AN ABSTRACT OF THE DISSERTATION OF

Xiumei Wu for the degree of Doctor of Philosophy in Genetics presented on March 13, 2008.

Title: Functional Analysis of the Biosynthetic Gene Cluster of the Antitumor Agent Cetoniacytone A.

Abstract approved: _____

Taifo Mahmud

A gene cluster responsible for the biosynthesis of the antitumor agent cetoniacytone A was identified in *Actinomyces* sp. strain Lu 9419, an endosymbiotic bacteria isolated from the intestines of the rose chafer beetle. The nucleotide sequence analysis of the 26 Kb DNA region revealed the presence of 17 complete ORFs, including genes predicted to encode a 2-*epi*-5-*epi*-valiolone synthase (CetA), a glyoxalase (CetB), an FAD/FMN-dependent dehydrogenase (CetF), an oxidoreductase (CetG), two aminotransferases (CetH, CetM), and a pyranose oxidase (CetL). BLAST search analysis using the newly isolated *cet* biosynthetic pathway revealed a homologous biosynthetic pathway in the genome of *Frankia alni* ACN14a, suggesting that this organism is capable of producing a metabolite related to the cetoniacytones.

The 2-*epi*-5-*epi*-valiolone synthase (CetA) was cloned and expressed in *E. coli* and biochemical characterization of the gene product revealed that CetA is

capable of catalyzing the cyclization of sedoheptulose 7-phosphate to 2-*epi*-5-*epi*-valiolone. In addition, three other 2-*epi*-5-*epi*-valiolone synthase genes, from different natural product biosynthetic pathways have also been recombinantly expressed and biochemically characterized. These include *BE-orf9* from the BE-40644 biosynthetic gene cluster, *prlA* from the pyralomicin biosynthetic gene cluster, and *salQ* from the salbostatin biosynthetic gene cluster. Comparative analysis of the gene products with other related cyclases that are involved in natural product biosynthesis revealed that the 2-*epi*-5-*epi*-valiolone synthases uniquely represent a class of sugar phosphate cyclases (SPCs) that has a catalytic mechanism similar to that of dehydroquinate synthase in the shikimate pathway. Enzymes that belong to the SPC superfamily catalyze the cyclization of sugar phosphates to produce a variety of cyclitol intermediates that serve as the building blocks of many primary and secondary metabolites. Further phylogenetic analysis of SPC sequences revealed a new clade of SPCs, consisting a group of proteins with unknown function, that may regulate the biosynthesis of a novel set of secondary metabolites.

The product of *cetB*, which has high identity to glyoxalases (members of the vicinal oxygen chelate (VOC) superfamily), has also been characterized. Members of the VOC superfamily catalyze a large range of divalent metal ion-dependent reactions. Enzymatic characterization of CetB revealed that this enzyme was able to catalyze the second metabolic step in cetoniacytone biosynthesis, mediating the epimerization of 2-*epi*-5-*epi*-valiolone to 5-*epi*-valiolone. Therefore, CetB may be

designated as a new member of the VOC superfamily. Site directed nutagenesis and metal binding analysis showed that CetB is a Ni^{2+}-dependent protein with four metal binding sites. Similar to other members of the VOC superfamily, CetB contains the common structural βαβββ scaffold, as predicted through Phyre program. Native protein gel and size exclusion analyses have shown that CetB exists as a two-module protein dimer. The results provide important insight into the mode of formation of this unique aminocyclitol natural product, and will contribute to future studies that aim to create new aminocyclitol analogs.

Functional Analysis of the Biosynthetic Gene Cluster of the Antitumor Agent Cetoniacytone A

by

Xiumei Wu

A DISSERTATION

submitted to

Oregon State University

in partial fulfillment of

the requirements for the

degree of

Doctor of Philosophy

Presented March 13, 2008

Commencement June 2008

Doctor of Philosophy dissertation of Xiumei Wu presented on March 13, 2008.

APPROVED:

_____

Major Professor, representing Genetics

_____

Director of the Genetics Program

_____

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.

_____

Xiumei Wu, Author

ACKNOWLEDGEMENTS

I express my sincere and deep appreciation to my academic advisor, Dr. Taifo Mahmud, for his guidance and encouragement throughout my study at Oregon State University. His guidance is a great help in the progress of finishing this research and dissertation work.

I would like to thank Dr. Patricia M. Flatt for her relentless help and her hands-on training on constructing and screening cosmid library. I am grateful to Valerie Peterson for showing me the HPLC working system.

I am indebted to my committee members, Dr. Mark Zabriskie, Dr. Mahfuz Sarker, Dr. Xihou Yin, and Dr. Arup Indra, for their expertise, time, and encouragement in directing my thesis.

Thanks to Dr. Axel Zeeck for providing *Actinomyces* strain LU9419, Dr. Soon-Kwang Hong for providing *salQ* gene, and Dr. Tohru Dairi for providing *BE-orf9* gene. Thanks to my group members and members in Dr. Philip J. Proteau's lab and Dr. Kerry L. McPhail's lab for their help on my research work.

Thanks to National Institute of Health for financial support during my study here.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

Page

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

LIST OF FIGURE

LIST OF FIGURE (Continued)

Figure Page

LIST OF FIGURE (Continued)

LIST OF TABLES

LIST OF ABBREVIATIONS

AHBA	3-amino-5-hydroxybenzoic acid
aminoDHS	5-deoxy-5-amino-3-dehydroshikimic acid
aminoDHQ	5-deoxy-5-amino-3-dehydroquinic acid
BRP	bleomycin resistance protein
COG	clusters of orthologous groups
DAHP	3-deoxy-D-*arabino*-heptulosonate 7-phosphate
DHBD	dihydroxybiphenyl dioxygenase
DHMEQ	dehydroxymethyl epoxyquinomicins
DHQ	dehydroquinate
DHQS	DHQ synthase
DOI	2-deoxy-*scyllo*-inososе
DOIS	2-deoxy-*scyllo*-inososе synthase
E4P	erythrose 4-phosphate
EVS	2-*epi*-5-*epi*-valiolone synthase
FAD	flavin adenine dinucleotide
FOS	fosfomycin resistance protein
GLO	glyoxalase
HPLC	high performance liquid chromotography
IPTG	isopropyl-β-D-thiogalactopyranoside
MIPS	*myo*-inositol 1-phosphate synthase
MFS	major facilitator superfamily

LIST OF ABBREVIATIONS (Continued)

MMCE	methylmalonyl-CoA epimerase
MS	mass spectrometry
NAD	nicotinamide adenine dinucleotide
NAT	*N*-acetyltransferase
NF-κB	nuclear factor-κB
NMR	nuclear magnetic resonance
NSAID	non-steroidal anti-inflammatory drugs
ORF	open reading frame
PEP	phosphoenolpyruvate
PLP	pyridoxal 5′-phosphate
SPCs	sugar phosphate cyclases
TLC	thin layer chromatography
VAO	vanillyl-alcohol oxidase
VOC	vicinal oxygen chelate

DEDICATION

I dedicate this work to my husband, Chengwei Zhang, my parents, Qinfa Wu and Jiying Zhang, my sister, and my brother, for their constant encouragement and support throughout my study towards this Ph.D degree.

Functional Analysis of the Biosynthetic Gene Cluster of the Antitumor Agent Cetoniacytone A

Chapter One

Introduction

Biosynthetic studies of natural products have emerged as one of the more exciting areas of research and have become an important part of modern drug discovery and development efforts. An in-depth knowledge of the biosynthesis of natural products, including identification of genes and characterization of enzymes that are involved, has proved to be crucial for the success of improving the production and creating new analogs of natural products.[1-3] The advances of molecular genetics, protein science, biotechnology, and analytical instrumentation have accelerated the pace of natural products biosynthetic studies and provided significant contributions to this area of research during the last two decades. Using combinations of contemporary molecular genetic approaches, enzymology, and chemistry, coupled with powerful and sophisticated mass spectroscopy and NMR, it is now possible to dissect mechanisms and processes involved in natural product biosynthesis at the molecular level. In addition, dozens of structurally altered natural products have recently been generated,[4] ranging from analogs of the antibiotic erythromycin[5, 6] to the anthelmintic avermectins[7] and the anti-tumor agents, the epothilones.[8] In addition, genetic methods can be used to generate analogs that are difficult to make using traditional synthetic methodologies.

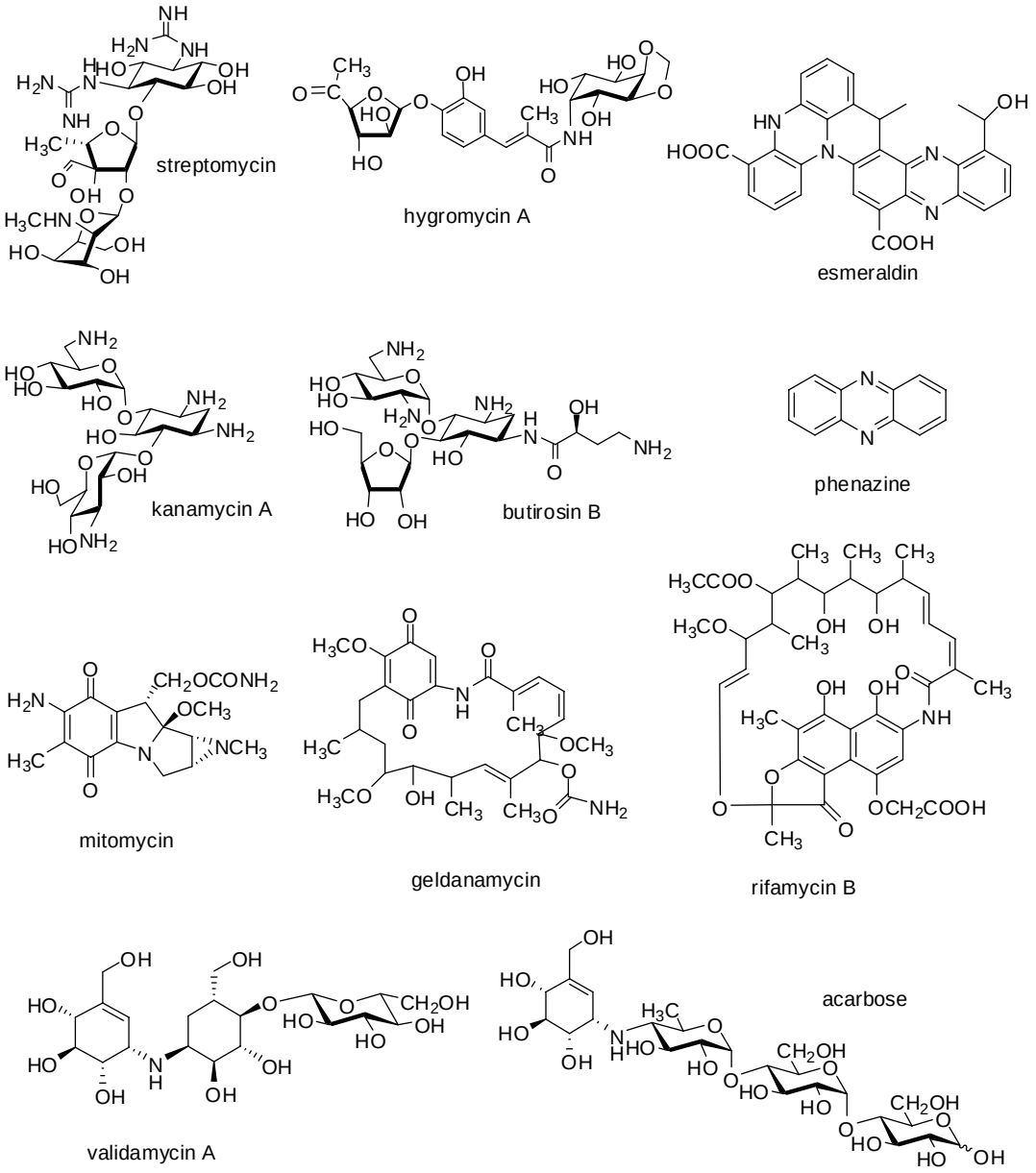


Figure 1-1. Chemical structures of diverse natural products derived from sugar phosphate intermediates.

Among the various classes of natural products are those derived from cyclitol intermediates made through sugar phosphate cyclases, which represent a

large family of enzymes involved in primary and secondary metabolism (Figure 1-1). They have formed a broad class of biologically active compounds and have come to occupy a prominent place in natural product research. These include the clinically important aminoglycoside antibiotics, *e.g.*, kanamycin, neomycin, and streptomycin, the ansamycin antibiotics, *e.g.*, rifamycin and geldanamycin, and the C_7N-aminocyclitol-containing compounds, *e.g.*, acarbose, validamycin A, and pyralomicin. Many of these compounds are used widely for the treatment of diseases in humans, animals, and plants. For example, streptomycin, kanamycin, neomycin, and rifamycin have long been used in the clinic against bacterial infections,[2] whereas acarbose is a potent α-glucosidase inhibitor used for the treatment of insulin-independent type-II diabetes.[3] On the other hand, validamycin A is an important crop protectant, particularly against sheath blight disease of rice plants.[4] Based on the type of sugar phosphate cyclases involved in their biosynthesis, they can be divided into five major groups:[1] (1) the *myo*-inositol-derived aminoglycoside antibiotics [*e.g.*, streptomycin and hygromycin], which require *myo*-inositol 1-phosphate synthase (a sugar phosphate cyclase); (2) the shikimate pathway-derived natural products [*e.g.*, phenazine and esmeraldine], which require dehydroquinate (DHQ) synthase; (3) the 2-deoxy-*scyllo*-inosose (DOI)-derived aminoglycosides [*e.g.*, kanamycin], which require DOI synthase; (4) the aminoshikimate-derived antibiotics [*e.g.*, rifamycin, geldanamycin, and mitomycin], which require aminodehydroquinate (aminoDHQ) synthase; and (5) the 2-*epi*-5-*epi*-valiolone-derived aminocyclitols [*e.g.*, acarbose and validamycin

A], which require 2-*epi*-5-*epi*-valiolone synthase for their biosynthesis.

1.1. The myo-*inositol 1-phosphate synthases*

The first group of secondary metabolites made through a sugar phosphate cyclase intermediate is the *myo*-inositol-derived aminoglycoside antibiotics. Members of this group include streptomycin, bluensomycin, spectinomycin and fortimicin, which contain aminocyclitol units originating from a *myo*-inositol intermediate. The biosynthesis of inositol has been known as an evolutionarily conserved pathway and its importance in the biological kingdom has been recognized for a long time. The cyclitol and its metabolites are not only involved in the biosynthesis of aminoglycoside antibiotics, they are also involved in growth regulation, membrane biogenesis, osmotolerance and many other biological functions including signal transduction pathways.[9, 10] Phylogenetic analysis of *myo*-inositol 1-phosphate synthase (MIPS) has shown that *myo*-inositol not only has a diverse set of functions but also is abundantly distributed throughout the biological kingdoms, *e.g.*, in higher plants and animals, parasites, fungi, green algae, eubacteria and archaea.[11]

All *myo*-inositol-producing organisms studied to date produce the cyclitol via dephosphorylation of *myo*-inositol 1-phosphate generated from glucose 6-phosphate by an oxidoreduction/aldol cyclization reaction catalyzed by MIPS.[10, 11] MIPSs serve as the first and the rate-limiting enzyme in the synthesis of all

inositol-containing compounds. A number of MIPS genes have been cloned from different organisms, and their catalytic function has been studied in detail. In addition, the crystal structure of yeast MIPS has been determined.[12] Based on both the NAD^+-bound enzyme and the inhibitor-bound enzyme, the mechanism of the conversion of glucose 6-phosphate to *myo*-inositol 1-phosphate has been proposed (Figure 1-2). It involves three distinct steps including an oxidation, an intramolecular aldol cyclization, and a reduction.[11, 13] NAD^+ and an ammonium cation are required as cofactors in the reaction (Figure 1-2). Although *myo*-inositol 1-phosphate is also a cyclic product of an SPC enzyme and is incorporated into a wide array of secondary metabolites, the MIPSs are thought to utilize a separate reaction mechanism and do not share significant homology with other classes of SPCs described below.

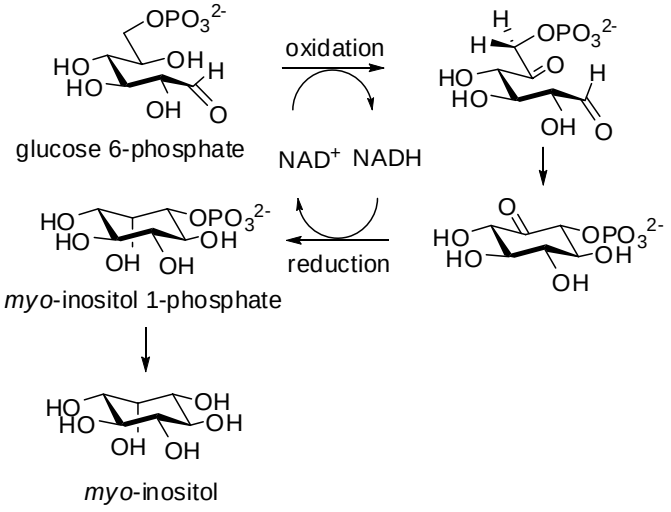


Figure 1-2. Proposed mechanism of *myo*-inositol 1-phosphate synthase.

1.2. Dehydroquinate synthases (DHQ synthases)

DHQ synthases are sugar phosphate cyclases that are involved in the shikimate pathway. The shikimate pathway is a common route leading to the production of aromatic amino acids, *e.g.*, phenylalanine, tyrosine, and tryptophan. Seven metabolic step are involved in the conversion of phosphoenolpyruvate (PEP) and erythrose 4-phosphate (E4P) to chorismic acid, the precursor of the three aromatic amino acids and a number of other aromatic compounds which are critical to sustaining the primary function of living organisms (Figure 1-3). A number of very good reviews have been published during the last three decades.[14-18]

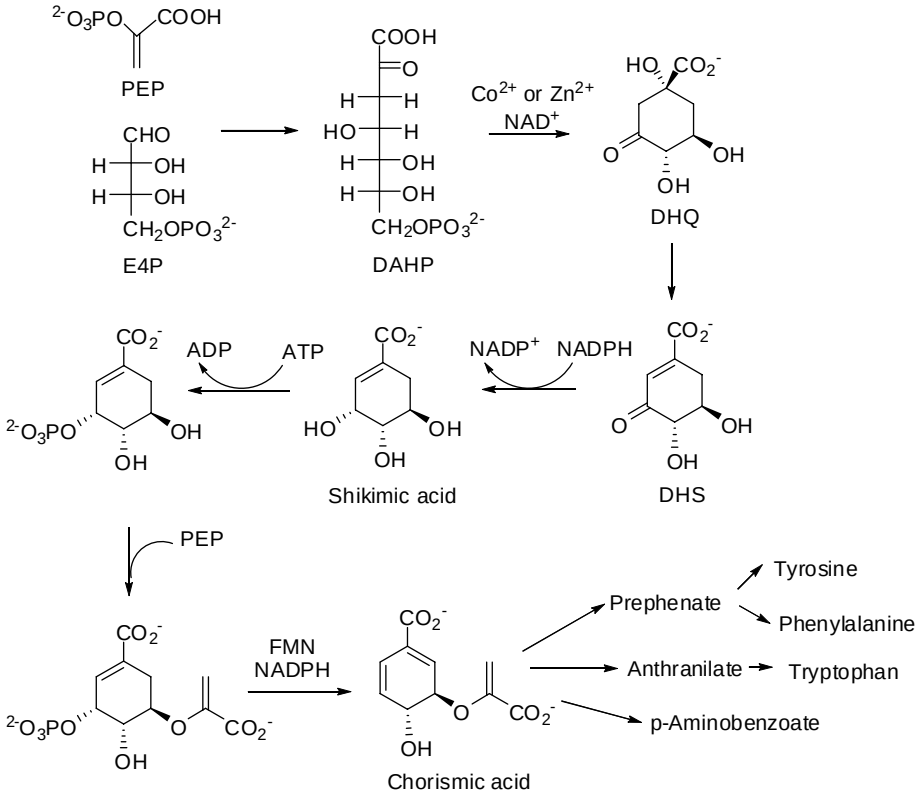


Figure 1-3. The shikimate pathway.

DHQ synthase catalyzes the second step in the shikimate pathway, converting DAHP to DHQ. In the fungus *Neurospora crassa*, the DHQ synthase reaction is the first of five sequential steps in the shikimate pathway that are catalyzed by a multifunctional super enzyme called AROM. The multifunctional AROM enzyme, comprising DHQ synthase and the four subsequent shikimate pathway activities, appears to be restricted to fungi. In bacteria and plants, each biosynthetic step during shikimate biosynthesis is mediated by a separate, independent enzyme. It appears that the *arom* locus evolved by gene fusion and convergent evolution.[19] DNA encoding the entire AROM complex has been cloned and sequenced from *Aspergillus nidulans*,[19-22] *Saccharomyces cerevisiae*,[23, 24] *N. crassa*,[25] and *P. carinii*.[26] Crystal structure analysis of the *A. nidulans* DHQ synthase showed that DHQ synthase adopts an open form in the absence of substrate[27] and that large-scale domain movement occurs during the normal catalytic cycle.[28-31] Compared to corresponding fungal enzymes, the DHQ synthases of bacteria are smaller and generally separable from other shikimate-pathway activities. Since the discovery of the AROM protein, a substantial number of DHQ synthases from different source including *E coli* K-12,[32, 33] *Corynebacterium glutamicum*,[34] *Neisseria gonorrhoeae*,[35] *Helicobacter pylori*,[36] and *Bacillus subtilis*[37] have been purified and characterized.

DHQ synthases require divalent cations for activity,[33, 38] showing the most robust activity with either Co^{2+} or Zn^{2+}. In addition, the enzyme also requires

catalytic amounts of NAD^+ for activity.[25] There is evidence for two functionally distinct metal binding sites per polypeptide chain.[39] Conversion of DAHP to DHQ involves an oxidation, a ß-elimination of inorganic phosphate, a reduction, a ring opening, and an intramolecular aldol condensation (Figure 1-4).[29, 40] The molecular mechanism of the overall reaction has been studied extensively by the use of substrate analogs. The crystal structure of DHQ synthase in five liganded states suggests that the enzyme is actively involved in all of the five steps of the reaction.[29, 30] Comparison of the available *A. nidulans* and *S. aureus* DHQ synthase structures suggested that despite the low overall amino acid sequence identity, there is a 100% identity between substrate-contacting residues.[28]

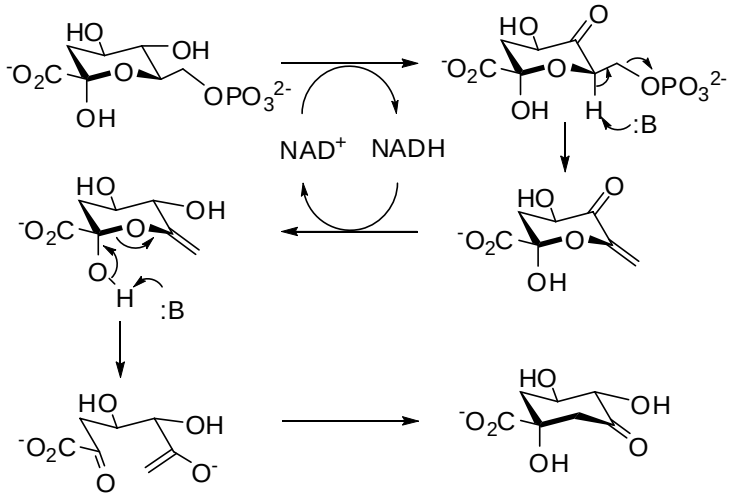


Figure 1-4. Proposed mechanism of DHQ synthase.

1.3. 2-Deoxy-scyllo-*inososе synthases*

One of the most-well-studied families of secondary metabolites made through sugar phosphate cyclase is the 2-deoxy-*scyllo*-inosose-derived

aminoglycosides. The 2-deoxystreptamine aglycon is a common structural feature found in this class of aminoglycoside antibiotics including neomycin, kanamycin,[41] tobramycin,[42] gentamicin, sisomicin, butirosin and ribostamycin.[43] This moiety plays a critical role in the biological function of this class of aminoglycoside antibiotics and has been shown to interact directly with the 16S ribosomal RNA subunit.[44-46] The biosynthesis of 2-deoxystreptamine commences with the initial carbocycle formation step from D-glucose 6-phosphate to 2-deoxy-*scyllo*-inosose (DOI). This crucial step is known to be catalyzed by DOI synthase, an enzyme that shares high homology with DHQ synthase. Three additional steps including oxidation and two transaminations[47, 48] are required to complete the production of 2-deoxystreptamine (Figure 1-5).

