyama, M., Sakamoto, M., Mochizuki, H., Nimi, O. & Nomi, R. 1983 Purification and Characterization of Streptomycin 6-kinase, an Enzyme Implicated in Self-protection of a Streptomycin-producing Micro-organism. *J. Gen. Micorbiol.* **129**, 1683-1687.

Sugiyama, M., Takeda, A., Paik, S. Y., Nimi, O. & Nomi, R. 1986 Acetylation of blasticidin S by its producing actinomycetes. *J. Antibiot. (Tokyo)* **39**, 827-32.

Suhadolnik, R. J. 1970 *Nucleoside Antibiotics*. New York: Wiley-Interscience.

Swaminathan, V., Smith, J. L., Sundaralingam, M., Coutsogeorgopoulos, C. & Kartha, G. 1981 Crystal and molecular structure of the antibiotic blasticidin S hydrochloride pentahydrate. *Biochim. Biophys. Acta.* **655**, 335-41.

Takana, Y., Sakagami, Y., Nishimura, T., Yamaki, H. & Umezawa, H. 1961 Activity of cytomycin and blasticidin S against transplantable animal tumors. *J. Antibiot.* **14**, 123-126.

Takeuchi, H., Ara, G., Sausville, E. A. & Teicher, B. 1998 Jasplakinolide: interaction with radiation and hyperthermia in human prostate carcinoma and Lewis lung carcinoma. *Cancer Chemother. Pharmacol.* **42**, 491-6.

Takeuchi, S., Hirayama, K., Ueda, K., Sakai, H. & Yonehara, H. 1958 Blasticidin S, a new antibiotic. *J. Antibiot. (Tokyo)* **11**, 1-5.

Talpir, R., Benayahu, Y., Kashman, Y., Pannell, L. & Schleyer, M. 1994 Hemiasterlin and Geodiamolide TA; Two New Cytotoxic Peptides from the Marine Sponge *Hemiasterella minor*. *Tetrahedron Lett.* **35**, 4453-4456.

Tamura, K., Kimura, M. & Yamaguchi, I. 1995 Blasticidin S deaminase gene (BSD): a new selection marker gene for transformation of *Arabidopsis thaliana* and *Nicotiana tabacum*. *Biosci. Biotechnol. Biochem.* **59**, 2336-8.

Tapuhi, Y., Schmidt, D. E., Lindner, W. & Karger, B. L. 1981 Dansylation of amino acids for high-performance liquid chromatography analysis. *Anal. Biochem.* **115**, 123-9.

Taylor, D. P., Eckhardt, T. & Fare, L. R. 1983 Gene cloning in the actinomycetes. *Ann. N. Y. Acad. Sci.* **413**, 47-56.

Tercero, J. A., Espinosa, J. C., Lacalle, R. A. & Jimenez, A. 1996 The biosynthetic pathway of the aminonucleoside antibiotic puromycin, as deduced from the molecular analysis of the pur cluster of *Streptomyces alboniger*. *J. Biol. Chem.* **271**, 1579-90.

Tercero, J. A., Lacalle, R. A. & Jimenez, A. 1993 The pur8 gene from the pur cluster of *Streptomyces alboniger* encodes a highly hydrophobic polypeptide which confers resistance to puromycin. *Eur. J. Biochem.* **218**, 963-71.

Thiruvengadam, T. K., Gould, S. J., Aberhart, D. J. & Lin, H. 1983 Biosynthesis of Streptothricin F. 5. Formation of β-Lysine by *Streptomyces* L-1689-23. *J. Am. Chem. Soc.* **105**, 5470-5476.

Thompson, C. J., Skinner, R. H., Thompson, J., Ward, J. M., Hopwood, D. A. & Cundliffe, E. 1982 Biochemical characterization of resistance determinants cloned from antibiotic-producing streptomycetes. *J. Bacteriol.* **151**, 678-85.

Thompson, J., Cundliffe, E. & Stark, M. J. 1982 The mode of action of berninamycin and mechanism of resistance in the producing organism, *Streptomyces bernensis*. *J. Gen. Microbiol.* **128 (Pt 4)**, 875-84.

Thompson, J., Schmidt, F. & Cundliffe, E. 1982 Site of action of a ribosomal RNA methylase conferring resistance to thiostrepton. *J. Biol. Chem.* **257**, 7915-7.

Trauger, J. W., Kohli, R. M., Mootz, H. D., Marahiel, M. A. & Walsh, C. T. 2000 Peptide cyclization catalysed by the thioesterase domain of tyrocidine synthetase. *Nature* **407**, 215-8.

Trauger, J. W., Kohli, R. M. & Walsh, C. T. 2001 Cyclization of backbone-substituted peptides catalyzed by the thioesterase domain from the tyrocidine nonribosomal peptide synthetase. *Biochemistry* **40**, 7092-8.

Tseng, C. C., Bruner, S. D., Kohli, R. M., Marahiel, M. A., Walsh, C. T. & Sieber, S. A. 2002 Characterization of the surfactin synthetase C-terminal thioesterase domain as a cyclic depsipeptide synthase. *Biochemistry* **41**, 13350-9.