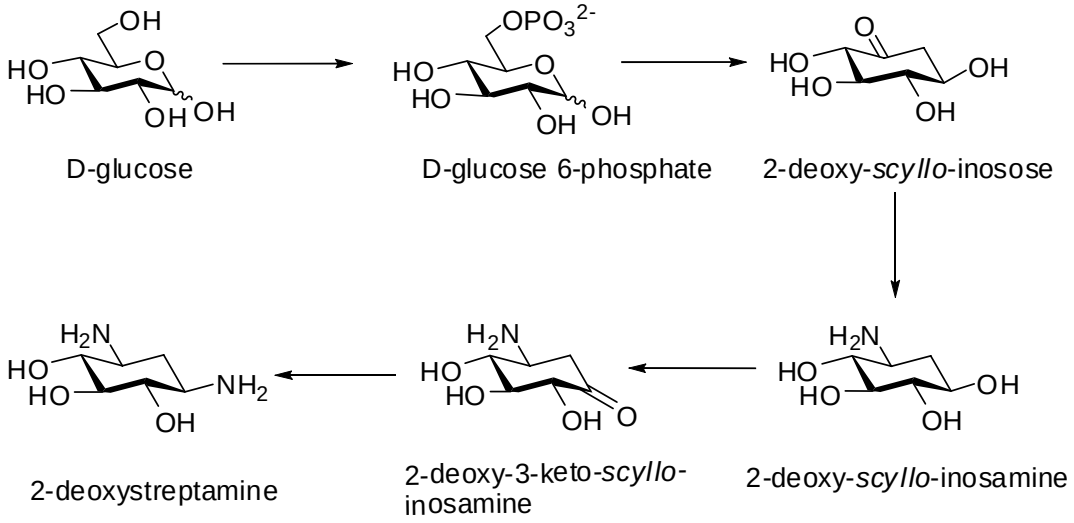


Figure 1-5. The formation of 2-deoxystreptamine.

Kakinuma and coworkers reported the first isolation and functional

characterization of DOI synthase from the butirosin-producing bacteria, *Bacillus circulans,* by cloning and overexpression the gene, *btrC*.[49-51] A gene (*btrC2*) encoding a 20-kDa subunit of DOI synthase was also identified from butirosin-producer *B. circulans*. Although heterologous expression of BtrC in the absence of BtrC2 gave full enzymatic activity, the simultaneously over-expressed BtrC2 and BtrC showed that BtrC2 and BtrC spontaneously form a heterodimer structure, and that the heterodimer was much more stable than BtrC homodimer. So, the major role of BtrC2 appears to be involved in the stabilization and regulation of BtrC.[52] In addition to BtrC, a number of DOI synthases from different aminoglycoside biosynthetic gene clusters, including the ribostamycin,[53] neomycin,[54] kanamycin,[55-57] tobramycin,[42, 58] and gentamicin,[59] have been isolated.

The mechanism of the DOI synthase reaction was investigated by X-ray crystallography analysis with a potent irreversible inhibitor.[60] A comparison of the active site residues of the DHQ synthase from *A. nidulans* with DOI synthases showed several key differences within the active site, suggesting that these residues may be involved in substrate binding.[61] However the reaction mechanism of DOI synthase is predicted to be very similar to that of DHQ synthase. This single enzyme is able to catalyze the full carbocycle-forming process from glucose 6-phosphate to DOI in the presence of NAD^+ and Co^{2+} as cofactors. The multistep process includes, (i) oxidation at C-4 and elimination of phosphate, (ii) stereoselective reduction at C-4, (iii) deprotonation and ring-opening of the hemi-

acetal to an enolate-aldehyde, and, finally, (iv) stereoselective intramolecular aldol condensation between C-1 and C-6 (Figure 1-6). The presence of Co^{2+} was essential for enzyme activity, and Zn^{2+} was totally inhibitory.[49] While the reaction mechanisms are quite similar, the DOI synthases appear to be distinct from the DHQ synthases with respect to their quaternary structures, metal cation requirements, and kinetic parameters.[49] In the DHQ synthase reaction, the phosphate elimination takes place in *syn* fashion by intramolecular self-catalysis. On the other hand, anti-elimination of the phosphate is postulated for the DOI synthase reaction.[50]

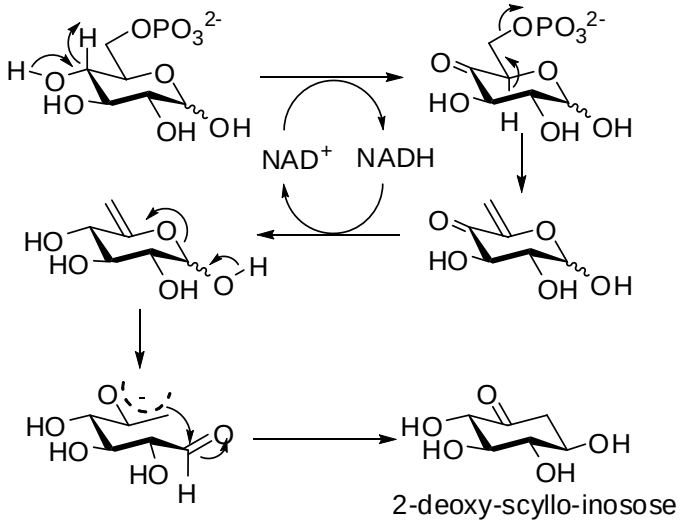


Figure 1-6. Proposed mechanism of 2-deoxy-*scyllo*-inosose synthase.

1.4. AminoDHQ synthases

The third family of secondary metabolites derived from the cyclization of sugar phosphate is the large family of ansamycin antibiotics. They all contain a

common C_7N core unit, which is derived from 3-amino-5-hydroxybenzoic acid (AHBA). Members of this family include rifamycin B,[62-64] geldanamycin,[65] ansamitocin, and mitomycin.[66] In the case of rifamycin B, geldanamycin, and ansamitocin, the AHBA represents the starter unit from which the ansa chain is assembled by a polyketide-type biosynthesis. Inactivation of the AHBA biosynthetic genes in *Amycolatopsis mediterranei* and *Streptomyces hygroscopicus* genome resulted in the loss of rifamycin [67] and geldanamycin[68, 69] formation, respectively.

The biosynthesis of AHBA has been studied at the chemical, biochemical, and genetic levels in organisms producing various ansamycins,[70] and in the heterologous host *E. coli*.[71, 72] Floss and coworkers have probed the biosynthetic pathway to AHBA through feeding experiments with stable isotopes, cell-free assays,[67, 73] and mutational analyses.[74] Their results confirmed that the formation of AHBA occurs by a novel parallel route to the normal shikimate pathway. The operation of the new variant of the shikimate pathway in the formation of AHBA starts from phosphoenolpyruvate (PEP) and imino-erythrose 4-phosphate (imino-E4P) to 3,4-dideoxy-4-amino-D-arabino-heptulosonate-7-phosphate (aminoDAHP). The pathway then continues via 5-deoxy-5-amino-3-dehydroquinic acid (aminoDHQ) and 5-deoxy-5-amino-3-dehydroshikimic acid (aminoDHS) to AHBA (Figure 1-7).[75-77] The key intermediate, aminoDAHP, was first proposed by Hornemann *et al*.[78] Heterologous expression of the biosynthetic gene cluster of

AHBA from the rifamycin producer and the geldanamycin producer further confirmed the proposed AHBA biosynthetic pathway.[71] In AHBA biosynthesis, an aminoDHQ synthase catalyzes the cyclization of aminoDAHP to aminoDHQ (RifG in rifamycin pathway,[79] GdmO in geldanamycin pathway,[80] Asm47 in ansamitocin pathway,[81] and MitP in mitomycin[66] pathway). These enzymes show high sequence similarity with DHQ synthase in the shikimate pathway. AHBA synthase, the terminal enzyme in the formation of AHBA was studied in detail and shows the dependency on pyridoxal 5′-phosphate (PLP).[82, 83] Mechanistic studies showed that the enzyme-bound pyridoxal 5′-phosphate forms a Schiff's base with the amino group of aminoDHS and catalyzes both an α,β-dehydration and a stereospecific 1,4-enolization of the substrate.

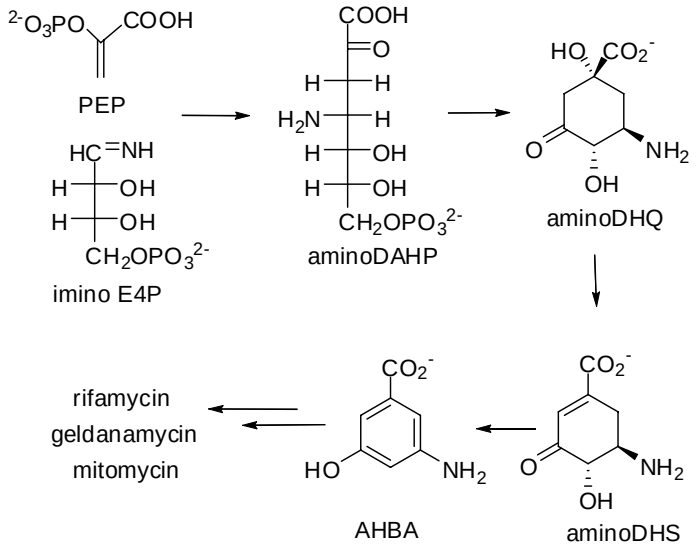


Figure 1-7. Proposed pathway for the formation of AHBA.

1.5. 2-epi-*5*-epi-*Valiolone synthases*

The last group of secondary metabolites made through sugar-phosphate-cyclase is the 2-*epi*-5-*epi*-valiolone derived natural products (Figure 1-8). Many of these compounds are used widely to combat diseases in humans, animals, and plants. For example, acarbose, which is produced by *Actinoplanes* strain SE 50/110, is a potent α-glucosidase inhibitor used for the treatment of insulin-independent type-II diabetes.[84] Validamycin A, which was isolated from the fermentation broth of *Streptomyces hygroscopicus* var. *limoneus* and *Streptomyces hygroscopicus* subsp. *jinggangensis* 5008, is a trehalase inhibitor widely used in Asia as a crop protectant for sheath blight disease.[85, 86] Other members of this family include the anti-bacterial agent pyralomicin, the antitumor agent cetoniacytone A, the trehalase inhibitor salbostatin, and the human thioredoxin system inhibitor BE-40644.[87] Epoxyquinomicins, which were isolated from the culture broth of *Amycolatopsis* sp. strain MK 299-95 F4, share a similar C_7N-aminocyclitol moiety with cetoniacytone A, although the biosynthetic origin of its core C_7N moiety has not been investigated.[88-90] Since the central part of this Ph.D thesis is most related with the biosynthesis of C_7N-aminocyclitols, this family of natural products will be discussed in more detail.

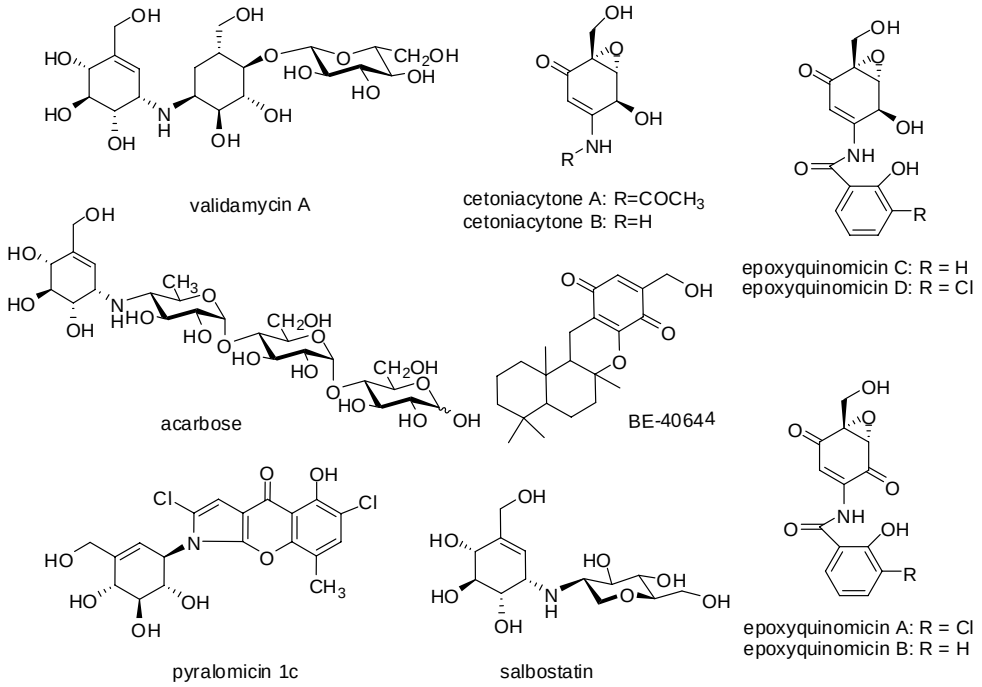


Figure 1-8. Chemical structures of 2-*epi*-5-*epi*-valiolone-derived C_7-cyclitols.

1.5.1. Biosynthetic origin of the C_7N-aminocyclitol core unit

The carbon skeleton of the aminocyclitols in acarbose and validamycin has been established by the groups of Rinehart[91] and Floss.[92] Feeding experiments with ^{13}C-labeled glucose and [U-$^{13}C_3$]glycerol demonstrated that the C_7N moiety is derived from a 7-carbon sugar phosphate, sedoheptulose 7-phosphate, which is formed by C_2+C_2+C_3 related to the pentose phosphate pathway. Later, Floss and co-workers synthesized a number of ^{2}H-, ^{3}H-, or ^{13}C-labeled potential precursor cyclitols and fed them to the acarbose and validamycin producer.[93-95] The feeding experiments showed that while valiolone, which has the same stereochemistry as the aminocyclitol core unit of acarbose, failed to act as a precursor of acarbose, its

epimer, 2-*epi*-5-*epi*-valiolone, was efficiently incorporated into the compound. It suggested that 2-*epi*-5-*epi*-valiolone is the initial cyclitol precursor of this core moiety. Additional feeding experiments showed that 2-*epi*-5-*epi*-valiolone was also efficiently incorporated into validamycin and labels both cyclitol moieties, suggesting that 2-*epi*-5-*epi*-valiolone is the precursor of all the aminocyclitol moieties in this compound. Similar results have also been shown in the biosynthetic studies of pyralomicin and cetoniacytone A, in which 2-*epi*-5-*epi*-valiolone was incorporated into both compounds.[96-98]

The use of 2-*epi*-5-*epi*-valiolone in the biosynthesis of acarbose and validamycin has been confirmed by genetic and biochemical approaches.[99-101] The putative biosynthetic gene cluster of acarbose was identified in the producer strain *Actinoplanes* sp. SE50/110 by cloning a DNA fragment containing the conserved sequences for dTDP-glucose 4,6-dehydratases (*acbB*) and using this fragment as a hybridization probe. The two flanking genes were *acbA*, which showed homology to dTDP-glucose synthases, and *acbC*, which encodes a protein homologous to DHQ synthases, amino-DHQ synthases, and DOI synthases. The *acbC* gene was overexpressed heterologously in *Streptomyces lividans* and the product was shown to be a C_7-cyclitol synthase using sedoheptulose 7-phosphate, but not ido-heptulose 7-phosphate, as its substrate.[99] The cyclization product was identified as 2-*epi*-5-*epi*-valiolone. Based on the analysis of the three dimensional structure of an eukaryotic DHQ synthase domain, a five-step reaction mechanism of AcbC,

including dehydrogenation at C-5 by NAD^+, elimination of phosphate group, reduction of the C-5 ketone, ring opening and rotation along the C-5-C-6 bond, and intramolecular aldo condensation, was proposed (Figure 1-9).[99]

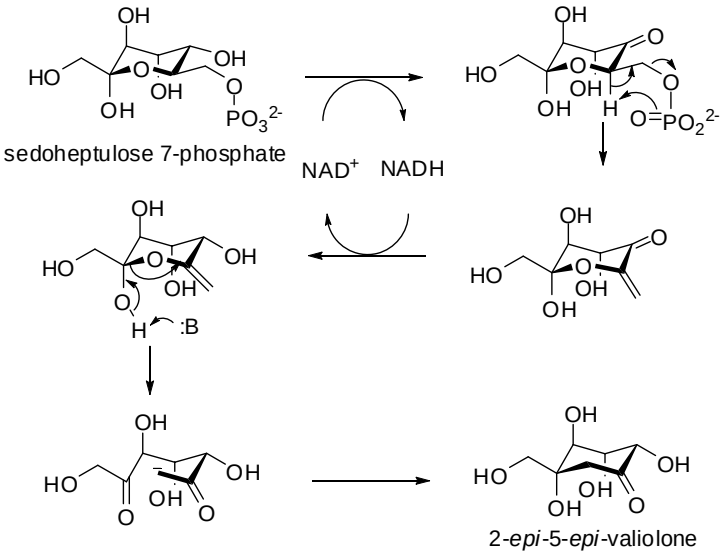


Figure 1-9. Proposed mechanism of 2-*epi*-5-*epi*-valiolone synthase.

1.5.2. Biosynthesis of acarbose

The α-glucosidase inhibitor acarbose is produced by *Actinoplanes* sp. SE50/110 and inhibits various α-glucosidases including sucrase, maltase, dextrinase, and glucoamylase. It is built-up from three distinct moieties, an unsaturated cyclitol (valienamine), a deoxyhexose and a maltose unit, linked by (pseudo-)α-1,4 linkages. The valienamine and the deoxyhexose are linked *via* an N-glycosidic bond. This acarviosyl moiety, the combination of valienamine and the deoxyhexose, is primarily responsible for the inhibitory effect on α-glucosidases.

The biosynthesis of acarbose has been studied by a combination of conventional feeding experiments using isotopically labeled compounds and by genetic and biochemical approaches. Feeding experiments with *Actinoplanes* using various ^{15}N-labeled compounds have been done and showed that glutamate is the most efficient nitrogen source among the substrates tested and that the mode of nitrogen introduction into the acarbose must be a pyridoxal phosphate-dependent transamination.[102] Feeding experiments using ^{2}H- or ^{3}H-labeled maltose or maltotriose showed that the maltose unit of acarbose arises from maltotriose which is present in abundance in the fermentation medium.[103] Feeding experiments with [U-$^{13}C_3$]glycerol demonstrated that the aminocyclitol moiety of acarbose is derived from the pentose phosphate pathway and that 2-*epi*-5-*epi*-valiolone is a key intermediate in acarbose biosynthesis.[92, 104]

With the identification of the acarbose biosynthetic gene cluster, the formation of acarbose becomes more evident. Biosynthesis of acarbose is proposed to begin with the cyclization of sedoheptulose 7-phosphate to 2-*epi*-5-*epi*-valiolone by the 2-*epi*-5-*epi*-valiolone synthase, AcbC. AcbM has been verified as a 2-*epi*-5-*epi*-valiolone 7-kinase that converts 2-*epi*-5-*epi*-valiolone to its 7-phosphate derivative.[105] The 2-*epi*-5-*epi*-valiolone 7-phosphate was further converted by the AcbO epimerase to 5-*epi*-valiolone 7-phosphate.[105, 106] The NADH-dependent dehydrogenase AcbL has been proposed to mediate the reduction of the C-1 keto group to form 5-*epi*-valiolol 7-phosphate. The remaining steps in acarbose

biosynthesis have been predicted from BLAST search analysis and include a dehydration reaction by AcbN to synthesize 1-*epi*-valienol 7-phosphate. A second phosphorylation by an unidentified kinase and a subsequent nucleotidylation by the NDP-transferase, AcbR, are predicted to yield NDP-1-*epi*-valienol 7-phosphate. The AcbA (homology to dTDP-glucose synthase), AcbB (homology to dTDP-glucose 4,6-dehydratases), and AcbV (homology to aminotransferases) proteins are predicted to catalyze the conversion of glucose 6-phosphate to 4-amino-4,6-dideoxy-dTDP-glucose, which is then coupled with NDP-1-*epi*-valienol 7-phosphate by either AcbS or AcbI to give dTDP-acarviose 7-phosphate. Further extension with two units of glucose yields acarbose (Figure 1-10).

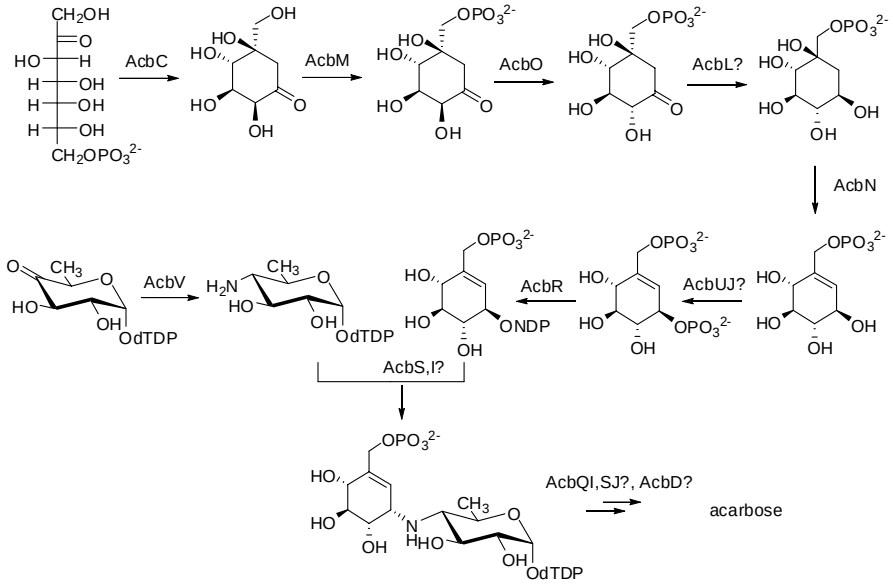


Figure 1-10. Proposed biosynthetic pathway to acarbose.

1.5.3. Biosynthesis of validamycin

The validamycins were found to be effective for controlling sheath blight disease of rice plants and preventing damping-off of cucumber seeding caused by the fungus *Rhizoctonia solani*. As mentioned above, the valienamine moieties in acarbose and validamycin both utilize 2-*epi*-5-*epi*-valiolone as a precursor during biosynthesis. However, the downstream steps required for valienamine moiety formation are substantially different.[94] In the case of acarbose biosynthesis, 2-*epi*-5-*epi*-valiolone is phosphorylated directly and further reactions in the acarbose pathway involve phosphorylated intermediates. In the case of validamycin, several nonphosphorylated intermediates, including 5-*epi*-valiolone, valienone, and validone, were incorporated into validamycin, suggested that the phosphorylation step happens later than that of acarbose pathway.[95] Validamycin A, the most abundant validamycin metabolite, contains validoxylamine A as a core moiety, which is assembled from two C_7N-aminocyclitols, the unsaturated residue, valienamine, and the saturated one, validamine, which are connected through a single nitrogen bridge shared by both units. Feeding experiments with isotopically labeled precursors to the validamycin producer *S. hygroscopicus* var. *limoneus* revealed that 2-*epi*-5-*epi*-valiolone is specifically incorporated into validamycin A and labels both cyclitol moieties. A more proximate precursor of validamycin A is valienone, which is also incorporated into both cyclitol moieties. The first step involves epimerization at C-2 of 2-*epi*-5-*epi*-valiolone to give 5-*epi*-valiolone, which is also efficiently incorporated into validamycin A in isotopic feeding experiments. Epimerization is followed by dehydration between C-5 and C-6, to

yield valienone which is the specific precursor of both the unsaturated and saturated cyclitol moieties of validamycin A. Reduction of valienone forms validone, which is also incorporated specifically into validamycin A, but labels only the reduced cyclitol moiety(Figure 1-11).[95]

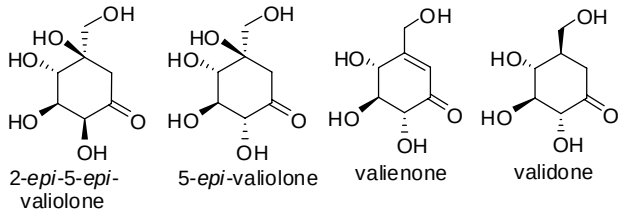


Figure 1-11. Precursors used for feeding experiments with validamycin A producer.

The biosynthetic gene cluster of validamycin has been identified in *S. hygroscopicus* subsp. *jinggangensis* 5008 by a collaboration between the Deng and Mahmud laboratories.[100] A 45 Kb DNA sequence revealed 16 structural genes, 2 regulatory genes, 5 genes related to transport, transposition/integration or resistance, and 4 genes with no obvious identity.[101] Deletion of a 30 Kb DNA fragment from this cluster in the chromosome resulted in the loss of validamycin production, confirming a direct involvement of the gene cluster in the biosynthesis of validamycin.[100] Within the 16 structural genes, only seven genes, *valABCKLMN*, from the two original operons are required for the synthesis of validoxylamine A in the heterologous host *S. lividans*. The predicted function of these seven genes include *valA* as the 2-*epi*-5-*epi*-valiolone synthase, *valB* as a nucleotidyltransferase, *valC* as a cyclitol kinase, *valK* as an epimerase/dehydratase, *valL* as a validoxylamine A 7-phosphate synthase, *valM* as an aminotransferase, and *valN* as

a cyclitol reductase. With one more gene, *valG* (encoding a glycosyltransferase), the formation of validamycin has been observed in *S. lividans*.[101] *In vivo* inactivation of the *valG* gene abolished the final attachment of glucose for validamycin production, and resulted in the accumulation of the precursor, validoxylamine A, while the normal production of validamycin A was restored by the complementation with *valG*. *In vitro* enzymatic assays using the recombinant ValG protein also demonstrated the glycosylation of validoxylamine A to validamycin A.[101] Further genetic and biochemical studies revealed that a number of other genes, *e.g.*, ValD and ValO, are also essential for the optimal production of validamycin A (Figure 1-12).

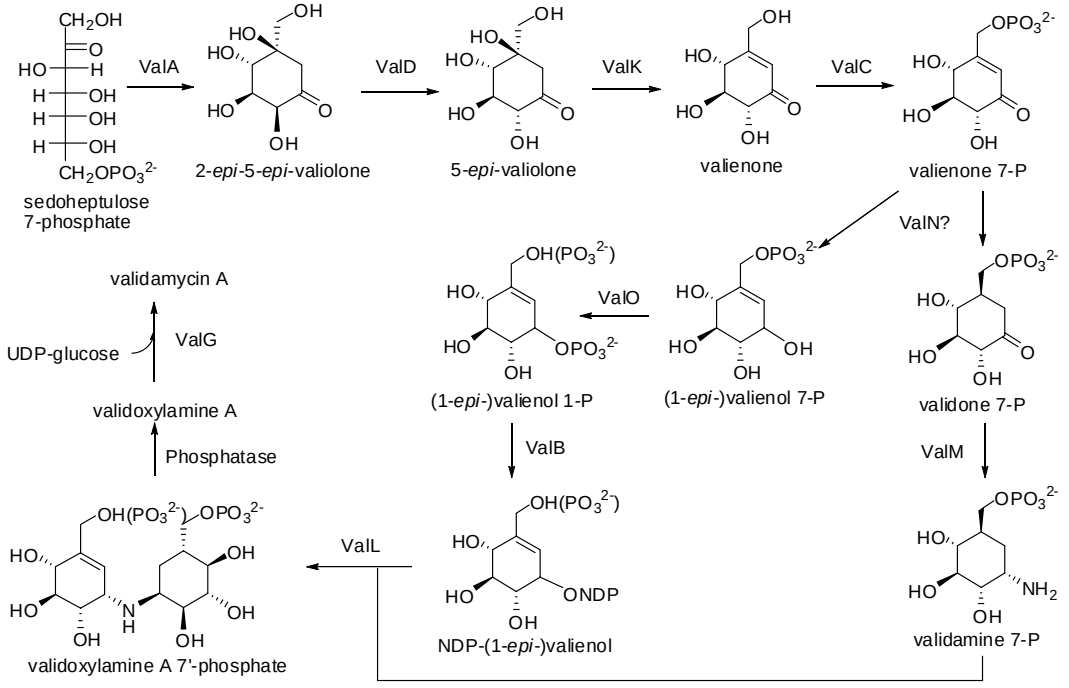


Figure 1-12. Proposed biosynthetic pathway to validamycin A.

Recently, two other validamycin gene clusters were identified from *S. hygroscopicus* 10-22[107] and *S. hygroscopicus* var. *limoneus*.[108] Three open reading frames, *orf1*-*orf3*, in *S. hygroscopicus* 10-22 genome were identified and showed 99% identity with *valA*, *valB*, and *valC*. Direct comparison of the validamycin cluster from *S. hygroscopicus* subsp. *jinganggensis* 5008 (the *val* cluster) and that from *S. hygroscopicus* var. *limoneus* KCCM 11405 (the *vld* cluster) has shown that both clusters contain an almost identical set of genes necessary for the biosynthesis of validamycin A, or at least for that of validoxylamine A.

1.5.4. Biosynthesis of pyralomicin

The pyralomicins are a set of antibiotics isolated from the soil bacterium *Nonomuraea spiralis* by Takeuchi and coworkers in 1995.[109-111] Pyralomicin 1a has a unique structure, a benzopyranopyrrole that is *N*-glycosylated by a C_7-cyclitol.[109] Studies on the origin of the benzopyranopyrrole core unit in pyralomicin 1a by feeding experiments have revealed that it is derived from two units of acetate, one unit of propionate and one molecule of proline.[111] The mechanism required for the formation of the unique benzopyranopyrrole core structure during pyralomicin biosynthesis is unknown, but is suggested to occur in a manner similar to the related compound, pyoluteorin.[112] On the other hand, biosynthesis of the cyclitol moiety is thought to occur in a similar manner to that of acarbose. Initial feeding experiments with D-[U-$^{13}C_6$]glucose showed incorporation into pyralomicin 1a with labeling and coupling patterns similar to those in acarbose and validamycin,

suggesting that it is derived from the pentose phosphate pathway. Further analysis of the cyclitol moiety of pyralomicin 1a has been studied using a series of ^{2}H-labeled potential precursors. The results demonstrate that 2-*epi*-5-*epi*-valiolone, a common precursor for acarbose and validamycin A biosynthesis, is also an immediate precursor of pyralomicin 1a. 5-*epi*-Valiolone was also incorporated into pyralomicin 1a, although much less efficiently than 2-*epi*-5-*epi*-valiolone. Other potential intermediates, such as valiolone, valienone, valienol, 1-*epi*-valienol, 5-*epi*-valiolol, and 1-*epi*-5-*epi*-valiolol were not incorporated into pyralomicin 1a. It was suggested that either 2-*epi*-5-*epi*-valiolone is specifically activated (*e.g.*, to its phosphate) or the transformation of 2-*epi*-5-*epi*-valiolone involves a substrate-channeling mechanism in which enzyme-bound intermediates are directly transferred from one enzyme active site to the next in a multi-enzyme complex.[97]

1.5.5. Biosynthesis of BE-40644

BE-40644 is a new human thioredoxin system inhibitor produced by *Actinoplanes* sp. A40644.[113] The structure of the novel metabolite was established by NMR analysis and was revealed as an isoprenoid-cyclitol-hybrid compound. Feeding experiments with [1-^{13}C]acetate, [1,2-$^{13}C_2$]acetate, and [1-^{13}C]glucose showed that the isoprenoid portion of the compound is derived from the mevalonate pathway, whereas the cyclitol unit is formed *via* the pentose phosphate pathway. It was suggested that the cyclization of the sedoheptulose 7-phosphate proceeds in a similar manner as reported for the biosynthesis of the cyclitol unit in

acarbose and validamycin.[114] The biosynthetic gene cluster of BE-40644 has been identified in *Actinoplanes* sp. A40644 by Dairi and co-worker.[87] A 22 Kb DNA sequence revealed a set of genes related to the mevalonate pathway that is responsible for the biosynthesis of isoprene units (*orfs 2-7*). Additional genes responsible for the formation of the cyclitol unit (*orfs 9-13*), genes encoding polyprenyl diphosphate synthase (*orf1*), prenyltransferase (*orf14*), and several hypothetical and regulatory proteins (*orfs 15-18*) were also revealed. In this study, *orf9*, a gene sharing high homology with 2-*epi*-5-*epi*-valiolone synthases has been identified. The lack of a cyclitol kinase in the gene cluster suggests that the cyclitol moiety may be transferred to the isoprenoid portion without an intermediate phosphorylation step (Figure 1-13).

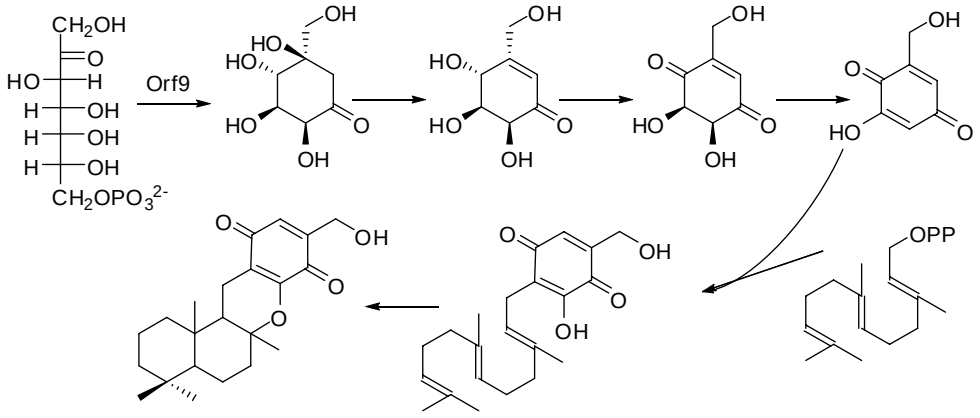


Figure 1-13. Proposed biosynthetic pathway to BE-40644.

1.5.6. Biosynthesis of salbostatin

Salbostatin is a basic, non-reducing pseudodisaccharide that consists of valienamine moiety linked to 2-amino-1,5-dihydro-2-deoxy-D-glucitol by an imino

bridge.[115] It was isolated from the fermentation culture of the polyether antibiotic salinomycin producer *Streptomyces albus* ATCC 21838. Despite having a similar chemical structure as acarviosine and other pseudodisaccharides, salbostatin does not inhibit the enzymes saccharase (EC 3.2.1.26) or maltase (EC 3.2.1.20) from the mucus of porcine small intestine, and only weakly (IC50 = 0.68 mM) inhibits murine liver aldose reductase (EC 1.1.1.21). Salbostatin also shows no antibiotic activity; however, it does effectively inhibit trehalase from porcine kidney (EC 3.2.1.28) with a Ki of 1.8×10^{-7} M. By using degenerate primers designed from alignment of *acbC*, *valA* and *aroB*, a biosynthetic gene cluster of salbostatin was isolated recently by Hong and co-worker in Myongji University.[116] Twenty one ORFs including an AcbC homolog, SalQ, was proposed to mediate the biosynthesis of salbostatin (Figure 1-14).

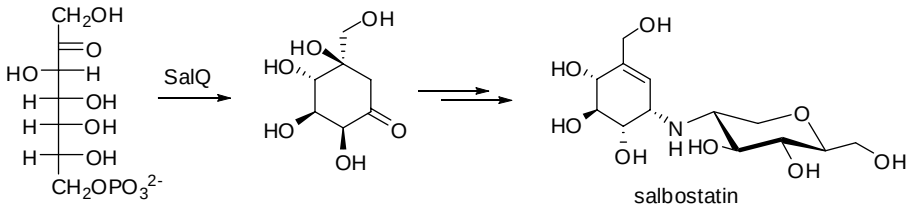


Figure 1-14. Proposed biosynthetic pathway to salbostatin.

1.5.7. Biosynthesis of cetoniacytone and epoxyquinomicins

Cetoniacytone A is a moderate anti-tumor agent produced by an endosymbiotic *Actinomyces* sp. strain Lu9419, which was isolated from the intestines of the rose chafer beetle (*Cetonia aureata*). Biological activity assays of cetoniacytone A showed no antimicrobial activity against Gram-positive and Gram-

negative bacteria, but it showed a moderate growth inhibition against the human tumor cell lines, HEPG2 (GI_{50} = 3.2μM) and MCF7 (GI_{50} = 4.4μM).[98] The structure of the novel metabolite was established by detailed spectroscopic analysis. Although cetoniacytone A belongs to the C_7N-aminocyclitol family, in contrast to most of the secondary metabolites belonging to this family, it contains an unusual C_7N-aminocyclitol structure, in which the amino group is acetylated and located at the C-2 position instead of the C-1 position (Fig 1-8).

Feeding experiments with *Actinomyces* sp. strain Lu9419 using [U-^{13}C]glycerol showed that the labeling and coupling patterns are similar to those of the valienamine moiety of acarbose and validamycin A, suggesting that the core unit of cetoniacytone is derived via the pentose phosphate pathway.[98] Further experiments with isotopically labeled 2-*epi*-5-*epi*-valiolone confirmed the involvement of this common precursor in the biosynthesis of cetoniacytone A.[96] The uniqueness of the cetoniacytone core structure has encouraged us to study their formation at the molecular level. Similar core structures have also been found in the epoxyquinomicins which were isolated from the culture broth of *Amycolatopsis* sp. strain MK299-95F4. These compounds have been found to be potential drug candidates for rheumatoid arthritis.[88-90, 117, 118] Structure determination showed that epoxyquinomicins C and D have the same core unit as cetoniacytone A, whereas epoxyquinomicins B and C are dechloro analogs of epoxyquinomicins A and D, respectively. In addition, epoxyquinomicins A and B possess a keto group at the

C-1 position of their core units instead of the hydroxyl group seen in the epoxyquinomicins C and D (Figure 1-8). The mode of action of epoxyquinomicins appears to be different from that of common non-steroidal anti-inflammatory drugs (NSAIDs), since epoxyquinomicins C and D did not inhibit cyclooxygenase-I at the concentration of 300 uM. Subsequent studies on their biological activities have revealed that these compounds inhibit histidine decarboxylase which is considered to be involved in the mechanism of inflammation by producing histamine.[119] On the other hand, the dehydroxymethyl derivatives of epoxyquinomicins C (DHMEQ) have been synthesized and several studies on the biological activities of DHMEQ showed that they are novel inhibitors of NF-κB which is a transcription factor that induces inflammatory cytokines and anti-apoptotic proteins.[120-127]

As part of our long term goal to generate analogs of bioactive aminocyclitol-containing natural products, we have investigated the biosynthesis of cetoniacytone A at the molecular level. This dissertation describes a functional analysis of the biosynthetic gene cluster of cetoniacytone A, which include the identification of the gene cluster (Chapter 3) and characterization of a 2-*epi*-5-*epi*-valiolone synthase (CetA) from the cetoniacytone pathway and its functional relationship within the sugar phosphate cyclase superfamily (Chapter 4). In addition, we have identified a new member of the vicinal oxygen chelate superfamily, CetB, as a 2-*epi*-5-*epi*-valiolone epimerase (Chapter 5).

Chapter 2

Materials and Methods

2.1. Bacterial strains and culture conditions

The bacterial strains and plasmids used in this study are listed in Table 2-1. *Actinomyces* sp. strain Lu 9419 was kindly provided by Dr. Zeeck's group at the University of Gottingen. This strain was maintained on YMG agar and cultured in YMG liquid medium consisting of malt extract (1%), yeast extract (0.4%), and glucose (0.4%), pH = 7.3 at 30 ^{o}C. Oatmeal medium consisting 2% oatmeal and 2.5 ml/l trace element solution [contains $CaCl_2.2H_2O$ (3 g/L), Fe(III)-citrate (1 g/L), $MnSO_4$ (0.2 g/L), $ZnCl_2$ (0.1 g/L), $CuSO_4.5H_2O$ (25 mg/L), $Na_2B_4O_7.10H_2O$ (20 mg/L), $CoCl_2$ (4 mg/L), $Na_2MoO_4.2H_2O$ (10 mg/L)], and sodium acetate (1g/L) as a supplement was used as the production medium for cetoniacytone A. *E. coli* strains XL-1-Blue and DH-10B (Stratagene) were used as hosts for library construction and routine subcloning, respectively. Over-expression of recombinant His-tagged proteins was carried out in *E. coli* BL21Gold(DE3)pLysS (Stratagene). *E. coli* VCS257 was used as a host for the preparation of pOJ446 cosmid library. *Escherichia coli* strains were cultured in Luria-Bertani Miller (LB) or LB containing betaine (2.5 mM) and sorbitol (1 M) (LBBS)[67] medium supplemented with appropriate amount of antibiotics whenever necessary. *S. lividans* T7, TK-24 and 1326 were used as hosts for the heterologous expression of the cetoniacytone A

biosynthetic gene cluster. *S. lividans* were routinely maintained on SFM agar at 30 ^{o}C and grown in YEME medium for protoplast preparation.[128] The production medium for *S. lividans* harboring the cetoniacytone A biosynthetic gene cluster is the same oatmeal medium as used for *Actinomyces* sp. strain Lu 9419.

Table 2-1. Bacteria strains and plasmids used in this study.

Strain or plasmid	Relevant characteristics	Source or Ref.
Actinomyces sp. strain Lu9419	Cetoniacytone producer	Dr. Zeeck's group
S. lividans 1326	SLP2+ SLP3+	Hopwood *et al.,* 1985
S. lividans T7	*S. lividans* 1326 with T7 promoter	Hopwood *et al.* 1985
S. lividans TK24	*Str-6* SLP2- SLP3-	Hopwood *et al.* 1985
E. coli DH10B	F-, *lacZ, recA1*	Stratagene
E. coli BL21Gold(DE3)pLys	Protein expression strain	Stratagene
E. coli VCS257		Stratagene
ET12567/pUZ8002	*Dam dcm hsd*S pUZ8002	MacNeil *et al.*, 1992
plasmid		
pOJ446	SCP2*-derived *E. coli-streptomyces* shuttle cosmid *aac(3)IV ori*T	Bierman *et al.*, 1992
pBluescript SK(-)	*bla*, *lacZ*-α	Stratagene
pRSET-B	*pUC* origin, T7 promoter, *bla*	Invitrogen
pSET152	*aac(3)IV int(□C31) oriT ori*(pUC18)	Bierman *et al.*, 1992
pCet2	pOJ446 with 40 Kb *Actinomyces* sp genome fragment	This study
pCet25	pOJ446 with 40 Kb *Actinomyces* sp genome fragment	This study
pCet26	pOJ446 with 40 Kb *Actinomyces* sp genome fragment	This study
pCet27	pOJ446 with 40 Kb *Actinomyces* sp genome fragment	This study
pWUX12b	pSET152 with *bla* and MCS from pACYC-duet	This study
pWUX –Cet2	pWUX12b with 23 Kb pCet2 fragment	This study
pRSET-CetA	pRSET-B with *cetA* gene	This study
pRSET-PrlA	pRSET-B with *prlA* gene	This study
pRSET-Orf9	pRSET-B with *BE-ORF9* gene	This study
pSQ2	pET30a carrying a *salQ* gene	This study

2.2. Cetoniacytone A production and isolation

Actinomyces sp. strain Lu 9419 was maintained as a stock culture on YMG agar. Pre-cultures were prepared by inoculating 50 ml YMG medium with a $1cm^2$ piece of agar from 7 day old cultures in a rotary shaker and incubated for 48 hours at 28 ^oC. The main cultures were inoculated with 15 ml of these pre-cultures and incubated for 96 hours at 28 ^oC. Fermentations were carried out in 500 ml Erlenmeyer flasks with a spring coil inside containing 150 ml of oatmeal medium. At 48 hours, glucose (1 g/ml) was added into the culture to increase the yield of cetoniacytone A.

The harvested culture broths were separated from the mycelia by centrifugation (2500 rpm, 25 minutes) and the supernatant was passed through Amberlite XAD-2 (40/20). The column was washed twice with dd H_2O and the more lipophilic constituents were eluted with methanol. The methanol extract was evaporated to dryness using a rotary evaporator. The crude residue was subjected to silica gel column chromatography (CH_2Cl_2/methanol 9:1; 16/50) and subsequently to Sephadex LH-20 column chromatography (methanol; 16/70). The fractions containing cetoniacytone A were identified by TLC and MS. The chemical structure was confirmed by ^1H NMR in DMSO-d_6 on a Bruker topspin 300 (300 MHz). Mass spectra were taken by ThermoFisher LC-Q advantage.

2.3. DNA isolation and manipulations

Routine genetic procedures such as plasmid DNA isolations, restriction endonuclease digestions, alkaline phosphatase treatments, DNA ligations, and other DNA manipulations were performed according to standard techniques.[129] DNA fragments were excised from agarose gels and residual agarose was removed with the QiaQuick Gel Extraction Kit (Qiagen). PCR was carried out using *Pfx* or *taq* high fidelity or GC-rich DNA polymerase (Invitrogen) according to the manufacturer's protocol. For construction of the genomic library, the genomic DNA of *Actinomyces* sp. was isolated using standard techniques.[129] Purified DNA was partially digested with dilute *Sau*3AI in various concentrations (0.0005 – 0.5 IU/µl). The best partial digestion conditions (0.002 IU/µl, 30min) gave DNA fragments within the range of 35-45 Kb in size. The digested sample containing 35-45 Kb fragments was pooled and ligated to pOJ446 predigested with BamHI and HpaI. *In vitro* packaging was carried out using Gigapack III Gold packaging extract (Stratagene).

2.4. Hybridization of probe for the isolation of cetoniacytone A biosynthetic genes

The *valA* gene from *S. hygroscopicus* var. *limoneus* and the *acbC* gene from *Actinoplanes* sp. encoding 2-*epi*-5-*epi*-valiolone synthases were used as heterologous probes to identify the equivalent gene in the genomic DNA of *Actinomyces* sp. by means of DNA-DNA hybridization experiments. The *valA* gene was obtained from a *valA* expression vector[100] previously constructed in our lab by

cutting it with restriction enzymes NdeI and EcoRI. The *acbC* gene was amplified using PCR reaction from plasmid pAS8/7[99] by primers AcbC-F and AcbC-R which is listed in Table 2-2. Southern blot analysis was performed on Hybond-N nylon membranes (Amersham) with digoxigenin-labeled probes by using the DIG high prime DNA labeling and detection starter kit II (Roche). For this purpose, a 1.24 Kb *valA* gene and a 1.15 Kb *acbC* gene were labeled with digoxigenin-dUTP. First screening was performed by colonies lifting within the whole library (6000 clones) and nine positive clones were identified. The cosmid DNA of these nine positive clones were isolated and digested with BamHI. Subsquently, a second screening was performed by Southern blot and 3 real positive candidates were identified. The identified 3 Kb BamHI fragment was further subcloned into pBluescript SK^- for sequencing.

Sequencing was performed using the Big Dye RR terminator cycle sequencing kit (PerkinElmer Biosystems), and the gels were run on Applied Biosystems capillary 3730 Genetic Analyzer available at the Center for Genome Research and Biocomputing (CGRB) Facilities of Oregon State University.

2.5. Cloning, expression, and purification of cetA, cetB, prlA, BE-orf 9, *and* salQ

cetA, *cetB*, *prlA*, and *BE-orf9*, were overexpressed as histidine-fusion proteins (6XHis) at the N-terminus using the expression vector pRSETB (Invitrogen). The genes encoding CetA, CetB, PrlA, and BE-Orf9 were PCR

amplified using the *Pfx* PCR kit (Invitrogen). The primers used for amplification of these genes are listed in Table 2-2. The *cetA* gene (1.09 Kb) obtained from the cosmid pCet26 by the primers CetA-F and CetA-R was cloned into pRSETB digested with EcoRI and XhoI. The *cetB* gene (552 bp) obtained from the cosmid pCet2 by the primers CetB-F2 and CetB-R2 was cloned into pRSETB digested with EcoRI and BamHI. Similarly, the *prlA* gene (1.24 Kb) and the *BE-orf9* gene (1.25 Kb) were subcloned into pRSETB in frame with the 6XHis tag at BamHI/KpnI and BamHI/EcoRI restriction sites, respectively. The *salQ* overexpressing vector was constructed by the Hong group in Myongji University and was subcloned into pET30a at BamHI and HindIII restriction sites. All expression constructs were confirmed by DNA sequencing. These prepared recombinants were transferred into the expression host *E. coli* BL21(DE3)pLysS. The transformants were cultured in 500 ml of LBBS medium supplemented with ampicillin and chloramphenicol. When the cells reached an OD_{600} of 0.5 at 37 ^{o}C, the growth temperature was reduced to 25 ^{o}C and protein expression was induced by the addition of 0.2 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Cells were harvested by centrifugation after 15 hours incubation and were washed three times with LB broth to remove excess sorbitol from the growth medium. Bacterial pellets were stored at –80 ^{o}C until further use.

Table 2-2. Primers for overexpression and site directed mutagenesis of genes in different pathway.

Name	**Sequence (5' to 3')**
PrlA-F2	GAGGATCCCCATATGCCGGTATCGGCGCAGA
PrlA-R2	CGGGGTACCTCAGAGGGCGACGTCCTCG
BE-orf9-F	GAAGATCTGCATATGAATGACCCAGGACCGATCACCGT
BE-orf9-R	GGAATTCTCACGCATGGGCACGCTCCATCTCCCGC
Cet A-F	CCGCTCGAGCCATATGGCCAATCAGTGGCAGG
Cet A-R	GGAATTCTCATTCGCCGCCTCCCAG
SQF	GGATCCATGACCGGTACGAGCCTGA
SQR	AAGCTTCTAAAGGTTCCTGCCGCCGC
Acbc-F	GAAGATCTGCATATGAGTGGTGTCGAGACGG
Acbc-R	GGAATTCTCAGCGCGGGCGGCCTG
Cet B-F2	CCGCTCGAGCCATATGACCGGGCGGGGCATCC
Cet B-R2	GGAATTCTCATGAAGCTCCTCGGTGATC
LF	ACAGTTCGACGCTTTCCACGAT
LR	TGACCTTGCTCGTGATGAAC
RF	CGGCCACCGAATAGCACAT
RR	CGCCCATCCGCACGACATCTTCC
CF	CCGCTCGAGCCATATGACATTCATCGAGCTGAACCCGTT
CR	GGAATTCTCACCGCGCTCCGGTGTCCTCAG
CetBMF1	CCGGAGCGGTCGCGGTGCACGGCGTCGCCTACACGGTGCCCGA
CetBMR1	TCGGGCACCGTGTAGGCGACGCCGTGCACCGCGACCGCTCCGG
CetBMF2	ACAGCGACTGGGGCGGGCACGGCCTGGCCATCCACGTCGCGGA
CetBMR2	TCCGCGACGTGGATGGCCAGGCCGTGCCCGCCCCAGTCGCTGT
CetBMF3	TGGGCCCGGTGACCAACCTGGGGCTCTTCGAGTACGCCGCGCC
CetBMR3	GGCGCGGCGTACTCGAAGAGCCCCAGGTTGGTCACCGGGCCCA
CetBMF4	CGCCGTGGGGGATGCAGCTGGGACTGATCAACCTGCCCGCGGG
CetBMR4	CCCGCGGGCAGGTTGATCAGTCCCAGCTGCATCCCCCACGGCG

2.6. Enzyme assay for the activities of CetA, PrlA, BE-Orf9, and SalQ

To examine the enzyme activity of putative 2-*epi*-5-*epi*-valiolone synthases, bacterial cell pellets were lysed in K_2HPO_4/KH_2PO_4 (25 mM), NaCl (300 mM), pH 7.4, containing 0.2 mM $CoCl_2$ and the protease inhibitor 4-(2-aminoethyl)-benzenesulfonylfluoride (AEBSF) (200 ug/ml). The cell lysate was sonicated for two 30 sec bursts on a Microson Ultrasonic Cell Disruptor XL on setting 6, and

clarified by centrifugation to yield the cell free extract. This mixture was used directly in the cell-free protein assay or the enzymes were further purified using Ni-NTA agarose beads according to manufacturer's specifications (Qiagen). The purified protein was dialyzed in 2 L dialysis buffer [K_2HPO_4/KH_2PO_4 pH 7.5 (25 mM), $CoCl_2$ (0.2 mM)] and visualized using 10% SDS-PAGE followed by Coomassie staining.