Tsuji, S. Y., Cane, D. E. & Khosla, C. 2001 Selective protein-protein interactions direct channeling of intermediates between polyketide synthase modules. *Biochemistry* **40**, 2326-31.

Tsuji, S. Y., Wu, N. & Khosla, C. 2001 Intermodular communication in polyketide synthases: comparing the role of protein-protein interactions to those in other multidomain proteins. *Biochemistry* **40**, 2317-25.

Uhlen, M. & Moks, T. 1990 Gene fusions for purpose of expression: an introduction. *Methods Enzymol.* **185**, 129-43.

Unson, M. D. & Faulkner, D. J. 1993 Cyanobacterial symbiont of chlorinated metabolites from *Dysidea herbacea* (Porifera). *Experientia* **49**, 349-352.

Vara, J., Malpartida, F., Hopwood, D. A. & Jimenez, A. 1985 Cloning and expression of a puromycin N-acetyl transferase gene from *Streptomyces alboniger* in *Streptomyces lividans* and *Escherichia coli*. *Gene* **33**, 197-206.

Vo, N. V., Young, K. C. & Lai, M. M. 2003 Mutagenic and inhibitory effects of ribavirin on hepatitis C virus RNA polymerase. *Biochemistry* **42**, 10462-71.

Wakamiya, T. T., T.; Sakakibara, H.; Fukukawa, K.; and Shiba, T. 1977 Chemical Studies on Tuberactinomycin. XI. 1. Semisynthesis of Tuberactinomycin Analogs with Various Amino Acids in the Branched Part. *Bull. Chem. Soc. Japan* **50**, 1984-1989.

Wakamiya, T. U., H.; Teshima, T.; and Shiba, T. 1975 Synthesis of Acyl Derivatives of B-Lysine for Peptide Synthesis. *Bull. Chem. Soc. Japan* **48**, 2401-2402.

Walker, M. S. & Walker, J. B. 1971 Streptomycin Biosynthesis: separation and substrate specificities of phosphatases acting on guanidino-deoxy-*scyllo*-inositol posphate and streptomycin -(streptidino)-phosphate. *J. Biol. Chem.* **246**, 7034-7040.

Walsh, C. T. 1979 Reactions at the alpha carbon of amino acids. In *Enzymatic Reaction Mechanisms*, pp. 801. San Francisco: W.H. Freeman and Company.

Walsh, C. T. 2004 Polyketide and nonribosomal peptide antibiotics: modularity and versatility. *Science* **303**, 1805-10.

Walsh, C. T., Gehring, A. M., Weinreb, P. H., Quadri, L. E. & Flugel, R. S. 1997 Post-translational modification of polyketide and nonribosomal peptide synthases. *Curr. Opin. Chem. Biol.* **1**, 309-15.

Weber, G., Schorgendorfer, K., Schneider-Scherzer, E. & Leitner, E. 1994 The peptide synthetase catalyzing cyclosporine production in *Tolypocladium niveum* is encoded by a giant 45.8-kilobase open reading frame. *Curr. Genet.* **26**, 120-5.

Weinreb, P. H., Quadri, L. E., Walsh, C. T. & Zuber, P. 1998 Stoichiometry and specificity of in vitro phosphopantetheinylation and aminoacylation of the valine-activating module of surfactin synthetase. *Biochemistry* **37**, 1575-84.

Weisblum, B. & Davies, J. 1968 Antibiotic Inhibitors of the Bacterial Ribosome. *Bacteriol. Rev.* **32**, 493-528.

Williamson, J. M. & Brown, G. M. 1979 Purification and properties of L-Aspartate-alpha-decarboxylase, an enzyme that catalyzes the formation of beta-alanine in *Escherichia coli*. *J. Biol. Chem.* **254**, 8074-82.

Winter, E., Brummel, M., Schuch, R. & Spener, F. 1997 Decarboxylation of malonyl-(acyl carrier protein) by 3-oxoacyl-(acyl carrier protein) synthases in plant fatty acid biosynthesis. *Biochem. J.* **321 (Pt 2)**, 313-8.

Wu, G., Williams, H. D., Zamanian, M., Gibson, F. & Poole, R. K. 1992 Isolation and characterization of *Escherichia coli* mutants affected in aerobic respiration: the cloning and nucleotide sequence of ubiG. Identification of an *S*-adenosylmethionine-binding motif in protein, RNA, and small-molecule methyltransferases. *J. Gen. Microbiol.* **138 (Pt 10)**, 2101-12.

Wu, N., Tsuji, S. Y., Cane, D. E. & Khosla, C. 2001 Assessing the balance between protein-protein interactions and enzyme-substrate interactions in the channeling of intermediates between polyketide synthase modules. *J. Am. Chem. Soc.* **123**, 6465-74.

Xiang, L. & Moore, B. S. 2003 Characterization of benzoyl coenzyme A biosynthesis genes in the enterocin-producing bacterium "*Streptomyces maritimus*". *J. Bacteriol.* **185**, 399-404.