Enzyme activity was analyzed by the addition of cell-free extract or purified protein to a reaction mixture, as described below. The enzyme assay (100 µl volume) was performed at 30 °C for 2 h in 25 mM phosphate buffer (pH 7.5) containing 50 µl of protein (0.3 mg/ml), 5 mM substrate, 0.05 mM $CoCl_2$, 1 mM NAD^+ and 2 mM KF. AminoDAHP was a gift from Prof. Heinz G. Floss[130]. Sedoheptulose 7-phosphate, DAHP, glucose 6-phosphate, fructose 6-phosphate, mannose 6-phosphate, and ribose 5-phosphate were purchased from Sigma. After incubation, samples were visualized using silica gel thin layer chromatography (TLC) (butanol/ethanol/water 9:7:4), and the substrate and reaction product were detected as blue spots with a cerium sulfate/phosphomolibdate containing reagent. The reaction mixture was lyophilized and the reaction products were extracted with methanol. After drying, the extracts were mixed with drops of SIGMA-SIL-A and dried under argon. The products were reextracted with *n*-hexane and analyzed by GC-MS (Hewlett Packard 5890 series II gas chromatograph).

2.7. Enzyme assay for the activity of CetB

To examine CetB enzyme activity, bacterial cell pellets were lysed in K_2HPO_4/KH_2PO_4 (25 mM), 150 mM NaCl, pH 7.5. The cell lysate was sonicated for two 30 sec bursts (Misonix, microson ultrasonic cell disruptor) and clarified by centrifugation to yield the cell free extract. This mixture was further purified using Ni-NTA agarose beads with elution buffer containing 300 mM imidazole. The purified protein was dialyzed in 2 L dialysis buffer [K_2HPO_4/KH_2PO_4 pH 7.5 (25 mM)] and visualized using 10% SDS-PAGE followed by Coomassie staining.

The enzyme assay (100 µl volume) was performed at 30 °C for 3 h in 25 mM phosphate buffer (pH 7.5) containing 30 µl of purified CetA and 30 µl of cell free or purified CetB (0.3 mg/ml), 5 mM sedoheptulose 7-phosphate, 0.05 mM $CoCl_2$, 1 mM NAD^+, and 2 mM KF. The reaction products were extracted with methanol using the same procedure as CetA enzymatic assay. The substrate and the reaction product were detected on TLC (butanol/ethanol/water, 9:7:4) followed by staining with cerium sulfate. The reaction products were mixed with drops of SIGMA-SIL-A and dried under argon. The products were reextracted with *n*-hexane and analyzed by GC-MS (Hewlett Packard 5890 series II gas chromatograph). CetB was also incubated with different substrates including 1L-*epi*-2-inosose, D-(+)-gluconic acid delta-lactone, mannose, shikimic acid, dehydroshikimic acid, and amino-dehydroshikimic acid. The reaction conditions is 5 mM substrate, 0.05 mM $CoCl_2$, 1 mM NAD^+, and 2 mM KF with 30 µl of CetB

cell free extract at 30 °C for 3 h in 25 mM phosphate buffer (pH 7.5). If sedoheptulose 7-phosphate was used, it was also incubated with CetA to convert it to the proposed CetB substrate, 2-*epi*-5-*epi*-valiolone.

2.8. Gel filtration chromatography

CetB (1 mg/ml) in 200 µl potassium phosphate (25 mM, PH = 7.4) was applied to a HiLoad 26/60 Superdex 75 prep grade gel filtration column (GE Healthcare) equilibrated with 25 mM potassium phosphate pH = 7.4, 150 mM NaCl, fitted to a System Gold HPLC work station (BECKMAN). The column flow rate was 0.5 ml/min. Retention time of proteins, determined by measurement of OD 280 nm, was compared to those for protein standards. The protein size was estimated by making the standard curve of K_{av} to molecular weight. The K_{av} for the individual proteins can be calculated as follows: $K_{av} = (V_R-V_O)/(V_C-V_O)$ where V_O = void volume of the column. The void volume was determined using Blue Dextran 2000. V_R = elution volume of the proteins, V_C = the geometric bed volume which was 320 ml in this study.

2.9. CetB site-directed mutagenesis and metal cofactor analysis

Circular PCRs were performed to create all CetB single mutants (H14G, E76G, H98G, E151G) and double mutants (H14G/E76G, H14G/H98G, H14G/E151G, E76G/H98G, E76G/E151G, H98G/E151G). The composition of the PCR was as follows: 5 µl of 10× *Taq* Hi-fidelity polymerase buffer; 125 nM each

of dATP, dTTP, dGTP, and dCTP; 125 ng of each primer; 10 ng plasmid pRSETB carrying the *cetB* gene; H_2O to 49 µl; and 1 µl (2.5 U) *Taq* Hi-fidelity polymerase (Invitrogen). The primers used for this study are listed in Table 2-2 (CetBMF1 and CetBMR1 for H14G; CetBMF2 and CetBMR2 for H98G; CetBMF3 and CetBMR3 for E76G; CetBMF4 and CetBMR4 for E151G). For the double mutants, the single mutants were used as templates and the same PCR strategy was used to create second mutations. Following PCR, the mixtures were incubated with 1.5 µl (15 U) of DpnI (Promega) for 3.0 h to selectively digest the methylated parent plasmids, and the resulting PCR products were analyzed by 0.8% agarose gel electrophoresis. The products were transferred into DH-10B competent *E. coli* cells with selection for resistance to ampicillin (100 µg/ml), and successful mutagenesis was confirmed by DNA sequencing. Mutated plasmids were transferred into BL21(DE3) pLysS *E. coli* cells with selection for resistance to ampicillin (100 µg/ml) and chloramphenicol (20 µg/ml).

All enzymatic characterizations of CetB mutants were performed in a similar manner as that for CetB wild type. For the metal cofactor analysis, CetB mutant H98G was used as this protein binds less strongly to the metal cofactor. After 3 hours reaction of CetA with sedoheptulose 7-phosphate, 1,10-phenanthroline (1 mM) was added and the mixture incubated on ice for 20 min. Different metal ions were tested by adding solution of their corresponding salts ($CoCl_2$, $ZnCl_2$, $MgCl_2$, $MnCl_2$, $FeSO_4$, $NiCl_2$, $CuSO_4$) (1.5 mM) into the reaction

mixture. CetB mutant H98G was then added to each reaction and incubated for 3 hours at 30 ºC.

2.10. DNA sequencing and analysis

DNA sequencing of *cetA, prlA* and *BE-orf9* was done with Applied Biosystems capillary 3730 Genetic Analyzer. *ValA*, *acbC* and *aroB* sequences were retrieved from genbank at the NCBI database. The complete genome sequence of *Frankia alni* ACN14a and the sequence of BE-40644 gene cluster were also obtained from NCBI database. Deduced amino acid sequence alignment was performed with T-COFFEE sequence analysis. The genome walking approach was employed to sequence the cosmid containing the cetoniacytone biosynthetic gene cluster. Potential open reading frames (ORF) were identified using ORF finder and were searched for homology using the BLAST sever.[131]

2.11. Protoplast transformation and heterologous expression of cetoniacytone A biosynthetic gene cluster

To prepare *S. lividans* protoplast, 0.1 ml of spore suspension was added into 25 ml YEME medium in a 125 ml baffled flask and incubated for 36-40 h at 30 ºC. The cells were harvested by centrifugation at 1000 x g for 10 min and washed twice with 10.3% sucrose. The mycelia were then treated with lysozyme solution (1 mg/ml lysozyme in P buffer[128]) at 30 ºC for 20 minutes. After lysozyme treatment, the cells were washed twice with P buffer to remove lysozyme and then were

passed through sterilized cotton wool. The collected protoplasts were pelleted by gently spinning in a centrifuge (1000 x g, 7 minutes) and resuspended in 1 ml P buffer. The protoplasts were stored at –80^{o}C until further use.

The entire cetoniacytone A gene cluster was cut out from the pCet2 by SpeI and AflII and ligated into pWUX12b at XbaI and AflII restriction site. pWUX12b is a pSET152 derivative with a *bla* gene inserted in ApaI site and pACYC Duet-1 multiple cloning site replaced from BamHI to EcoRV site. The resulting plasmid was designated pWUX-Cet2. pWUX-Cet2 was then transferred into three different *S. lividans* strains (1326, T7, TK24) by protoplasts transformation. All transformants were confirmed by PCR of three fragments in the gene cluster (N-terminal, mid-section, and C-terminal PCR products could be obtained) with three sets primers listed in Table 2-2 (LF, LR; CF, CR; and RF, RR). *S. lividans* harboring pWUX-Cet2 was grown in 500 ml Erlenmeyer flasks containing 100 ml oatmeal medium and apramycin (50 ug/ml). Fermentation was carried out in the same way as that of *Actinomyces* sp. (strain Lu 9419). Metabolites from 1 liter of culture broth were extracted with Amberlite XAD-2. The crude extracts were examined by LC-MS.

2.12. Sequence alignment and phylogenetic analysis of SPCs

SPCs sequences were compiled from GenBank and aligned using the ClustalW program. An unrooted maximum likelihood phylogenetic tree was

created using the Seqboot, Proml, and Consense programs in the Phylip software package. Sequence analysis was conducted using the Jones-Taylor-Thorton probability model with global rearrangements and randomized input order. The *E. coli* glycerol dehydrogenase protein sequence was used as an out-group. Support for each node was evaluated with 100 bootstrap replicates of a heuristic search with two random stepwise addition sequences for each replicate. Active site residues of SPCs protein sequences were predicted using comparative alignment with the previously crystallized DHQS from *A. nidulans* (CAA28836).

Chapter 3

Identification of the Cetoniacytone Biosynthetic Gene Cluster

3.1. Introduction

Cetoniacytone A is a structurally unusual C_7N-aminocyclitol produced by *Actinomyces* sp. strain Lu9419, which was isolated from the intestines of the rose chafer beetle. Different from other C_7N-aminocyclitols, such as validamycin and acarbose which have an amino group at the C-1 position, cetoniacytone A contains an amino group at the C-2 position. Biological activity assays of cetoniacytone A showed no antimicrobial activity against Gram-positive or Gram-negative bacteria, but it showed a moderate growth inhibitory activity against HEP G2 (GI_{50} = 3.2 μM) and MCF 7 (GI_{50} = 4.4 μM) human tumor cell lines.[98] The mechanism of action of cetoniacytone A as a cell growth inhibitor is still unclear. However, many studies have been conducted on related compounds, the epoxyquinomicins,[88-90] which share the same core cyclitol moiety as cetoniacytone A. Analogs of epoxyquinomicin have been found to inhibit the activity of the nuclear factor- KB (NF-KB).[120, 123] As a novel NF-KB inhibitor, dehydroxymethylepoxyquinomicin has been reported to inhibit the growth of several cancer cells including bladder cancer, renal cancer and breast cancer cells.[121, 124-126]

Preliminary feeding experiments with [U-^{13}C]glucose by the Zeeck

laboratory established the biosynthetic origin of cetoniacytone A, which proceeds *via* the pentose phosphate pathway. Addition of sodium [1-^{13}C]acetate to cultures of the producer led to the labeling of the carbonyl atom of the acetyl group of cetoniacytone A. This corresponds with the proposed pathway, in which the introduction of the acetyl side chain takes place during the late stages of biosynthesis.[98] In collaboration with the Zeeck group, we then carried out feeding experiments with 2-*epi*-5-*epi*-valiolone and valienone. The result revealed that only 2-*epi*-5-*epi*-valiolone, but not valienone, was incorporated into cetoniacytone A, suggesting that a significantly distinct pathway is involved in cetoniacytone biosynthesis, compared to those of acarbose and validamycin. In order to gain further information regarding the biosynthesis of cetoniacytone, genetic and biochemical approaches have been pursued. This chapter will focus on the identification and functional analysis of the biosynthetic gene cluster of cetoniacytone A.

3.2. Results and discussion

3.2.1. Validation of the cetoniacytone producing strain

The cetoniacytone producer, *Actinomyces* sp. strain Lu9419, was cultured in oatmeal medium supplemented with sodium acetate (1g/L). Without sodium acetate in the medium, strain Lu9419 did not show any color change during 96 h of fermentation and produced very low yields of cetoniacytone A. However, when sodium acetate was added into the medium, after 96 hours fermentation, the culture

broth became dark brown and the yields of cetoniacytone A increased to 5 mg/L. This data suggested that adding sodium acetate may have two effects on the fermentation. One is that sodium acetate is a limiting precursor utilized directly in cetoniacytone A biosynthesis. Secondly, sodium acetate may indirectly effect cetoniacytone production by altering the transcriptional regulation of secondary metabolite genes. Consistent with previous studies,[98] adding glucose (lg/L) at 48 hours also increased the yields of cetoniacytone A. It was previously suggested that feeding glucose at 48 h after fermentation initiation correlates with the onset of cetoniacytone biosynthesis.[98] Early addition of glucose had no influence on the yield of cetoniacytone A since the strain would use it as energy source. Previous studies by the Zeeck laboratory revealed that cetoniacytone A was only found in the culture broth, rather than in the cell pellet. Thus, the culture broth was harvested by centrifugation (5000 rpm, 10 min) to separate the mycelia and culture filtrate. The filtrate was passed through Amberlite XAD-2 and eluted with methanol. The organic solvent was evaporated and the residue was further separated by successive silica gel and Sephadex LH-20 column chromatography. Detection of cetoniacytone A was carried out on silica gel TLC-plates and visualized with UV-light at 254 nm and color reaction with anisaldehyde-H_2SO_4 or cerium sulfate. On TLC plates, cetoniacytone A showed an Rf = 0.11 ($CHCl_3$/MeOH = 9:1). Cetoniacytone A was detected by TLC and MS with APCI positive or negative ion modes and showed m/z of 214 ($M+H^+$) (Figure 3-1) or 212 ($M-H^-$), respectively. ^1H NMR (300MHz, DMSO-d_6) showed signals at δ 2.05 (s, 3H), 3.72 (d, 1H), 3.75

(dd, 1H), 3.87 (dd, 1H), 4.49 (s, 1H), 4.73 (br t, 1H), 6.00 (br s, 1H), 6.65 (s, 1H), 9.87 (s, 1H) (Figure 3-1). This data is consistent with the results published by Zeeck and coworkers.[98]

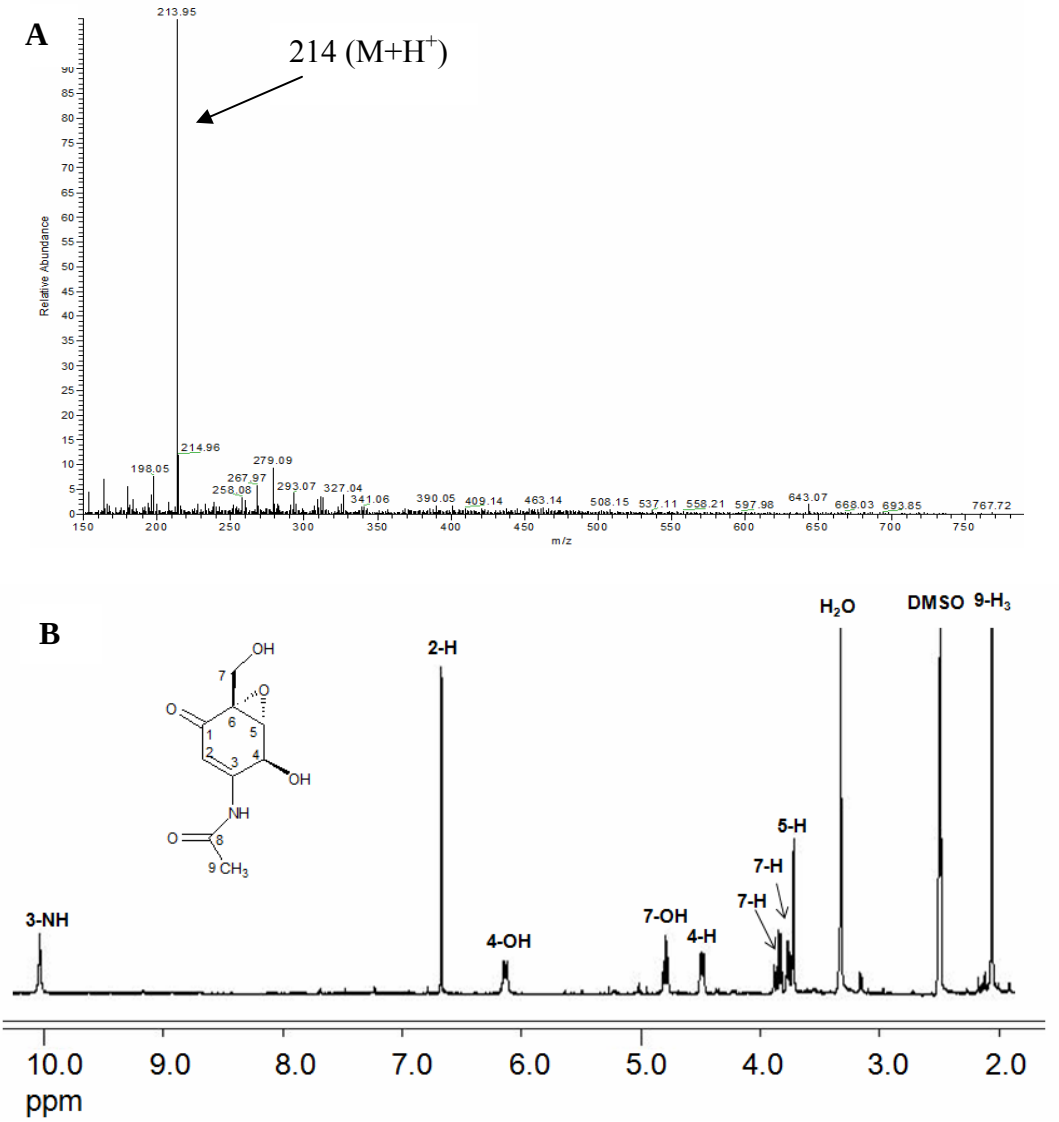


Figure 3-1. Mass and ^{1}H NMR spectra of cetoniacytone A. (A) APCI positive ion mass spectrum of cetoniacytone A. (B) ^{1}H NMR spectrum of cetoniacytone A.

3.2.2. Construction and screening of the cosmid library for the cetoniacytone A biosynthetic gene cluster.

A genomic library of *Actinomyces* sp. strain Lu 9419 was constructed in the *E. coli–Streptomyces* shuttle vector pOJ446[132] using the DNA fragments derived from partial digestion with Sau3A. *In vitro* packaging yielded approximately 6000 apramycin-resistance clones. As cetoniacytone, acarbose, and validamycin share the same initial cyclization reaction by a common enzyme which has similar activity, the genes encoding the 2-*epi*-5-*epi*-valiolone synthases from the acarbose and validamycin pathways were used as heterologous probes to screen the *Actinomyces* sp. Lu 9419 genomic library. Screening with *valA* as probe did not yield any positive colonies. However, screening with the *acbC* probe revealed nine positive cosmids. Cosmid DNA was isolated from the nine positive clones. Restriction enzyme digestion and Southern blot analysis revealed that three out of the nine positive clones contained a 3 Kb positive BamHI fragment that hybridized strongly with the *acbC* heterologous probe (Figure 3-2). The other six clones did not have any fragments that hybridized with *acbC*, suggesting that they were false positives in the earlier screening. The common 3 Kb BamHI fragment shared by the three clones (named as pCet25, pCet26 and pCet27) was subcloned, sequenced and analyzed using the BLAST database. It was found that the 3 Kb BamHI DNA fragment contains a full-length reading frame of a 2-*epi*-5-*epi*-valiolone synthase gene. Since this gene was the first gene to be identified in the cetoniacytone gene cluster, we named it *cetA*.

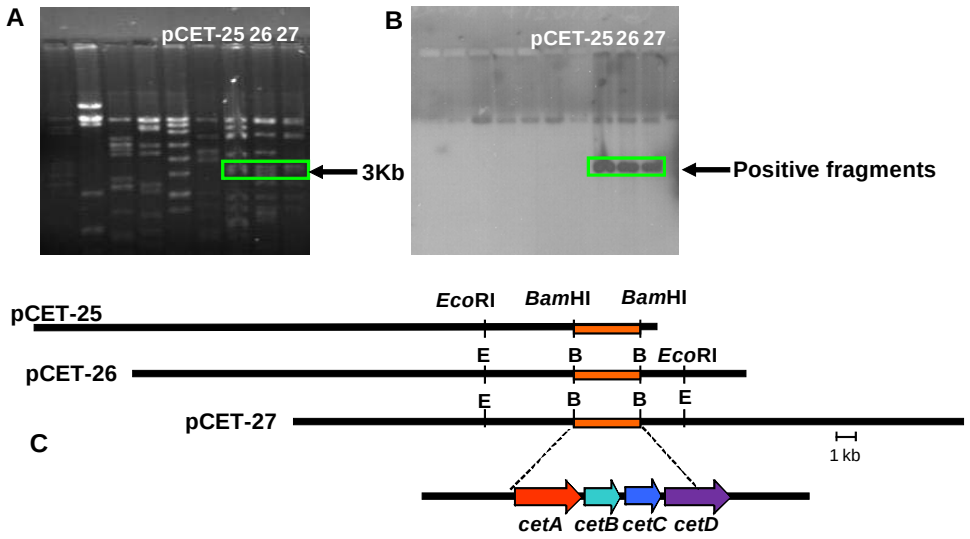
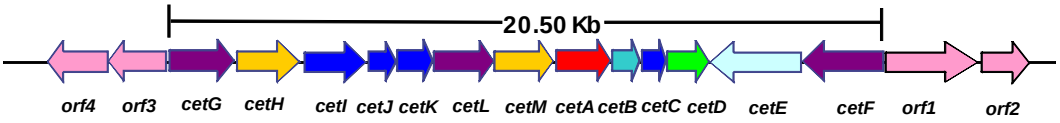


Figure 3-2. Screening of *Actinomyces* sp. genomic library. (A) DNA electrophoresis of the nine positive clones digested with BamHI. (B) Southern blot analysis using DIG-labelled *acbC* as a probe. (C) Digestion pattern of the three positive clones. The 3 Kb positive bands are shown in orange.

3.2.3. Sequence analysis of the cetoniacytone A biosynthetic gene cluster

Following the identification of the *cetA* gene, the nucleotide sequence of the 26 Kb DNA region was determined using chromosomal walking. The ORFs were identified using ORF finder analysis (NCBI) and analyzed with software in the BLAST server. The computer-aided analysis of the 26 Kb sequence revealed the presence of 17 complete ORFs. According to the structure of cetoniacytone A and the BLAST results, 20.5 Kb of the sequenced DNA encoding 13 ORFs are predicted to be involved in the biosynthesis of cetoniacytone A. Four ORFs located on the outer region of the 20.5 Kb DNA seem to be unrelated to the biosynthesis of cetoniacytone A.[133-136] The product of each ORF and its deduced function are shown in Figure 3-3.



gene	homology	species	Size (bp)	% identi	% sim
cetA	2-*epi*-5-*epi*-valiolone synthase	*Streptomyces hygroscopicus*	1146	54%	71%
cetB	putative glyoxalase	*Frankia alni* ACN14a	552	76%	83%
cetC	hypothetical protein	*Frankia alni* ACN14a	495	54%	71%
cetD	putative arylamine *N*-acetyltransferase	*Streptomyces murayamaensis*	867	40%	53%
cetE	drug resistance transporter	*Burkholderia thailandensis* E264	1860	42%	56%
cetF	FAD/FMN-containing dehydrogenases	*Brevibacterium linens* BL2	1650	47%	67%
cetG	glucose-methanol-choline oxidoreductase	*Pseudomonas fluorescens*	1359	35%	52%
cetH	adenosylmethionine-8-amino-7-oxononanoate aminotransferase	*Magnetospirillum magneticum* AMB-1	1299	33%	50%
cetI	hypothetical protein	*Streptomyces achromogenes*	1278	35%	51%
cetJ	hypothetical protein	*Actinoplanes* sp. A40644	570	59%	72%
cetK	conserved hypothetical protein	*Frankia alni* ACN14a	720	61%	74%
cetL	putative Pyranose oxidase precursor (Glucose 2-oxidase)	*Frankia alni* ACN14a	1248	66%	76%
cetM	L-alanine:N-amidino-3-keto-*scyllo*-inosamine aminotransferase	*Streptomyces griseus*	1224	54%	66%
orf1	ASPIC/UnbV	*Herpetosiphon aurantiacus* ATCC 23779	1914	40%	54%
orf2	UnbU	*Micromonospora chersina*	1062	46%	61%
orf3	NikS	*Streptomyces tendae*	1179	41%	56%
orf4	NocH (membrane transport protein)	*Nocardia uniformis subsp. tsuyamanensis*	1221	31%	46%

Figure 3-3. Genetic organization and deduced function of the cetoniacytone biosynthetic gene cluster. The ORFs proposed involving in cetoniacytone A biosynthesis are named as *cetA*-*cetM*.

The deduced product of *cetA* showed significant similarity to 2-*epi*-5-*epi*-valiolone synthase, (54% identity, 71% similarity to ValA, and 56% identity, 72% similarity to AcbC).[96] The degree of identity of CetA with AcbC is higher than that of CetA with ValA. It may partially explain why the *acbC* probe hybridized with the cetoniacytone *cetA* gene, whereas the *valA* probe did not work during the course of library screening.

CetB bears similarity with a putative glyoxalase from *Frankia alni* ACN14a (76% identity, 83% similarity).[137] It also showed high similarity with ValD which is defined as a conserved hypothetical protein in the validamycin pathway. Glyoxalase I is a Zn^{2+}-metalloenzyme that catalyzes the glutathione-dependent inactivation of toxic methylglyoxal by catalyzing the conversion of cytotoxic methylglyoxal to *S*-D-lactoylglutathione via a 1,2-hydrogen transfer. The results of recent X-ray crystallographic analysis of glyoxalase I in complex with a transition state analogue and site-directed mutagenesis studies strongly support a base-mediated, proton-transfer mechanism in which the bound diastereomeric substrates undergo catalytic interconversion to product via a Zn^{2+}-coordinated, cis-enediolate intermediate. In cetoniacytone biosynthesis, this enzyme may be involved in the epimerization of 2-*epi*-5-*epi*-valiolone to 5-*epi*-valiolone.

CetD belongs to the family of arylamine *N*-acetyltransferases with the highest identity to a putative arylamine *N*-acetyltransferase from *Streptomyces*

murayamaensis (40% identity, 53% similarity).[138, 139] The arylamine *N*-acetyltransferase (NAT) enzymes have been found in a broad range of both eukaryotic and prokaryotic organisms. The NAT enzymes catalyze the transfer of an acetyl group from acetyl-CoA onto the terminal nitrogen of a range of arylamine, hydrazine and arylhydrazine compounds, and are proposed to mediate the transfer of an acetyl group to the C-2 position of cetoniacytone (Figure 3-4).

CetE appears to belong to a drug resistance transporter family, bearing 42% identity and 56% similarity to a drug resistance transporter of *Burkholderia thailandensis* E264.[140] In *Burkholderia thailandensis* E264, the putative drug resistance transporter belongs to the major facilitator superfamily (MFS) which is one of the two largest families of membrane transporters known. MFS drug transporters are antiporters, which have a unique antiporter motif, also called motif [$G(X_8)G(X_3)GP(X_2)GG$], necessary for the drug/H^+ antiport activity. They translocate substrates against their electrochemical gradient by coupling the movement of an ion, *e.g.*, H^+, along its gradient and function as a drug resistance transporter. Base on the high resolution structures of transporters, all members of the MFS share a similar structure, regardless of their low sequence identity.[141]

The deduced product of *cetF* showed significant similarity (47% identity, 67% similarity) to the known protein FAD/FMN-containing dehydrogenase of *Brevibacterium linens* BL2.[142] Based on CDD (Conserved Domain Database) and

COG (Clusters of Orthologous Groups) assignments and functional annotation, the gene product contains an FAD binding domain. This family of proteins consists of various enzymes that use FAD as a co-factor; most of the enzymes are similar to oxygen oxidoreductases such as vanillyl-alcohol oxidase (VAO). VAO catalyzes the oxidation of a wide variety of substrates, ranging from aromatic amines to 4-alkylphenols. Other members of this family include D-lactate dehydrogenase, which catalyses the conversion of D-lactate to pyruvate using FAD as a co-factor and radical oxidase, which oxidizes the reduced form of the mitomycins and is involved in mitomycin resistance.

CetG bears similarity to the glucose-methanol-choline (GMC) oxidoreductase from *Pseudomonas fluorescens* (35% identity, 52% similarity).[143] This family of proteins binds FAD as a cofactor. Similar to CetF, CetG was proposed to be involved in the redox reaction. However, it is still unclear about the timing of this step in cetoniacytone biosynthesis.

CetH is similar to adenosylmethionine-8-amino-7-oxononanoate aminotransferase from *Magnetospirillum magneticum* AMB-1 (33% identity, 50% similarity).[144] It is proposed to mediate the transamination step at the C-3 position.

CetL shares high sequence homology with a putative pyranose 2-oxidase from *Frankia alni* ACN14a (66% identity, 76% similarity).[145-148] Pyranose 2-

oxidase oxidizes D-glucose and other aldopyranoses regioselectively at C-2 to the corresponding 2-keto sugars. In cetoniacytone biosynthesis, this enzyme may catalyze the oxidation of the C-2 hydroxyl group of 5-*epi*-valiolone to a keto group.

CetM also shares homology with aminotransferases, bearing high similarity with the L-alanine:N-amidino-3-keto-*scyllo*-inosamine aminotransferase from *Streptomyces griseus* (54% identity, 66% similarity).[149] It belongs to the DegT/DnrJ/EryC1/StrS aminotransferase family. The members of this family are proposed to utilize pyridoxal 5′-phosphate as a cofactor, and contain aminotransferase enzymes with a variety of molecular functions. The family includes StsA, StsC and StsS from the streptomycin biosynthetic gene cluster. The aminotransferase activity was demonstrated for purified StsC protein as the L-glutamine:*scyllo*-inosose aminotransferase, which catalyses the first amino transfer in the biosynthesis of the streptidine unit of streptomycin. Since both CetH and CetM encode an aminotransferase, we propose these two genes may be redundant genes and only one is active during the cetoniacytone biosynthesis.

The deduced products of *cetC, cetI, cetJ,* and *cetK* are hypothetical proteins with unknown function.[87, 150-152] These enzymes are structurally related to the cupin superfamily of proteins, which is one of the most functionally diverse protein superfamily. However, those involved in the biosynthesis of secondary metabolites are most related to hydroxylation, epoxidation, decarboxyation, dehydration, and

halogenation reactions. Therefore, we predict that *cetC*, *cetI*, *cetJ*, and *cetK* are involved in cetoniacytone A biosynthesis as shown in Figure 3-4.

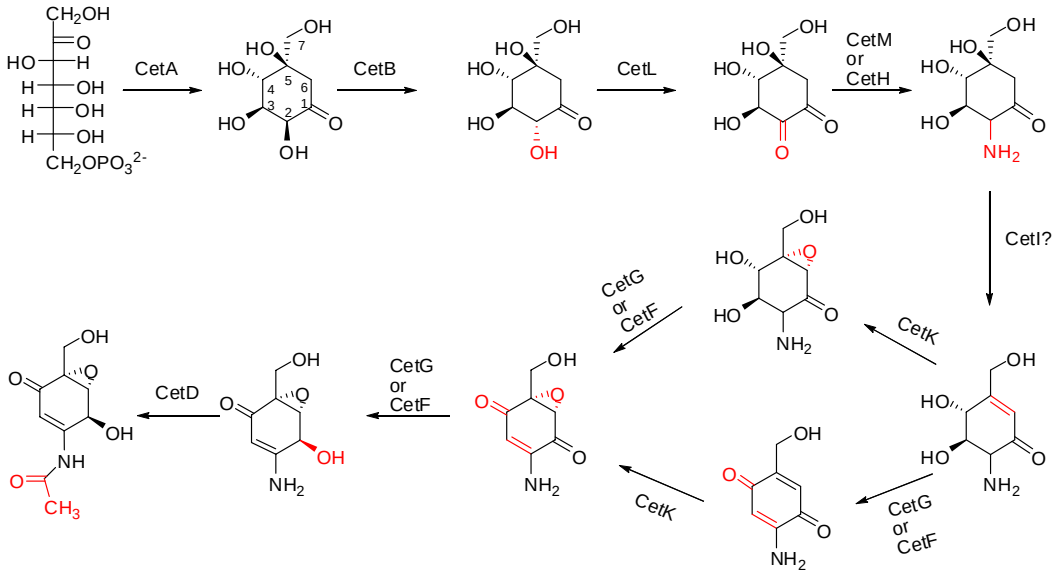


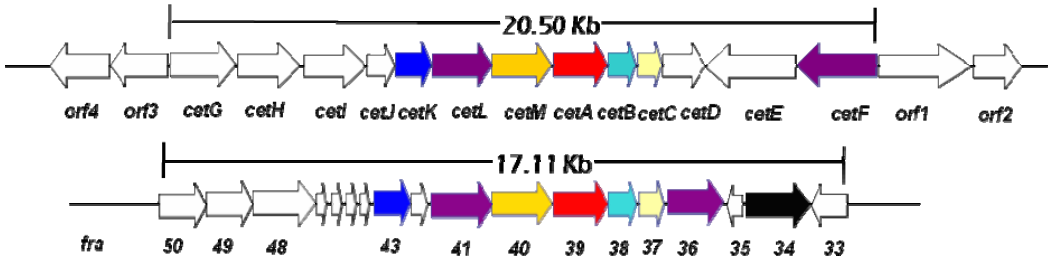
Figure 3-4. Proposed biosynthetic pathway to cetoniacytone A.

In acarbose and validamycin biosyntheses, the activity of a kinase is necessary to phosphorylate intermediates for the activation of the cyclitol ring. It is believed that most intermediates in the acarbose and validamycin biosynthetic pathways are phosphorylated.[105, 95, 153] The phosphorylation of 2-*epi*-5-*epi*-valiolone and valienone in the acarbose and validamycin pathways, respectively, is catalyzed by the newly identified cyclitol-kinases, AcbM (in acarbose biosynthesis) and ValC (in validamycin biosynthesis). However, to our surprise, no kinase gene was found in the cetoniacytone biosynthetic gene cluster suggesting that there is no phosphorylation step necessary in cetoniacytone biosynthesis. In contrast to the validamycin pathway, results of feeding studies indicate that valienone is not

involved in cetoniacytone biosynthesis. This suggests that the downstream modifications of 5-*epi*-valiolone in cetoniacytone biosynthesis are different from those in the validamycin pathway (Figure 3-4). While a number of possibilities may apply for 5-*epi*-valiolone, the most likely scenario would be the oxidation of the C-2 hydroxyl group by the action of CetL (similar to pyranose 2-oxidase) followed by a transamination catalyzed by CetM to give 2-amino-5-*epi*-valiolone. Reduction of C-1 ketone of 5-*epi*-valiolone to an alcohol is unlikely because it would eliminate the driving force for the subsequent dehydration reaction. A C-2/C-3 dehydration is also unlikely at this point, as this would cause the introduction of the amino group to become less possible. Finally, oxidation of C-4 hydroxyl group by CetL or other oxidoreductase (CetF and CetG) is possible, but subsequent oxidation at C-2 is energetically unfavorable. Previous studies have shown that strain LU9419 can produce both cetoniacytone A and cetoniacytone B; cetoniacytone B is the deacetylated form of cetoniacytone A. When sodium acetate was added in the production medium, more cetoniacytone A was isolated. So, it was proposed that acetylation should be the last step in the pathway. However, whether epoxidation occurs prior to C-2/C-3 double bond formation or vice versa cannot be confirmed only based on the genetic information. Further studies are needed to reveal this unique biosynthetic pathway.

3.2.4. Comparison of the cetoniacytone cluster with an unknown gene cluster in Frankia alni *ACN14a genome*

During the genetic analysis of the individual genes in the cetoniacytone A biosynthetic gene cluster, seven genes were found to have high identity to genes in the *Frankia alni* ACN14a genome. Interestingly, the seven homologous genes of *Frankia alni* ACN14a are also clustered together in the genome, indicating that *Frankia alni* ACN14a may have the capacity to produce compounds similar to cetoniacytone A (Figure 3-5). In addition to these seven homologs, a halogenase was also found in this unknown cluster which suggests that the compound may contain one or more halogen atoms. The identity of this gene cluster and the corresponding secondary metabolites may be a subject of further investigation.



cet gene	*fra* gene	Putitive function	identity	similarity
cetK	*fra43*	hypothetical protein	61%	74%
cetL	*fra41*	putative Pyranose oxidase precursor	66%	76%
cetM	*fra40*	aminotransferase	69%	80%
cetA	*fra39*	2-*epi*-5-*epi*-valiolone synthase	71%	84%
cetB	*fra38*	putative glyoxalase	76%	83%
cetC	*fra37*	hypothetical protein	54%	71%
cetF	*fra36*	putative secreted FAD-linked oxidase	52%	70%
	fra34	putative halogenase		

Figure 3-5. Comparison of the unknown cluster in *Frankia alni* ACN14a genome with the cetoniacytone biosynthetic gene cluster. The seven homologous genes are shown in the same color.

3.2.5. Inactivation attempts of the cet *genes*

To examine whether the identified 20.5 Kb gene cluster is involved in the biosynthesis of cetoniacytone A, several strategies, ***e.g.***, gene disruption, heterologous expression of the whole gene cluster, characterization of the individual genes in the cluster, have been attempted. In most cases, gene disruption is the method of choice because it would give more direct and convincing results. Therefore, we set out to create mutants of the cetoniacytone producer, in which one or more genes within the *cet* cluster have been disrupted. However, the attempts were hampered by the lack of an appropriate selection marker for the knockout experiments, as strain Lu9419 is literally resistant to all antibiotics commonly used in molecular genetics work, including apramycin, gentamicin, hygromycin, kanamycin, spectinomycin, streptomycin, thiostrepton, tetracycline, neomycin, puromycin, paromomycin, and chloramphenicol. Without selection markers, the double crossover constructs could not be made and subsequently, the gene inactivation approach was abandoned.

3.2.6. Heterologous expression of cetoniacytone A biosynthetic gene cluster

Since the inactivation of the gene cluster could not be pursued, the second strategy, expression of the whole gene cluster, were carried out in a heterologous host, *Streptomyces lividans*. Sequence analysis of the previously identified cosmid clones pCet25, pCet26 and pCet27, revealed that all three clones contain a truncated gene cluster. In order to find cosmid clones that harbor the complete gene

cluster, the genomic library from the cetoniacytone producer was rescreened using the *cetA* probe. An additional 8 positive cosmids were identified. End sequencing of these 8 positive clones showed only one clone contained the entire cluster, and was named pCet2. Since the library was constructed with the shuttle vector pOJ446 which can be replicated directly in *E. coli* and *Streptomyces*, the cetoniacytone A biosynthetic gene cluster may be expressed directly in *Streptomyces* by cosmid transformation. A number of successful heterologous expression of natural product biosynthetic genes using this vector have been reported.[154-156] However, there were also reports on the instability nature of this cosmid. This proved to be also the case in our system. When pCet2 was transferred into *S. lividans*, and the plasmid reisolated from the transformants, it showed a different digestion pattern from of the original construct (Figure 3-6), which may be due to structural instability of the pOJ446 cosmid vector. Vector pOJ446 was made through a modification of the pSCP103 cosmid that unfortunately removed a DNA fragment that may be responsible for the structural stability of pSCP103 (Figure 3-6).[132]

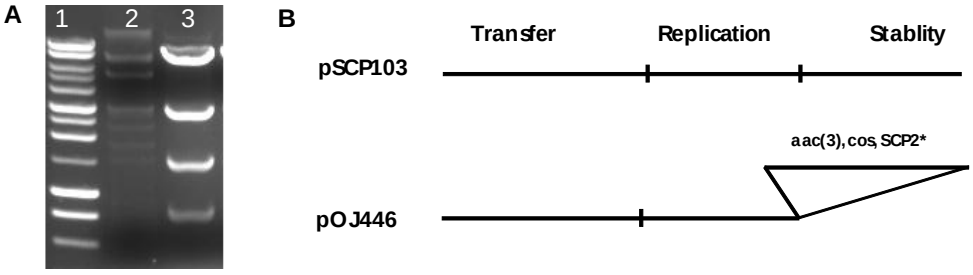


Figure 3-6. Instability of pOJ446. (A) Rearrangement of pOJ446 harboring a large DNA fragment. Lane 1 = 1 Kb molecular marker, lane 2 = BamHI digested pattern of original pCet2, lane 3 = BamHI digested pattern of pCet2 after one time transformation. (B) Structure of pOJ446. pOJ446 was made from pSCP103 and loss the stability part of pSCP103.

In order to overcome this problem, the gene cluster was cut out from the pCet2 by SpeI and AflII and ligated into pWUX12b at XbaI and AflII restriction site (Figure 3-7). pWUX12b is a derivative of the integration vector pSET152 that contains an additional multiple cloning site from pACYC-duet and the *bla* gene from pBluescript SK-. Ligation of the *cet* cluster into pWUX12b gave pWUX-Cet2. To confirm that the entire *cet* gene cluster has been transferred into pWUX12b, a PCR reaction (Figure 3-7) was carried out using three different sets of primers listed in Chapter 2 (Table 2-2). LF-LR is the primer set for the amplification of left end of the gene cluster, RF-RR and CF-CR are primer set for the amplification of right end and central part of the gene cluster, respectively. The results showed that all three fragments were amplified as expected, indicating that the whole gene cluster has been successfully transferred into pWUX12b.

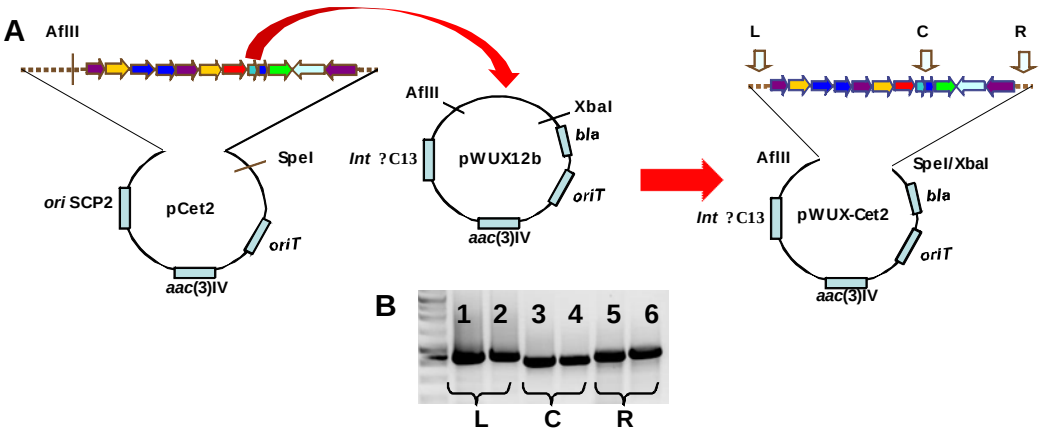


Figure 3-7. Construction strategy for pWUX-Cet2. (A) Transfering gene cluster from pCet2 to pWUX-Cet2. (B) PCR confirmation of the gene cluster in pWUX-Cet2. Lane 1,3,5 are PCR product using pCet2 as template and lane 2,4,6, are PCR product using pWUX-Cet2 as template. Lane 1 and lane 2 are the 875 bp left fragment of the gene cluster, lane 3 and lane 4 are the 867 bp center fragment of the gene cluster, and lane 5 and lane 6 are the 805 bp right fragment of the gene cluster.

The pWUX-Cet2 plasmid was then transferred into three different *S. lividans* strains (T7, TK-24, and 1326). To confirm that the whole gene cluster has been transferred into the genome of *S.lividans*, another series of PCR reactions were carried out using the same sets of primers as those described in Figure 3-7. The genomic DNAs of the transformants were used as templates. The PCR products indicated that pWUX-Cet2 has been successfully transferred and integrated into the genome of the all three *S. lividans* strains. (Figure 3-8).

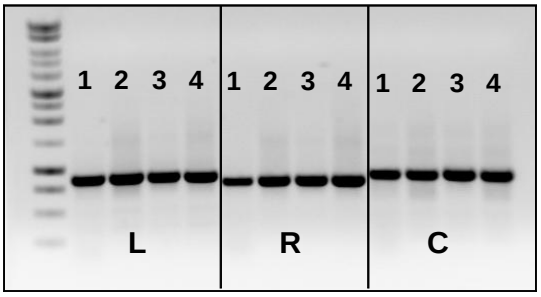


Figure 3-8. PCR confirmation of the *cet* gene cluster transferred into *S. lividans*. By using (1) original cetoniacytone A producer, (2) *S.lividans* T7 mutant, (3) *S.lividans* 1326 mutant, and (4) *S.lividans* TK24, as templates, PCR products were checked by DNA electrophoresis. L represent the left fragment of the gene cluster, C represent the central fragment of the gene cluster, R represent the right fragment of the gene cluster.

The three different *S. lividans* mutant strains harboring pWUX-Cet2 were grown in 500 ml Erlenmeyer flasks containing 100 ml oatmeal medium and the production of cetoniacytone A or related compounds was examined by LC-MS. However, no production of cetoniacytone A could be detected in culture broths of the mutant strains. While there is no clear explanation for the lack of cetoniacytone A production in these strains, it is widely accepted that different hosts may have different regulatory system or lack the necessary supporting system that in turn

may reduce or inhibit the expression of foreign genes and the production of new secondary metablites. Because gene inactivation and heterologous expression approaches did not provide sufficient information to directly confirm the identity of the cetoniacytone biosynthesis gene cluster, we then decided to pursue this objective via biochemical approaches. Detail descriptions on these investigations are given in Chapter 4 and 5.

Chapter 4

Recombinant Expression and Characterization of 2-*epi*-5-*epi*-Valiolone Synthase (CetA) and Its Functional Relationship Within the Sugar Phosphate Cyclase Superfamily

4.1. Introduction

As described in Chapter 1, the biosynthesis of a diverse array of clinically-relevant natural products involves pathways that parallel the initial steps in shikimate biosynthesis (the primary route to aromatic amino acids).[18] Similar to the shikimate pathway, each parallel pathway requires the cyclization of a specific sugar phosphate substrate. In the case of shikimate biosynthesis, 3-deoxy-D-arabinoheptulosonate 7-phosphate (DAHP) is converted to 3-dehydroquinic acid (DHQ), whereas in the biosynthesis of 3-amino-5-hydroxybenzoic acid (AHBA), the core unit of many ansamycin antibiotics, is mediated through the conversion of aminoDAHP to aminoDHQ (Figure 4-1).[74] Furthermore, the aminoglycoside antibiotics require the conversion of glucose 6-phosphate to 2-deoxy-*scyllo*-inosose,[49, 50] and the biosynthesis of the C_7-cyclitols is mediated through conversion of sedoheptulose 7-phosphate to 2-*epi*-5-*epi*-valiolone (Figure 4-1).[99] In each case the cyclization reaction is catalyzed by a distinct sugar phosphate cyclase (SPC) that shares high homology with DHQ synthase.[29]

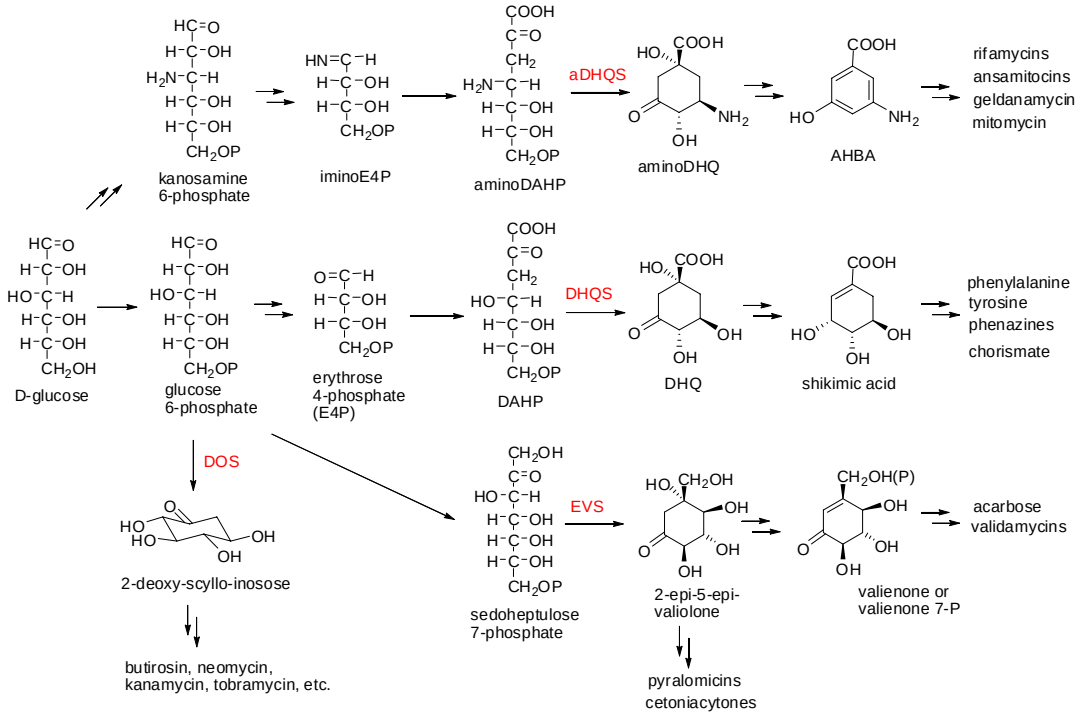


Figure 4-1. Biosynthetic pathways from different sugar phosphate intermediates. DOIS = 2-deoxy-*scyllo*-inosose synthase, EVS = 2-*epi*-5-*epi*-valiolone synthase, DHQS = dehydroquinate synthase, aDHQS = aminodehydroquinate synthase.