Xue, Q., Ashley, G., Hutchinson, C. R. & Santi, D. V. 1999 A multiplasmid approach to preparing large libraries of polyketides. *Proc. Natl. Acad. Sci. U S A* **96**, 11740-5.

Yamaguchi, H. & Tanaka, N. 1966 Inhibition of protein synthesis by blasticidin S. II. Studies on the site of action in *E. coli* polypeptide synthesizing systems. *J. Biochem. (Tokyo)* **60**, 632-42.

Yamaguchi, H., Yamamoto, C. & Tanaka, N. 1965 Inhibition of protein synthesis by blasticidin S. I. Studies with cell-free systems from bacterial and mammalian cells. *J. Biochem. (Tokyo)* **57**, 667-77.

Yamaguchi, I. & Misato, T. 1985 Active Center and Mode of Reaction of Blasticidin S Deaminase. *Agric. Biol. Chem.* **49**, 3355-3361.

Yamaguchi, I., Seto, H. & Misato, T. 1986 Substrate Binding by Blasticidin S Deaminase, an Aminohydrolase for Novel 4-aminopyrimidine Nucleosides. *Pest. Biochem. Phys.* **25**, 54-62.

Yamaguchi, I., Shibata, H., Seto, H. & Misato, T. 1975 Isolation and purification of blasticidin S deaminase from Aspergillus terreus. *J. Antibiot. (Tokyo)* **28**, 7-14.

Yamaguchi, I., Takagi, K. & Misato, T. 1972 The Sites of Degradation of Blasticidin S. *Agric. Biol. Chem.* **36**, 1719-1727.

Yang, C., Carlow, D., Wolfenden, R. & Short, S. A. 1992 Cloning and nucleotide sequence of the *Escherichia coli* cytidine deaminase (ccd) gene. *Biochemistry* **31**, 4168-74.

Yin, X., O'Hare, T., Gould, S. J. & Zabriskie, T. M. 2003 Identification and cloning of genes encoding viomycin biosynthesis from *Streptomyces vinaceus* and evidence for involvement of a rare oxygenase. *Gene* **312**, 215-24.

Yonehara, H. 1984 Blasticidin S: Properties, Biosynthesis, and Fermentation. In *Biotechnology of Industrial Antibiotics*, vol. 22 (ed. E. Vandamme). New York: Marcel Dekker, Inc.

Yonehara, H. & Otake, N. 1965 Biological and chemical reconstruction of blasticidin S. *Antimicrob. Agents Chemother.*, 855-857.

Yu, S., Fiss, E. & Jacobs, W. R., Jr. 1998 Analysis of the exochelin locus in *Mycobacterium smegmatis*: biosynthesis genes have homology with genes of the peptide synthetase family. *J. Bacteriol.* **180**, 4676-85.

Yu, T. W., Muller, R., Muller, M., Zhang, X., Draeger, G., Kim, C. G., Leistner, E. & Floss, H. G. 2001 Mutational analysis and reconstituted expression of the biosynthetic genes involved in the formation of 3-amino-5-hydroxybenzoic acid, the starter unit of rifamycin biosynthesis in *Amycolatopsis mediterranei* S699. *J. Biol. Chem.* **276**, 12546-55.

Zabriskie, T. M., Klocke, J., Ireland, C. M., Marcus, A., Molinski, T., Faulkner, J., Xu, C. & Clardy, J. 1986 Jaspamide, a Modified Peptide from a Jaspis Sponge, with Insecticidal and Antifungal Activity. *J. Am. Chem. Soc.* **108**, 3123-3124.

Zampella, A., Giannini, C., Debitus, C., Roussakis, C. & D'Auria, M. V. 1999 New jaspamide derivatives from the marine sponge *Jaspis splendans* collected in Vanuatu. *J. Nat. Prod.* **62**, 332-4.

Zhang, D. & Miller, M. J. 1999 Polyoxins and nikkomycins: progress in synthetic and biological studies. *Curr. Pharm. Des.* **5**, 73-99.

Zhang, J. H., Quigley, N. B. & Gross, D. C. 1997 Analysis of the syrP gene, which regulates syringomycin synthesis by *Pseudomonas syringae pv. syringae*. *Appl. Environ. Microbiol.* **63**, 2771-8.

Zhang, Q. 1998 Advanced Studies of Blasticidin S Biosynthesis. In *Chemistry*, pp. 159. Corvallis: Oregon State University.

Zhang, Q., Cone, M. C., Gould, S. J. & Zabriskie, T. M. 2000 Reevaluation of the Final Steps in the Biosynthesis of Blasticidin S by *Streptomyces griseochromogenes* and Identification of a Novel Self-Resistance Mechanism. *Tetrahedron* **56**, 693-701.

Zhang, Q., Gould, S. J. & Zabriskie, T. M. 1998 A new cytosine glycoside from *Streptomyces griseochromogenes* produced by the use in vivo of enzyme inhibitors. *J. Nat. Prod.* **61**, 648-51.