To investigate if CetA is indeed a 2-*epi*-5-*epi*-valiolone synthase that is involved in the biosynthesis of cetoniacytone A, biochemical characterization of the enzyme was carried out. In addition, because little is known about the various subclasses of the sugar phosphate cyclase superfamily, detailed comparative analysis of these enzymes has also been performed. The results have led to the identification of several related genes that are involved in the biosynthesis of other aminocyclitol-containing compounds, including a putative novel set of SPCs, which are more widely distributed in fungi and cyanobacteria.

4.2. Results and discussion

4.2.1. In silico *analysis of cetA and other sugar phosphate cyclases.*

Protein sequence comparisons revealed that CetA is highly similar with ValA and AcbC. The deduced CetA protein has 56% and 54% identity with the AcbC and ValA, respectively. It also has high similarity to DHQS,[157] which is known to be responsible for the cyclization of DAHP to DHQ (Figure 4-2). The alignment also showed a number of highly conserved residues between DHQS and 2-epi-5-epi-valiolone synthase, some of which are predicted to be involved in the catalysis. This is particularly true for the cofactors NAD^+ and divalent ion binding regions, in which Asp146, Glu194, His271, and His287 in the DHQS are also found in CetA (Asp145, Glu192, His263, His279). However, other catalytic residues of the DHQS, such as Lys250, Asn268, His275, and Lys356, have been substituted with different amino acid residues in CetA (Table 4-1). It is clear that these two groups have different conserved activity site sequences. This has also been confirmed by the enzyme analysis that AcbC does not cyclize DAHP[99] and AroB (DHQS) does not cyclize sedoheptulose 7-phosphate.[18]

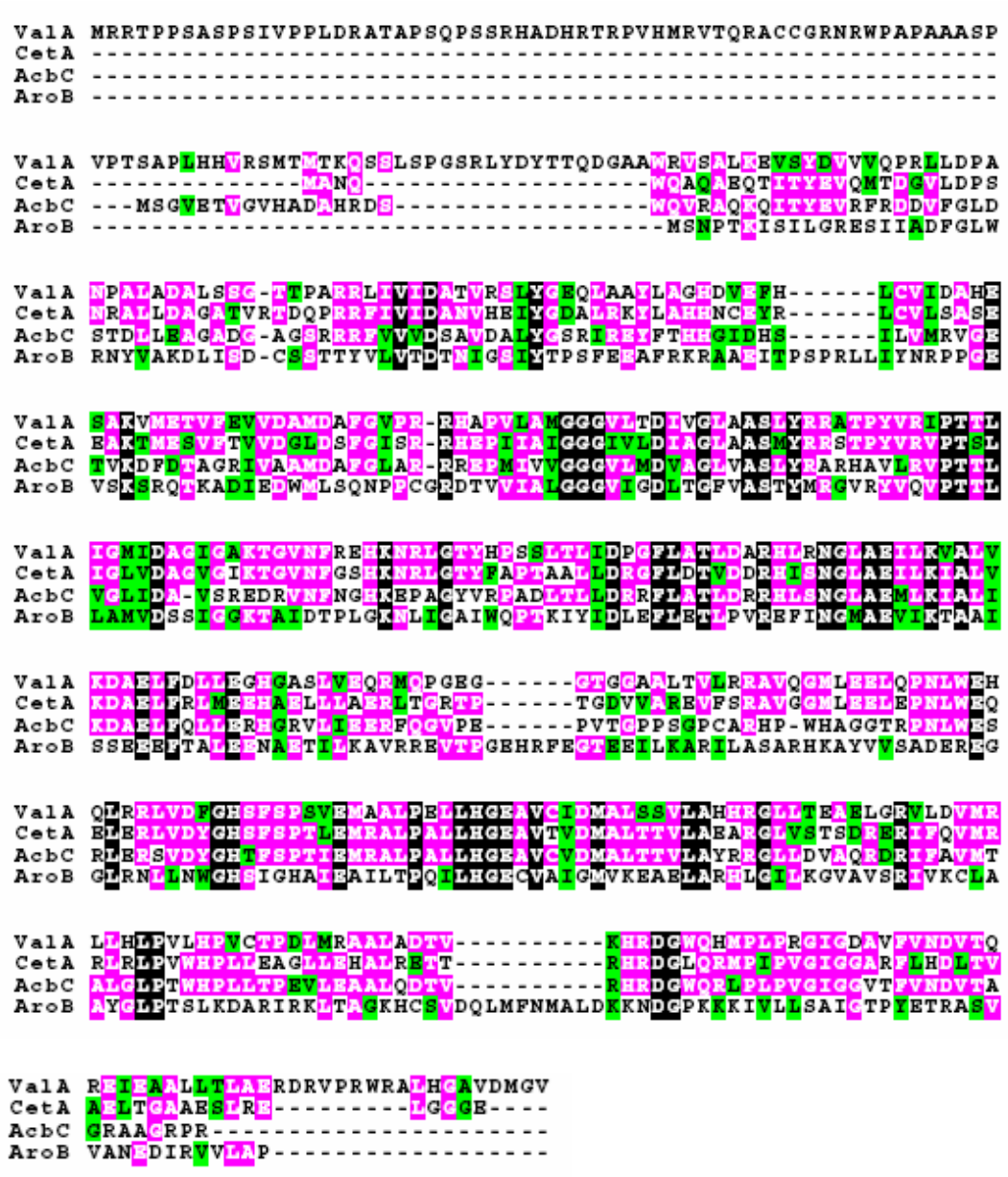


Figure 4-2. Sequence alignment of CetA, ValA, AcbC, and AroB.

Table 4-1. Sequence alignment of the sugar phosphate cyclase active-site residues. Colored boxes represent signature active-site residues specific for a subfamily of SPCs. Red boxes indicate outliers from conserved sequences. The indicated active-site residues are based on the crystal structure of the DHQ synthase from *Aspergillus nidulans*. Nos1-Nos3 are putative DHQsynthases from *Nostoc punctiforme*. Anb1 and Anb2 are putative DHQsynthases from *Anabaena variabilis*.

Protein	130	148	152	162	194	197	250	264	268	271	275	287	356	Proposed Function
DHQ	R	D	K	N	E	K	K	R	N	H	H	H	K	DHQ Synthase
RifG	R	D	K	N	E	R	K	R	N	H	H	H	K	Amino DHQ synthase
GdmO	R	D	K	N	E	R	K	R	N	H	H	H	K	Amino DHQ synthase
Asm47	R	D	K	N	E	R	K	R	N	H	H	H	K	Amino DHQ synthase
MitP	R	D	K	N	E	R	K	R	N	H	H	H	K	Amino DHQ synthase
BtrC	R	D	K	N	E	K	K	G	E	H	H	H	K	2-deoxy-*scyllo*-inosose synthase
KanA	R	D	K	N	E	K	K	G	E	H	H	H	K	2-deoxy-*scyllo*-inosose synthase
Neo7	R	D	K	N	E	K	K	G	E	H	H	H	K	2-deoxy-*scyllo*-inosose synthase
GntB	R	D	K	N	E	K	K	G	E	H	H	H	K	2-deoxy-*scyllo*-inosose synthase
AcbC	A	D	E	E	E	K	P	R	D	H	P	H	P	2-*epi*-5-*epi* valiolone synthase
ValA	R	D	K	N	E	K	P	R	D	H	P	H	P	2-*epi*-5-*epi* valiolone synthase
CetA	R	D	K	N	E	K	P	R	D	H	P	H	P	2-*epi*-5-*epi* valiolone synthase
Nos 1	R	D	K	N	E	K	K	R	N	H	H	H	K	DHQ synthase?
Nos 2	R	D	K	N	E	K	P	R	A	H	P	H	P	2-*epi*-5-*epi*-valiolone synthase?
Nos 3	R	D	K	N	E	K	P	R	D	H	H	H	E	Unknown?
Anb 1	R	D	K	N	E	K	K	R	N	H	H	H	K	DHQ synthase?
Anb 2	R	D	K	N	E	K	P	R	A	H	P	H	P	2-*epi*-5-*epi*-valiolone synthase?

4.2.2. Isolation of prlA *from the pyralomicin producer*

Sequence alignment of AcbC, ValA, CetA, and AroB, were used to develop degenerate primers specific for the SPC family of 2-*epi*-5-*epi*-valiolone synthases. Those degenerate primers were then used to PCR amplify a partial 2-*epi*-5-*epi*-valiolone synthase gene from the pyralomicin producer, *N. spiralis*. This partial gene fragment was utilized as a probe to screen a genomic DNA library from *N. spiralis* that was generated using the Copy Control Fosmid System (Epicentre). A total of 5000 colonies were screened, yielding 6 positive overlapping clones. Sequencing of approximately 40 Kb DNA from the overlapping cosmids revealed a

cyclitol biosynthetic gene cluster containing 27 open reading frames (*orfs*) that share high sequence identity with the acarbose, validamycin, and to a lesser extent, the BE-40644 biosynthetic gene clusters. All the work on the pyralomicin biosynthetic gene cluster was mainly done by Dr. Patricia M. Flatt in our laboratory.

Eight of the isolated *orfs* represent genes that are predicted to be involved in cyclitol biosynthesis. In addition, there are a number of genes required for the biosynthesis of the aglycon core unit of pyralomicin and several other *orfs* related to gene expression and/or other regulatory elements. The gene *prlA* encodes a 411 amino acid protein with high identity to the 2-*epi*-5-*epi*-valiolone synthase, showing 56% identity/72% similarity to the validamycin ValA protein, 55% identity/66% similarity with the acarbose AcbC protein, and 57% identity/69% similarity with the BE-40644 Orf9 protein.

4.2.3. Bioinformatic analysis of 2-epi-*5*-epi-*valiolone synthases within the sugar phosphate cyclase superfamily*

Since 2-*epi*-5-*epi*-valiolone synthases contain regions of unique amino acid sequence compared with the DHQ synthases, we reasoned that other families of SPCs would also have unique signature sequences that can be used to more accurately annotate sequences according to their function and to develop biological tools for assessing the metabolic potential of microorganisms. To further analyze this possibility, the Phylip Phylogenetic Package was used to generate a maximum

likelihood phylogenetic tree of the SPC superfamily (Fig. 4-3). Input sequences were selected from GenBank from each family of SPC and included several putative DHQ synthase sequences annotated from genome sequencing projects. *E. coli* glycerol dehydrogenase was used as an out-group and bootstrap analysis was performed to assess the variability at each node (Fig.4-3).

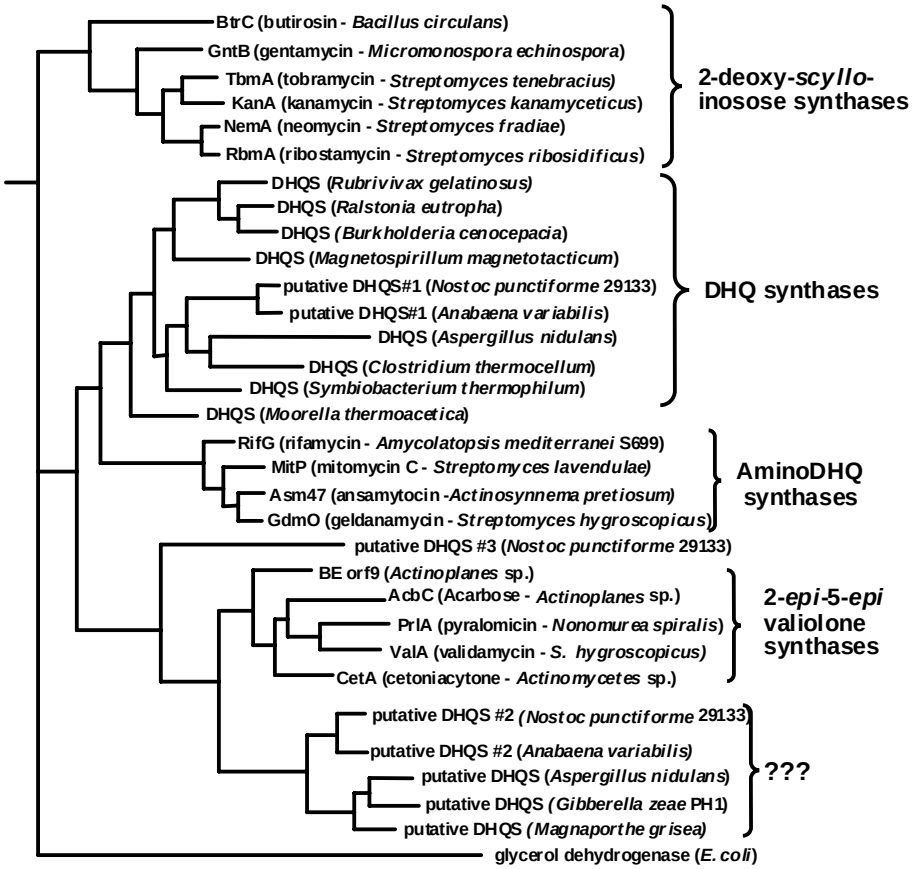


Figure 4-3. Phylogenetic analysis of the superfamily of SPCs. The Phylip software package was used to generate an unrooted maximum likelihood tree. *E. coli* glycerol dehydrogenase was used as an out group, and support for each node was evaluated with 100 bootstrap replicates of a heuristic search with two random stepwise addition sequences for each replicate. Species information and GenBank accession numbers are indicated for each sample.

SPCs from both primary and secondary metabolic pathways show high sequence similarity (25-70% identity), and maximum likelihood analysis has revealed that protein similarity is highly correlated with the predicted enzyme function (Fig. 4-3). Thus, each of the enzymatic functions listed in Figure 4-1 also has a corresponding set of gene products that form predictable clades within this related superfamily of enzymes (Fig. 4-3). Furthermore, alignment of the primary sequences from these pathways has revealed that each subclass of SPC has a unique signature of altered binding pocket residues when compared with DHQ synthases (Table 4-1).

DHQ synthases represent the most well-studied subclass of SPC enzymes.[14, 158] Pioneering work on the fungal species, *Aspergillus nidulans*[20, 159] and *Saccharomyces cerevisiae*,[23] has revealed the presence of a super protein, AROM that mediates five enzymatic steps in the shikimate metabolic pathway. The activity and structure of the DHQ synthase module from the multifunctional AROM protein has been extensively studied. The active site is formed within a cleft between the two domains and is lined by 13 amino acid residues listed in Table 4-1. The C-terminal domain contains most of the residues involved in substrate binding, coordination of the Zn^{2+} cofactor, and catalysis. The mechanism of catalysis was proposed to proceed through a multi-step process including alcohol oxidation, phosphate β-elimination, carbonyl reduction, ring opening and intramolecular aldol condensation.[29]

Since the discovery of the AROM protein, a substantial list of DHQ synthases have been isolated from other fungi, bacteria, and higher order plants, suggesting that shikimate biosynthesis is a universal metabolic pathway shared by all of these organisms.[14] Interestingly, all of the bacterial and plant DHQ synthase homologs isolated to date, exist as monofunctional enzymes containing only the DHQ synthase activity. Thus, the evolution of the AROM super protein appears to have come after the branch point of fungi with bacteria and higher order plants. In spite of the differences in the overall architecture of DHQ synthases, all of the enzymatically-characterized DHQ synthases that have been isolated from bacteria, fungi, and plants retain strict conservation of the 13 binding pocket residues (Table 4-1). Furthermore, site-directed mutagenesis studies converting R130A, K152A, R264A, or H275A as single point mutations, each abolished enzyme activity without disrupting the protein quaternary structure or the ability of the enzyme to bind with the substrate and zinc cofactor, suggesting that each of these residues plays a key role in the catalytic mechanism of DHQ synthase.[40]

The family most related to DHQ synthases are the amino-DHQ synthases from the rifamycin (RifG),[79] geldanamycin (GdmO),[80] ansamitocin (Asm47),[81] and mitomycin C (MitP)[66] biosynthetic pathways (Figure 4-1). Within each of these gene clusters, a cassette of ORFs, including the aminoDHQ synthases, is required to form the precursor, 3-amino-5-hydroxybenzoic acid (AHBA) (Figure 4-1). Interestingly, sequence alignment of the binding pocket residues revealed that DHQ

synthases and aminoDHQ synthases retain high sequence conservation, only differing at position K197. In all of the reported aminoDHQ synthases, the lysine corresponding to position 197 in DHQ synthase has been altered to a conserved arginine residue (Table 4-1).

The next well-studied subfamily of SPCs is the more distantly related 2-deoxy-*scyllo*-inososse synthases required for the biosynthesis of aminoglycoside antibiotics (Figure 4-1). The substrates for DHQ synthase (DAHP) and 2-deoxy-*scyllo*-inososse synthase (glucose 6-phosphate) have noticeable structural differences, leading to the production of cyclized products with different stereochemistry and positioning of hydroxyl functionalities as well as the absence of the carboxylic acid moiety in the case of 2-deoxy-*scyllo*-inososse (Figure 4-1). Thus, it is not surprising that the clade of 2-deoxy-*scyllo*-inososse synthases is more evolutionarily divergent from the other related subclasses. However, sequence alignment of the predicted active site residues reveals that 2-deoxy-*scyllo*-inososse synthases retain high conservation with DHQ synthases (Table 4-1). Differences within the active site include the conversion of R264 and N268 to a glycine residue and an acidic glutamate, respectively (Table 4-1). These alterations are highly conserved within all of the available sequences for 2-deoxy-*scyllo*-inososse synthases creating a novel sequence signature for this subfamily. Further analysis of the recent crystal structure reported for BtrC should lend valuable insight into

the resulting structural differences in the active site pocket of 2-deoxy-*scyllo*-inososе synthases.[60]

The newest family of SPCs includes the 2-*epi*-5-*epi*-valiolone synthases required for the biosynthesis of C_7-cyclitols. Comparisons of the active site residues from the available 2-*epi*-5-*epi*-valiolone synthases, including the newly isolated CetA and PrlA (GenBank accession numbers EF120454 and EF120453, respectively) sequences from the cetoniacytone and pyralomicin producers, reveal striking dissimilarities with DHQ synthases (Table 4-1). One third of the active site residues are consistently altered in 2-*epi*-5-*epi*-valiolone synthases including H275, which is thought to play a critical role in enzyme catalysis. Amino acid residues corresponding to DHQ synthase K250, H275, and K356, are all highly conserved proline residues, and the basic N268 is a conserved aspartate residue in 2-*epi*-5-*epi*-valiolone synthases (Table 4-1). Furthermore, the AcbC active site residues differ at three additional positions, suggesting that there is some flexibility in active site residues within this subfamily (Table 4-1). Clearly, substantial differences in active site residues combined with the low substrate tolerance of recombinant 2-*epi*-5-*epi*-valiolone synthases, provides compelling evidence to suggest that the active site cleft is significantly altered in this subfamily of SPCs.

4.2.4. Identification of a putative novel set of SPCs

Interestingly, results from the phylogenetic analysis have also revealed a novel class of homologous enzymes that were previously annotated as hypothetical DHQ synthases or unknown proteins from genome sequencing projects.[160, 161] It is clear from the phylogenetic analysis that these samples form a separate clade that is distinct from the other families of SPCs (Fig. 4-3). Since all the other clades identified in the phylogenetic analysis have unique substrates and products, we hypothesize that this new clade constitutes a novel class of SPCs that has previously been misidentified as DHQ synthase. This new clade is represented primarily by fungal and cyanobacterial counterparts. Current studies are underway in our laboratory to investigate the role for this new clade of SPCs in the production of a novel class of cyclitol-containing compounds.

4.2.5. The involvement of 2-epi-*5*-epi-*valiolone synthase in BE-40644 biosynthesis*

Phylogenetic analysis of the CetA sequence also revealed a putative 2-*epi*-5-*epi*-valiolone synthase from the BE-40644 biosynthetic gene cluster.[87] BE-40644, a product of terpenoid biosynthesis is a unique compound in that it is modified by the incorporation of a cyclitol-derived ring structure into the final product. Based on the analysis of the BE-40644 structure, the biosynthetic origin of the cyclitol ring structure is not obvious. However, BLAST analysis of the BE-40644 biosynthetic genes has enabled the prediction of a putative mechanism for cyclitol biosynthesis. *BE-orf9* encodes a putative 2-*epi*-5-*epi*-valiolone synthase that shares

considerable homology with AcbC (57% identity/70% similarity), CetA (54% identity/65% similarity) and ValA (56% identity/68% similarity) from the acarbose, cetoniacytone, and validamycin pathways, respectively. With such high sequence similarity, we predict that BE-Orf9 will function in a similar manner with AcbC and ValA and cyclize sedoheptulose 7-phosphate to the 2-*epi*-5-*epi*-valiolone intermediate to initiate cyclitol biosynthesis.

4.2.6. Isolation of salQ *from the salbostatin producer*

The biosynthetic gene cluster of salbostatin in *Streptomyces albus* ATCC 21838 has been identified by our collaborator in Myongji University from Korea. The complete nucleotide sequence of the cluster revealed 32 *orfs* including the *acbC* homolog, *salQ*. The SalQ protein shows 52% identity with ValA and 72% identity with AcbC, respectively. The SalQ protein also shows 36% identity to the AroB protein from *Emericella nidulans*.

4.2.7. Biochemical characterization of CetA, PrlA, BE-Orf9, and SalQ

To establish the involvement of 2-*epi*-5-*epi*-valiolone as an intermediate in cetoniacytone, BE-40644, salbostatin, and pyralomicin production, the activity of recombinant CetA, BE-Orf9, SalQ, and PrlA was assessed. A 1.09 Kb DNA fragment containing the *cetA* gene was amplified from pCet26 by PCR. The *cetA* gene was subcloned as a 6X-His-fusion protein in the pRSETB vector (Invitrogen) and recombinantly expressed in the *E. coli* strain BL21(DE3)pLysS (Stratagene)

under control of the T7 promoter. Expression of *cetA* was induced by IPTG and yielded large quantities of 46-kDa soluble polyhistidine-tagged protein (Figure 4-4). When the crude extracts from induced cells were incubated with sedoheptulose 7-phosphate, a rapid conversion of the substrate to the product was observed (Figure 4-4). In contrast, boiled extracts containing CetA and extracts without CetA were unable to form the cyclized product (Figure 4-4). The product demonstrated the same Rf value as authentic 2-*epi*-5-*epi*-valiolone on thin-layer chromatography (Figure 4-4). Affinity purification of the recombinant protein with Ni-NTA agarose gave a protein that was >85% pure (Figure 4-4). The purified CetA was also incubated with sedoheptulose 7-phosphate and the product was extracted with methanol and converted to its trimethylsilylated derivative, then analyzed by GC-MS. Comparison of the GC-MS profile with a deuterium labeled synthetic standard, tetra-TMS-2-*epi*-5-*epi*-[6,6-^{2}H]valiolone, revealed an elution profile and mass fragmentation pattern consistent with 2-*epi*-5-*epi*-valiolone (Figure 4-5). The only variation in the MS fragmentation pattern between the standard and the CetA generated reaction product was due to the standard fragments containing the deuterated label, i.e., fragment 276 in the reaction product was observed as fragment 278 in the standard. Similarly, the overall mass differed by 2 atomic mass units (the reaction product in the CetA sample had an observed molecular ion of 480 whereas the standard had 482) was also due to the deuterated standard.

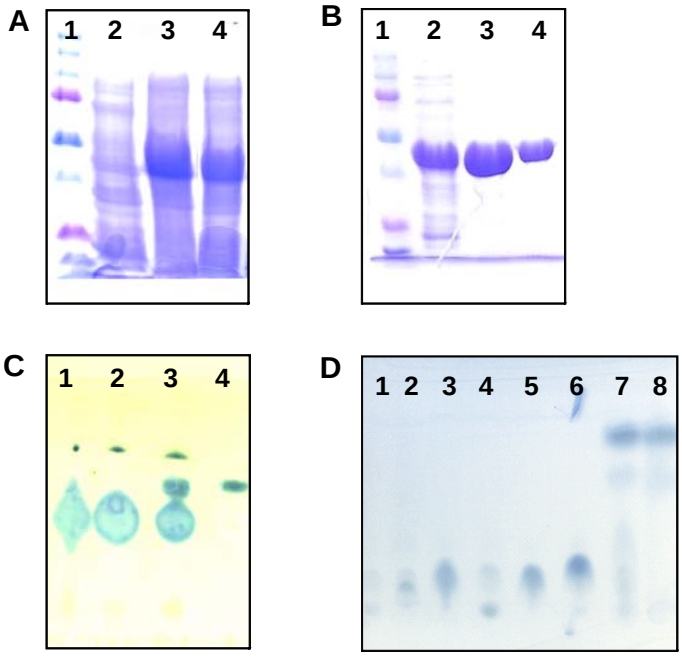


Figure 4-4. CetA overexpression and TLC analysis of CetA reactions. (A) CetA overexpression with *E. coli* expression system. Lane 1 = protein marker, lane 2 = CetA without IPTG induction, lane 3 = total proteins of CetA after 0.2 mM IPTG induction, lane 4 = soluble fraction of CetA after 0.2 mM IPTG induction. (B) Purification of CetA, lane 1 = protein marker, lane 2 = soluble fraction of CetA, lane 3 = nickel-agarose-purified CetA, lane 4 = additional Sephacryl S300 purified CetA. (C) TLC analysis of CetA enzyme activity, lane 1 = sedoheptulose 7-phosphate with boiled CetA, lane 2 = sedoheptulose 7-phosphate with cell free extract of *E. coli* harboring empty vector, lane 3 = sedoheptulose 7-phosphate with CetA, lane 4 = 2-*epi*-5-*epi*-valiolone standard. (D) CetA enzyme reaction mixtures contained 5 mM of the following substrate, lane 1 = aminoDAHP, lane 2 = DAHP, lane 3 = glucose 6-phosphate, lane 4 = fructose 6-phosphate, lane 5 = mannose 6-phosphate, lane 6 = ribose 5-phosphate, lane 7 = sedoheptulose 7-phosphate, and lane 8 = 2-*epi*-5-*epi*-valiolone standard.

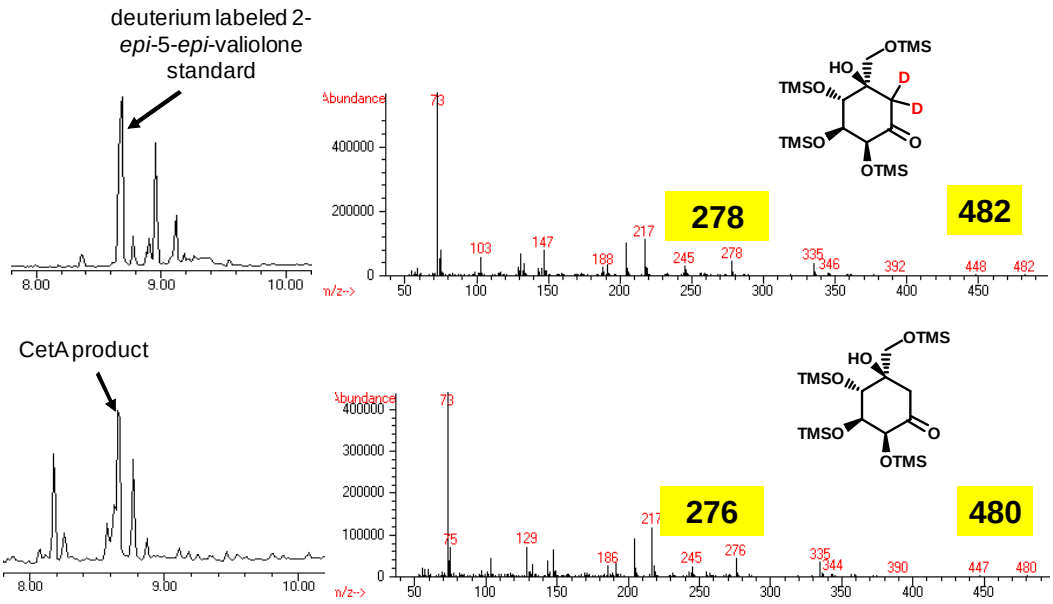


Figure 4-5. GC-MS profiles of silylated product of CetA reaction. Left panel are GC traces of deuterated silylated 2-*epi*-5-*epi*-valiolone standard (top) and silylated CetA reaction product (bottom). Right panel are the MS fragmentation patterns of the major peak eluting at 8.65-8.75 minutes.

To determine the substrate specificity of CetA, purified CetA was incubated with a number of different potential substrates, including DAHP, aminoDAHP, glucose 6-phosphate, fructose 6-phosphate, mannose 6-phosphate and ribose 5-phosphate. Analysis of reaction products using TLC revealed that purified CetA demonstrates restricted substrate specificity and only converts sedoheptulose 7-phosphate to its corresponding cyclized product (Figure 4-4).

For PrlA, a 1.23 Kb DNA fragment containing the *prlA* gene was amplified by PCR, cloned into the expression vector pRSET-B and transferred into *E. coli* BL21Gold(DE3)pLysS. Expression of *prlA* resulted in the production of a large

quantity of a 47 kDa soluble polyhistidine-tagged protein (Fig. 4-6). Enzyme analysis showed that the same as CetA, PrlA can convert sedoheptulose 7-phosphate to 2-*epi*-5-*epi*-valiolone (Fig. 4-6). The molecular mass and fragmentation pattern of the silylated derivative of the enzyme product [*m/z* 480 (M^+), 335, 276, 245, 217, 186, 147] was consistent with that of the authentic sample 2-*epi*-5-*epi*-[6-$^2H_2$]valiolone, which was synthesized in its dideuterated form [*m/z* 482 (M^+), 335, 278, 245, 217, 188, 147]. The two atomic mass unit higher molecular weight for fragments 482, 278, and 188 in the standard reflected the presence of two deuterium atoms in those fragments (Fig. 4-6). As with CetA, incubations of any alternate sugar phosphate substrate with purified PrlA did not show any appreciable levels of activity, suggesting that PrlA has restricted substrate specificity (Fig. 4-6).

Similarly, *BE-orf9* was subcloned into the plasmid pRSET-B and expressed in *E. coli* BL21Gold(DE3)pLysS cells (Fig. 4-7). Cell-free extracts housing the *BE-orf9* construct demonstrated 2-*epi*-5-*epi*-valiolone synthase activity (Fig 4-7). Further analysis of the purified recombinant protein revealed that BE-Orf9 is also highly selective and can only convert sedoheptulose 7-phosphate to the cyclized product (Fig. 4-7).

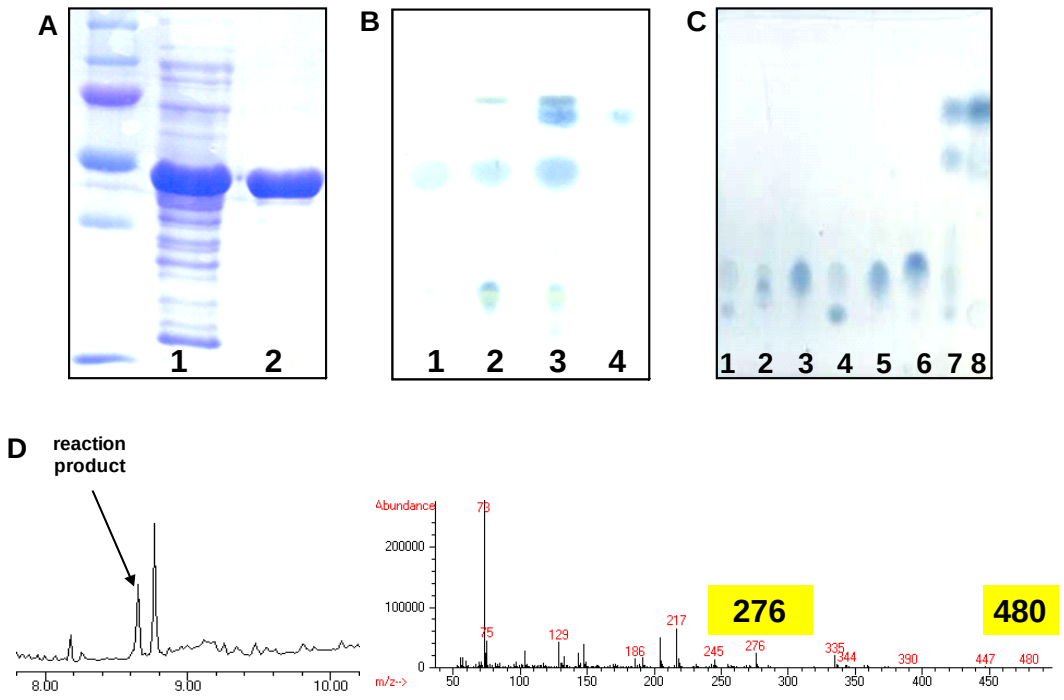


Figure 4-6. Recombinant expression and characterization of PrlA. (A) Heterologously expressed and purified PrlA. lane 1 = soluble fraction of PrlA, lane 2 = Ni-NTA agarose-purified PrlA. (B) TLC analysis of PrlA enzyme activity. lane 1 = boiled PrlA with sedoheptulose 7-phosphate, lane 2 = cell free extract of *E. coli* harboring empty vector with sedoheptulose 7-phosphate, lane 3 = PrlA with sedoheptulose 7-phosphate, lane 4 = 2-*epi*-5-*epi*-valiolone standard. (C) TLC analysis of the substrate specificity of PrlA. Enzyme reaction mixtures contained 5 mM of the following substrates, lane 1 = aminoDAHP, lane 2 = DAHP, lane 3 = glucose 6-phosphate, lane 4 = fructose 6-phosphate, lane 5 = mannose 6-phosphate, lane 6 = ribose 5-phosphate, lane 7 = sedoheptulose 7-phosphate, and lane 8 = 2-*epi*-5-*epi*-valiolone standard. (D) GC-MS profiles of silylated PrlA reaction product. Left panel shows the GC trace and right panel shows the MS fragmentation pattern of the major peak eluting at 8.65–8.75 minutes.

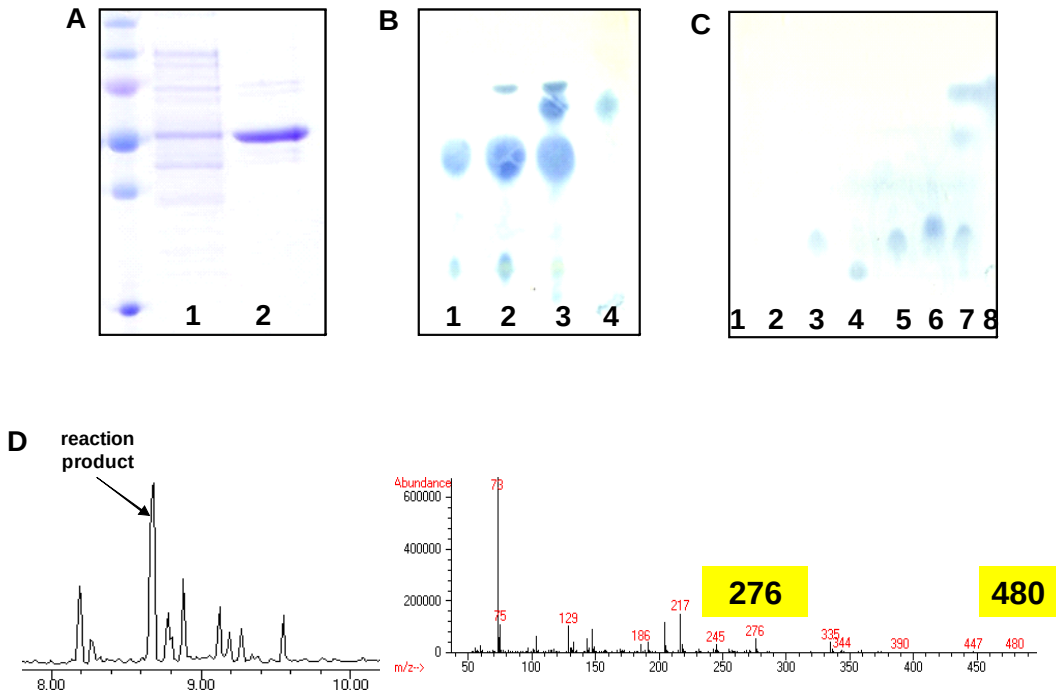


Figure 4-7. Recombinant expression and characterization of BE-Orf9. (A) Heterologously expressed and purified BE-Orf9. lane 1 = soluble fraction of BE-Orf9, lane 2 = Ni-NTA agarose-purified BE-Orf9. (B) TLC analysis of BE-Orf9 enzyme activity. lane 1 = boiled BE-Orf9 with sedoheptulose 7-phosphate, lane 2 = cell free extract of *E. coli* harboring empty vector with sedoheptulose 7-phosphate, lane 3 = BE-Orf9 with sedoheptulose 7-phosphate, lane 4 = 2-*epi*-5-*epi*-valiolone standard. (C) TLC analysis of the substrate specificity of BE-Orf9. Enzyme reaction mixtures contained 5 mM of the following substrates, lane 1 = aminoDAHP, lane 2 = DAHP, lane 3 = glucose 6-phosphate, lane 4 = fructose 6-phosphate, lane 5 = mannose 6-phosphate, lane 6 = ribose 5-phosphate, lane 7 = sedoheptulose 7-phosphate, and lane 8 = 2-*epi*-5-*epi*-valiolone standard. (D) GC-MS profiles of silylated BE-orf9 reaction product. Left panel shows the GC trace and right panel shows the MS fragmentation pattern of the major peak eluting at 8.65–8.75 minutes.

The *salQ* gene was amplified by PCR using Supercos 37-20 as template and primers SQF and SQR, and the product was subcloned into the expression vector pET-30a (Novagen) and expressed in *E. coli* BL21Gold(DE3)pLysS cells. Enzyme characterization was carried out in the same way as for PrlA and BE-Orf9. Similar

to the other 2-*epi*-5-*epi*-valiolone synthases, SalQ has 2-*epi*-5-*epi*-valiolone synthase activity and shows restricted substrate selectivity (Figure 4-8).

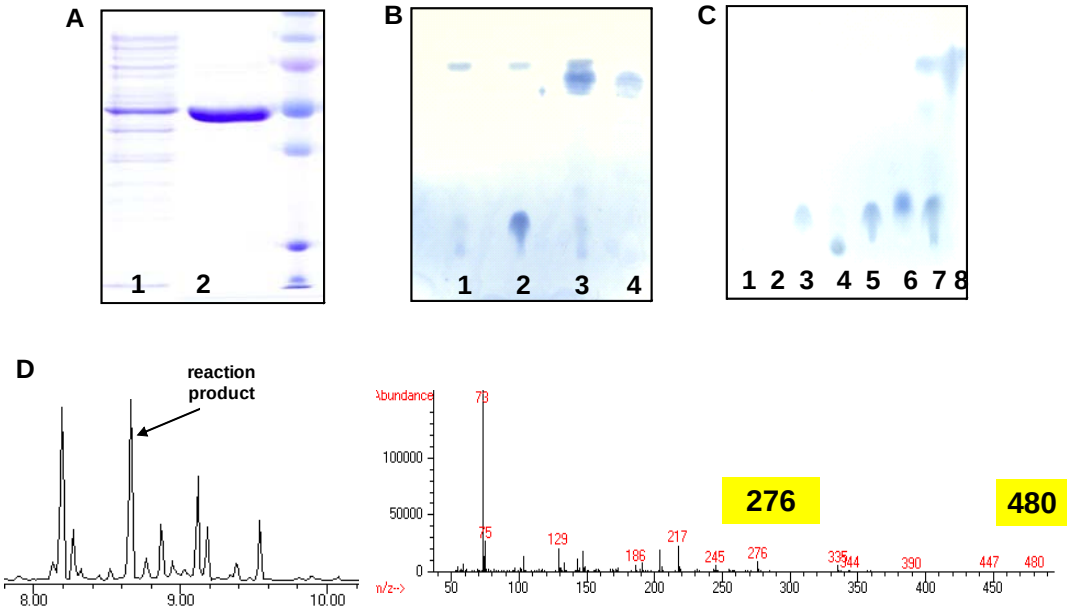


Figure 4-8. Recombinant expression and characterization of SalQ. (A) Heterologously expressed and purified SalQ. lane 1 = soluble fraction of SalQ, lane 2 = Ni-NTA agarose-purified SalQ. (B) TLC analysis of SalQ enzyme activity. lane 1 = boiled SalQ with sedoheptulose 7-phosphate, lane 2 = cell free extract of *E. coli* harboring empty vector with sedoheptulose 7-phosphate, lane 3 = SalQ with sedoheptulose 7-phosphate, lane 4 = 2-*epi*-5-*epi*-valiolone standard. (C) TLC analysis of the substrate specificity of SalQ. Enzyme reaction mixtures contained 5 mM of the following substrates, lane 1 = aminoDAHP, lane 2 = DAHP, lane 3 = glucose 6-phosphate, lane 4 = fructose 6-phosphate, lane 5 = mannose 6-phosphate, lane 6 = ribose 5-phosphate, lane 7 = sedoheptulose 7-phosphate, and lane 8 = 2-*epi*-5-*epi*-valiolone standard. (D) GC-MS profiles of silylated SalQ reaction product. Left panel shows the GC trace and right panel shows the MS fragmentation pattern of the major peak eluting at 8.65–8.75 minutes.

Results from these studies demonstrate that the primary sequence of SPCs represents a distinct fingerprint that is characteristic of the enzyme's catalytic activity and can be used as a tool for the directed isolation of biosynthetic gene

clusters and in evaluating the metabolic potential of previously uncharacterized organisms.[162]

Chapter 5

Characterization of the 2-*epi*-5-*epi*-Valiolone Epimerase (CetB), a New Member of the Vicinal Oxygen Chelate (VOC) Superfamily

5.1. Introduction

The VOC superfamily includes a set of structurally related proteins that are able to catalyze a large range of divalent metal ion-dependent reactions. Although the enzymes of this family use a variety of different divalent metal ions in catalysis (Co^{2+}, Zn^{2+}, Mn^{2+}, Ni^{2+}, and Fe^{2+}), there is one mechanistic feature common to all of the reactions, the chelation of vicinal oxygen ligands on a substrate or intermediate to promote a reaction. Members of this family include the Fe^{2+}-dependent extradiol dioxygenase (DHBD),[163-168] the bleomycin resistance protein (BRP),[169-172] the Co^{2+}-dependent methylmalonyl-CoA epimerase (MMCE),[173-177] the Zn^{2+}-dependent glyoxalase (GLO),[178-183] and the Mn^{2+}-dependent fosfomycin resistance protein (FOS)[184, 185] (Figure 5-1). DHBD exploits non-heme Fe^{2+} or Mn^{2+} (rarely) to catalyze the opening of aromatic rings which is a critical step in the microbial degradation of aromatic compounds. BRP binds and sequesters bleomycin and related compounds without degrading or transforming them. MMCE catalyzes the conversion of (2*R*)-methylmalonyl-CoA to (2*S*)-methylmalonyl-CoA, the substrate for the methylmalonyl-CoA mutase. GLO is also known as lactoylglutathione lyase and utilizes Zn^{2+} to catalyze the formation of an acylthioester during the first of two steps in the glutathione-dependent

inactivation of toxic methylglyoxal. The FOS family, FosA catalyzes the addition of GSH to fosfomycin, FosB catalyzes the addition of L-CySH to fosfomycin, and FosX catalyzes the hydration of fosfomycin. All of these enzymes belong to the VOC superfamily of enzymes.

Although BRP, DHBD, MMCE, FOS and GLO possess low sequence similarity and notable differences in function, metal preference, and oligomeric state, the overall three-dimensional structures of the MMCE, BRP, GLO, FOS dimers and the DHBD monomer are remarkably similar. All members in this family possess a characteristic common structural scaffold, comprised of βαββββ modules, two of βαββββ modules usually forming one metal-binding/active site (two modules protein). However, assembly of the domains occurs in several different ways, suggesting that the evolution of these proteins probably involved multiple gene duplication, gene fusion and domain swapping events. Structural comparisons, as well as phylogenetic analyses, support a general mechanism of metalloprotein evolution that exploits the symmetry of a βαββββ module to create a metal binding site followed by the relaxation of symmetry, as enabled by gene duplication, to establish and refine specific functions. A key feature in this scheme is the creation of a metal-binding site in a noncovalent βαββββ oligomer (Figure 5-2). The hypothesis of ancestral metal binding is challenged by the structure of BRP, which lacks a metal binding site by comparisons with GLO and DHBD. Rather than

binding a metal, it is likely that BRP binds and sequesters bleomycin in the same active-site cavity.

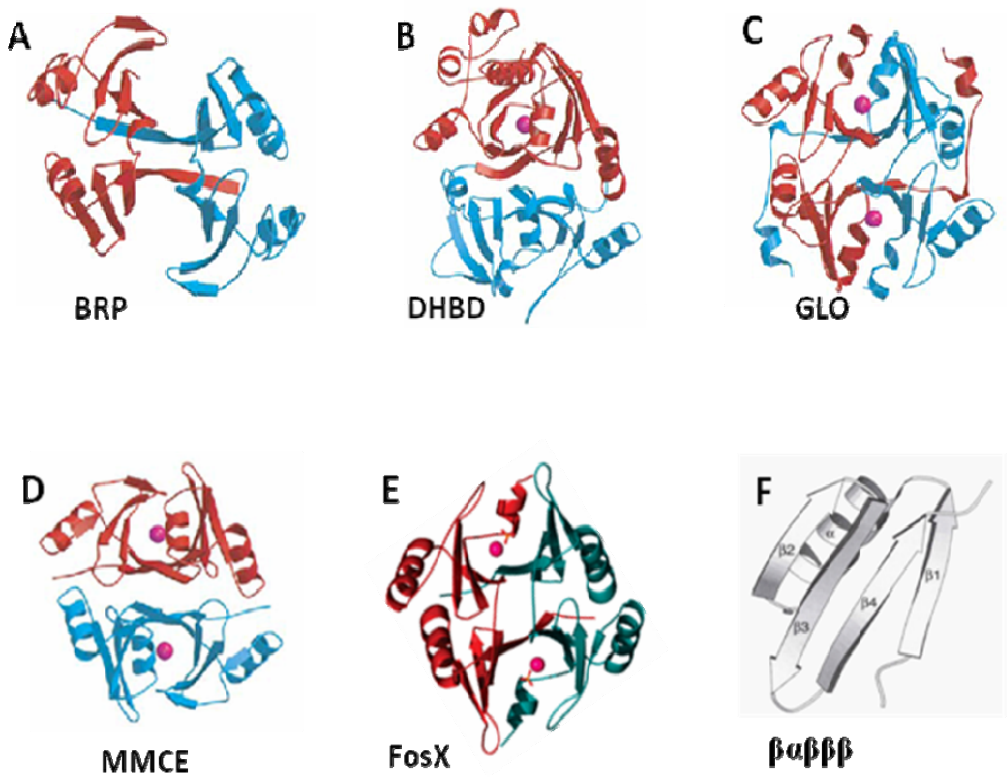


Figure 5-1. Reported crystal structures of various members of the VOC superfamily[173, 185, 186]. (A) Bleomycin resistance protein (BRP dimer). (B) Extradiol dioxygenases (DHBD monomer). (C) Glyoxalase (GLO dimer). (D) Methylmalonyl-CoA epimerase (MMCE dimer). (E) FosX dimer. (F) βαβββ module.

In BRP and GLO, the two βαβββ modules within each monomer are involved in a back-to-back contact; in MMCE, the two βαβββ modules have an edge-to-edge contact (Figure 5-2). The difference between the arrangement of modules in BRP and GLO versus that in MMCE is readily embraced by the concepts of three-dimensional domain swapping. In this case, swapped interactions

are observed at the level of the βαββββ modules. The structural similarities of the modules suggest that the gene for a two-module structure arose from the ancestral βαββββ gene by duplication and fusion. These events might have occurred independently for any subset of the known proteins. However, phylogenic analyses strongly support the notion that all two module units in the VOC superfamily share a common two-module ancestor.[186] Some of the two βαββββ modules proteins were further evolved by gene fusion event to give rise to the four modules monomer polypeptide such as some of the GLO[187] and most of dioxygenases (Figure 5-2).

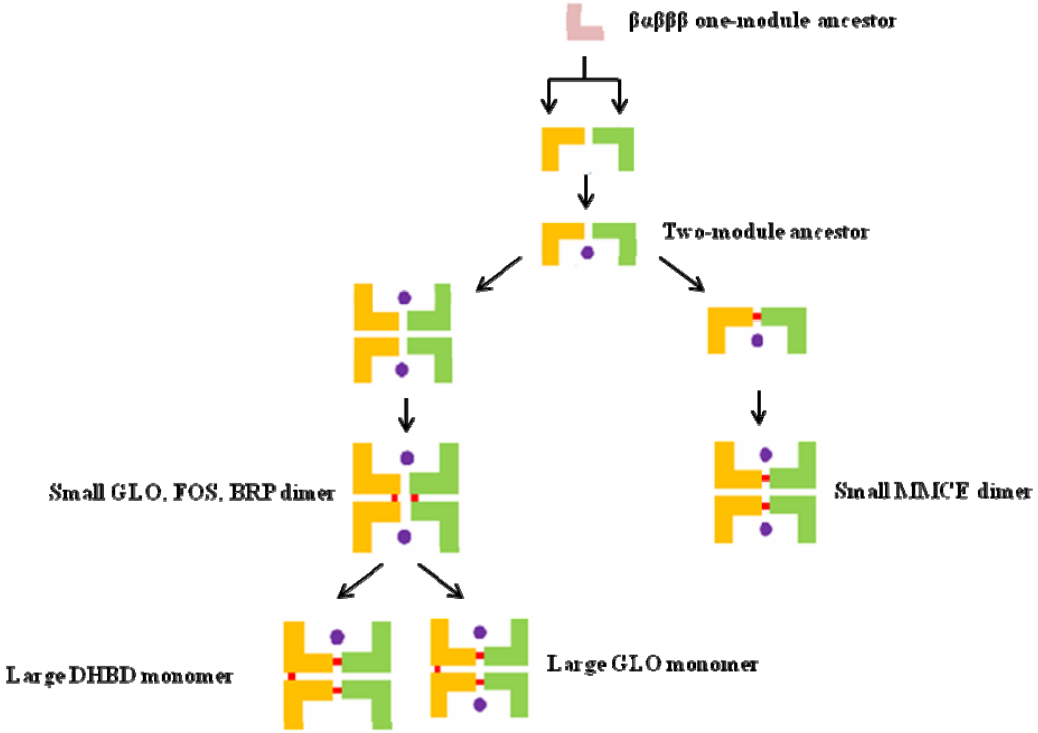


Figure 5-2. A potential evolutionary pathway leading to genes that code for the known VOC family. Each filled L shape symbolizes the gene for a βαββββ module.

Interestingly, within the cetoniacytone biosynthetic gene cluster we identified a putative glyoxalase gene (*cetB*), which may be involved in the formation of this unique natural product. Sequence analysis also suggests that CetB is similar to ValD, another homolog of glyoxalase in the validamycin pathway. However, CetB (183 amino acid) is only half of the size of ValD (451 amino acid), and amino acid sequence alignment of the two proteins revealed that ValD contains two similar domains, both of which have high identity to CetB. The involvement of such enzymes in cetoniacytone and validamycin biosynthesis could be predicted to occur early in the pathway. As based on our previous feeding experiments, the two pathways only share the first two biosynthetic steps. The first step, which is the cyclization of sedoheptulose 7-phosphate to 2-*epi*-5-*epi*-valiolone, has been demonstrated to be catalyzed by the 2-*epi*-5-*epi*-valiolone synthases, *e.g.*, ValA and CetA. The second step, which is a conversion of 2-*epi*-5-*epi*-valiolone to 5-*epi*-valiolone, has thus far not been clearly defined. The fact that CetB and ValD proteins belong to the VOC superfamily together with GLO and MMCE, it is likely that these proteins would catalyze the epimerization of 2-*epi*-5-*epi*-valiolone to 5-*epi*-valiolone. Therefore, we decided to investigate the catalytic activity of CetB as described below.

5.2. Results and discussion

5.2.1. Characterization of CetB activity using recombinantly expressed protein

To confirm the activity of CetB as a 2-*epi*-5-*epi*-valiolone epimerase, the *cetB* gene was subcloned and overexpressed in *E. coli* BL21(DE3)pLysS. Expression of *cetB* yielded large quantities of a 24-kDa soluble polyhistidine-tagged protein (Figure 5-3). As the potential substrate for CetB, 2-*epi*-5-*epi*-valiolone, was not readily available, characterization of CetB was carried out using a coupled enzyme assay using sedoheptulose 7-phosphate as substrate and the 2-*epi*-5-*epi*-valiolone synthase (CetA) as the first enzyme. The product 2-*epi*-5-*epi*-valiolone is expected to be converted to 5-*epi*-valiolone by CetB. When sedoheptulose 7-phosphate was incubated with CetA and then with the cell extracts containing CetB, the product gave the same Rf value as 5-*epi*-valiolone on silica gel TLC plates (Figure 5-3). GC-MS analysis of the silylated product of CetB reaction showed a retention time and mass fragmentation pattern consistent with that of the synthetic standard, tetra-TMS-5-*epi*-valiolone, suggesting that CetB is a 2-*epi*-5-*epi*-valiolone epimerase (Figure 5-3 C and D). In contrast, extracts from the empty vector pRSETB incubated with sedoheptulose 7-phosphate plus CetA gave only the 2-*epi*-5-*epi*-valiolone product (Figure 5-3 B). Interestingly, the boiled CetB can also give the epimerization product, which suggests that CetB is a heat-stable protein. This heat-stable property has been reported within the VOC family, such as MMCE, BRP, and GLO.[173]

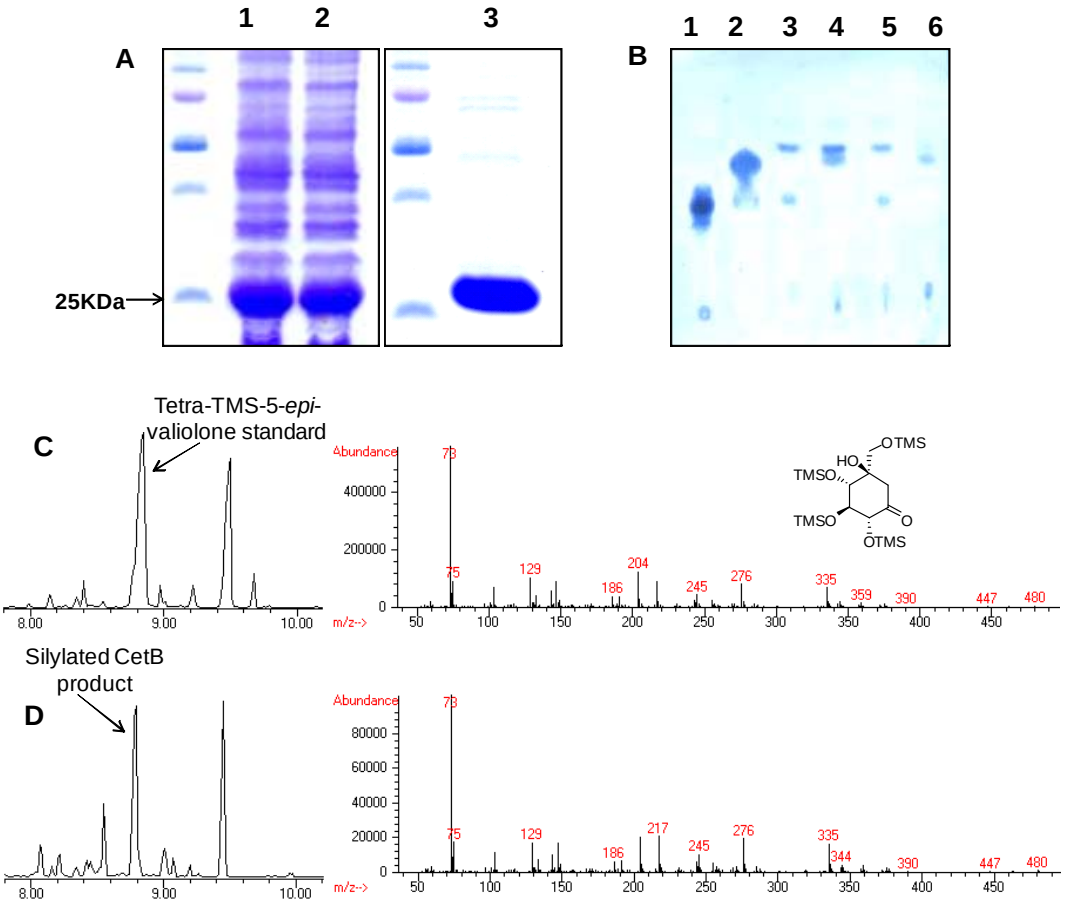


Figure 5-3. CetB production and enzymatic characterization. (A) SDS-PAGE of CetB expression. lane 1 = total proteins of CetB induced by 0.2 mM IPTG, lane 2 = soluble fraction of CetB induced by 0.2 mM IPTG, lane 3 = purified CetB by Ni-NTA agarose. (B) TLC analysis of CetB reaction. lane 1 = 5-*epi*-valiolone standard, lane 2 = 2-*epi*-5-*epi*-valiolone standard, lane 3 = boiled CetB with CetA reaction product, lane 4 = cell free extract of *E. coli* harboring empty vector with CetA reaction product, lane 5 = CetB with CetA reaction product, lane 6 = CetA reaction product. (C) GC-MS profile of silylated 5-*epi*-valiolone standard. (D) GC-MS profile of silylated product of CetB reaction. Left panel show the GC trace and right panel show the MS fragmentation pattern of the major peak eluting at 8.75–8.85 minutes.

To evaluated the substrate specificity of CetB, a number of unnatural substrates, that have chemical structures similar to 2-*epi*-5-*epi*-valiolone, *e.g.*, 1L-*epi*-2-inosose, D-(+)-gluconic acid delta-lactone, mannose, shikimic acid,

dehydroshikimic acid, and aminodehydroshikimic acid were tested. Analysis of the reaction products by TLC revealed that CetB demonstrates restricted substrate specificity and only epimerizes 2-*epi*-5-*epi*-valiolone to 5-*epi*-valiolone (Figure 5-4).

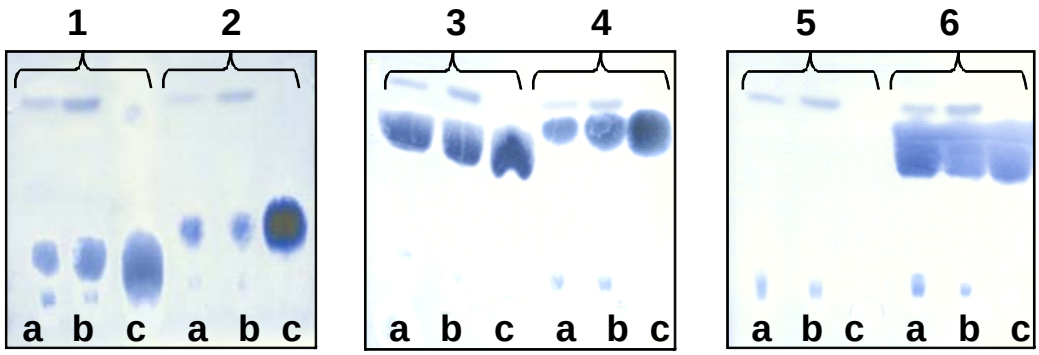


Figure 5-4. TLC analysis of reaction products of CetB with different substrates. CetB was incubated with six different substrates. 1 = 1L-*epi*-2-inosose; 2 = D-(+)-Gluconic acid delta-lactone; 3 = mannose; 4 = shikimic acid; 5 = dehydroshikimic acid; and 6 = aminodehydroshikimic acid. For each reaction, three spots (a = substrate with purified CetB, b = substrate with cell free extract of *E. coli* harboring empty vector, c = substrate only) were spotted on the TLC plate.

5.2.2. CetB metal binding analysis

Despite different metal preference, structural comparisons of MMCE, DHBD, and GLO showed that they have almost identical metal binding sites. MMCE is activated by divalent metal ions, with the most effective being Co^{2+} and Mn^{2+}, through four conserved metal binding sites (H12, Q65, H91, E141). Human GLO is a Zn^{2+} dependent metalloprotein and contains four conserved metal binding sites (Q33, E99, H126, E172). DHBD has four conserved metal binding sites (H146, A198, H210, E260) with the preference being Fe^{2+}. FosA is a Mn^{2+}

dependent enzyme containing four conserved metal binding sites (H7, T9, H67, E113). Sequence comparison of human GLO with CetB and ValD showed that both CetB and ValD contain metal binding sites (Figure 5-5). Since the size of ValD is twice as large as CetB and GLO, ValD was divided into two fragments in the alignment. The N-terminal portion is called ValDN and the C-terminal portion is called ValDC. To determine the metal preference of CetB, we attempted to develop an assay system involving the two enzymes as described above. First, sedoheptulose 7-phosphate was incubated with CetA, in the presence of Co^{2+}, for 2 hours to provide 2-*epi*-5-*epi*-valiolone, and then EDTA or 1,10-phenanthroline was added to the reaction mixture to remove Co^{2+} and other background divalent ions. Subsequently, CetB was added into the reaction and the mixture was incubated for another 3 hours without any divalent metal ions. This experiment was intended to be a negative control and no CetB activity was expected. Metal ion preference of CetB would be then analyzed by the addition of individual divalent metal ions into the reaction mixture. Surprisingly, TLC analysis of the reaction products suggests that even without addition of metal ion, CetB retains strong epimerase activity (Figure 5-6). Given that CetB contains a metal binding site similar to other members of the VOC superfamily, we proposed that CetB may bind to its metal cofactor very tightly that EDTA and 1,10-phenanthroline may not be able to remove the cofactor from the enzyme. In order to overcome this problem, the four putative metal binding sites of CetB were individually altered as single mutations (H14G, E76G, H98G, E151G) and double mutations (H14G/E76G, H14G/H98G,

H14G/E151G, E76G/H98G, E76G/E151G, H98G/E151G) and their epimerase activity was characterized (Figure 5-7). The results showed that, in the absence of 1,10-phenanthroline, the activity of the single mutant variants were relatively similar to that of the wild type (Figure 5-8). However, when 1,10-phenanthroline was added into the reaction mixtures, the activity of the single mutant variants diminished (Figure 5-9). On the other hand, all of the double mutant variants had no detectable activity even without 1,10-phenanthroline (Figure 5-8). The results suggest that a single mutation at the metal-binding site of *cetB* only modestly distorts the metal-binding affinity of the protein, whereas a double mutation of *cetB* completely disrupted the metal-binding ability of the enzyme.

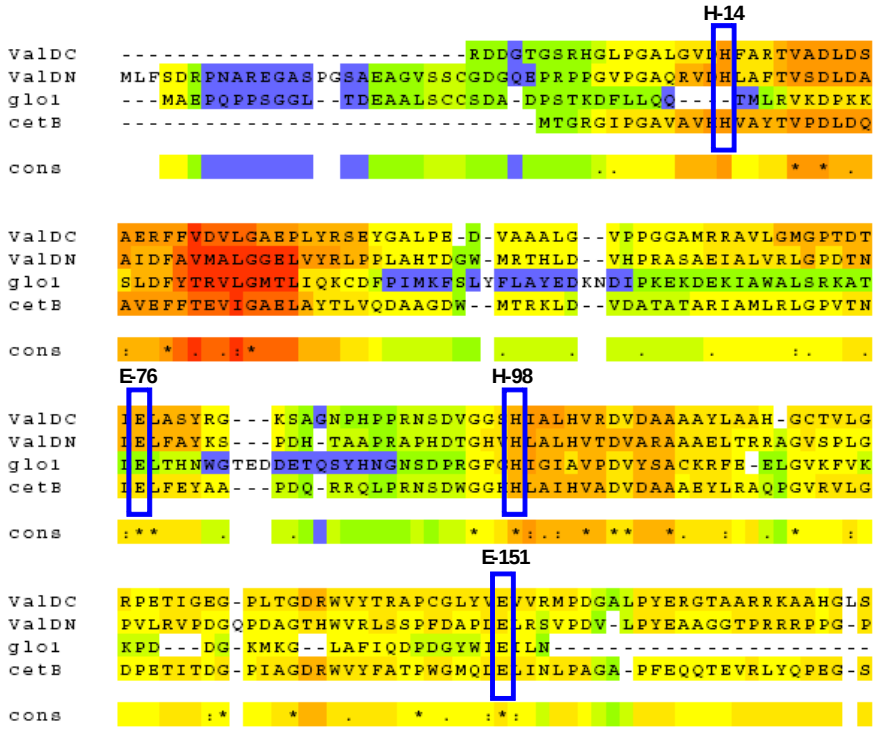


Figure 5-5. Sequence alignment of CetB, ValDN, ValDC, and human glyoxalase. The conserved metal binding sites are indicated in the frames.

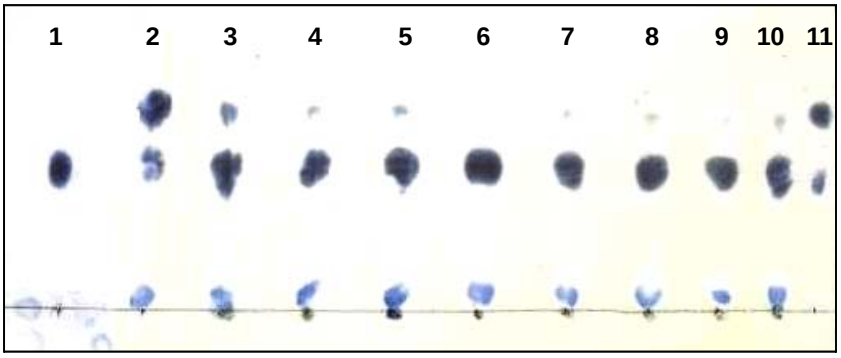


Figure 5-6. TLC analysis of reaction products of CetB enzyme with different metal ions. 1 = 5-*epi*-valiolone standard, 2 = CetA reaction product, 3 = CetB with chelated CetA reaction product, 4-10 = CetB with chelated CetA reaction produc plus different metal ion: $CoCl_2$ (4), $CuSO_4$ (5), $ZnCl_2$ (6), $MgCl_2$ (7), $MnCl_2$ (8), $FeSO_4$ (9), $NiCl_2$ (10), 11 = 2-*epi*-5-*epi*-valiolone standard.

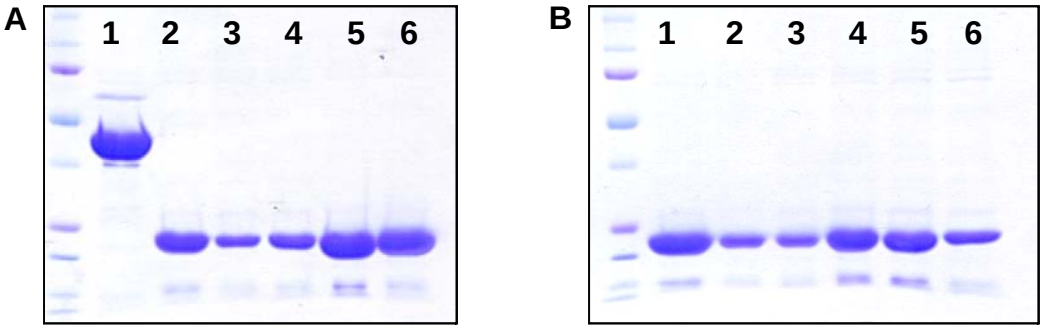


Figure 5-7. SDS-PAGE analysis of CetB mutants. (A) CetA and CetB single mutants. lane 1 = CetA, lane 2 = CetB wild type, lane 3 = H14G, lane 4 = H98G, lane 5 = E76G, lane 6 = E151G. (B) CetB double mutants. lane 1 = H14G/H98G, lane 2 = E76G/E151G, lane 3 = H14G/E76G, lane 4 = H14G/E151G, lane 5 = H98G/ E76G, lane 6 = H98G/E151G.

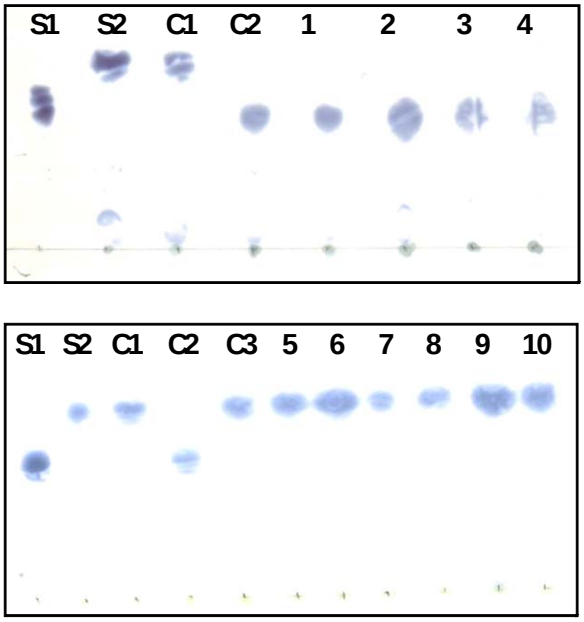


Figure 5-8. TLC analysis of reaction products of CetB mutants. S1 = 5-*epi*-valiolone standard, S2 = 2-*epi*-5-*epi*-valiolone standard, C1 = CetA reaction only, C2 = CetA reaction plus CetB, C3 = CetA reaction plus chelated CetB mutant H14G/H98G. 1-10 = CetA reaction plus CetB mutants: 1 = H14G, 2 = H98G, 3 = E76G, 4 = E151G, 5 = H14G/H98G, 6 = E76G/E151G, 7 = H14G/E76G, 8 = H14G/E151G, 9 = H98G/E76G, 10 = H98G/E151G.

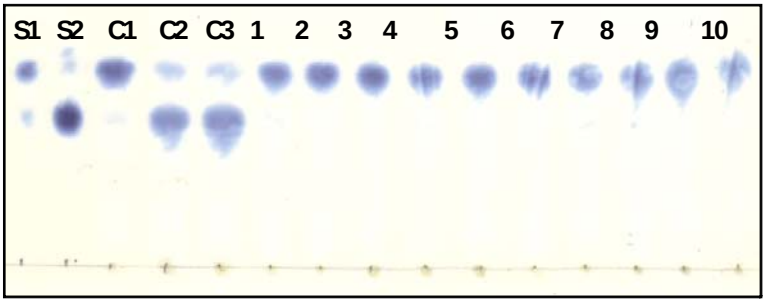


Figure 5-9. TLC analysis of reaction products of CetB mutants with 1,10-phenanthroline. S1 = 2-*epi*-5-*epi*-valiolone standard, S2 = 5-*epi*-valiolone standard, C1 = CetA reaction only, C2 = CetA reaction plus CetB without 1,10-phenanthroline, C3 = CetA reaction plus CetB with 1,10-phenanthroline. 1-10 = CetA reaction plus CetB mutants with 1,10-phenanthroline. 1 = H14G, 2 = H98G, 3 = E76G, 4 = E151G, 5 = H14G/H98G, 6 = E76G/E151G, 7 = H14G/E76G, 8 = H14G/E151G, 9 = H98G/E76G, 10 = H98G/E151G.

In addition to the diversity among the function and substrates/ligands of these proteins, MMCE, GLO, FOS, and DHBD all have metal binding site with different metal preference; one Fe^{2+} is bound to a DHBD monomer, two Zn^{2+} are bound to a GLO dimer, two Mn^{2+} are bound to a FOS dimer, and two Co^{2+} are bound to a MMCE dimer. The question is which metal ion is preferred for CetB? Because CetB has very tight metal-binding property, it is difficult to remove the metal ion from the enzyme and test its metal preference by adding external metal ions into the reaction mixture. However, the H98G mutant was found to have the same activity as the wild-type enzyme and yet the metal ion can be chelated by 1,10-phenanthroline. Therefore, the mutant H98G was selected for further analysis of its metal ion preference. After removing the residual metal ion in the CetA reaction with 1,10-phenanthroline, the H98G mutant protein, also pretreated with 1,10-phenanthroline, was added. Subsequently, a solution of metal ion, *e.g.*, Co^{2+}, Zn^{2+}, Mg^{2+}, Mn^{2+}, Fe^{2+}, Ni^{2+}, Cu^{2+} was individually added into the mixtures. TLC analysis showed that only Ni^{2+} can rescue the activity of H98G (Figure 5-10). While the replacement of the H98 residue with glycine may have caused a shift in metal ion preference of the enzyme, the fact that none of the other metal ions could in part rescue the activity of the enzyme suggests that Ni^{2+} is the most likely required metal ion for this subfamily of proteins.

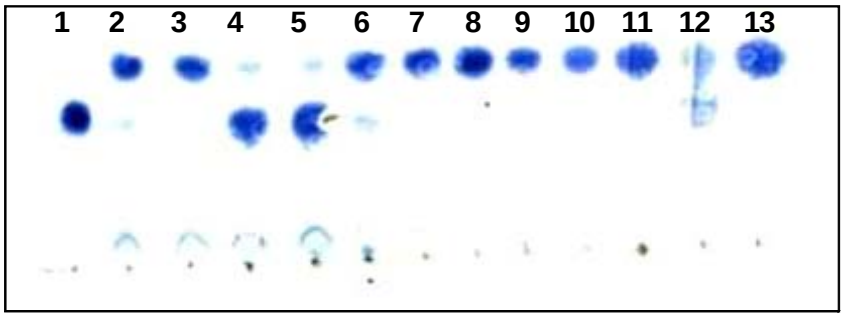


Figure 5-10. TLC analysis of CetB mutant proteins with different metal ions. 1 = 5-*epi*-valiolone standard, 2 = 2-*epi*-5-*epi*-valiolone standard, 3 = CetA reaction product, 4 = CetB with chelated CetA reaction product, 5 = CetB mutant H98G with CetA reaction product, 6 = CetB mutant H98G with chelated CetA reaction product, 7-13 = CetB H98G with chelated CetA reaction plus different metal ion, $CoCl_2$ (7), $ZnCl_2$ (8), $MgCl_2$ (9), $MnCl_2$ (10), $FeSO_4$ (11), $NiCl_2$ (12), $CuSO_4$ (13).

5.2.3. CetB modeling and dimerization analysis

All members of VOC superfamily contain either two or four βαβββ modules. The modeling of CetB using Phyre program[188, 189] showed that CetB also contains the common structural βαβββ scaffold (Figure 5-11 C). Crystal structure analysis of MMCE and GLO revealed that both enzymes form a dimer with the "back-to-back" packing in GLO and the "edge-to-edge" packing in MMCE. Also, it was reported that the four module yeast GLO is a monomeric enzyme with two active sites. Thus, we proposed that CetB would adopt a similar conformation as a homodimer. In order to confirm this hypothesis, the size of native CetB was determined by native-PAGE and high performance gel filtration chromatography. The native-PAGE indicated that CetB has a dimer structure (Figure 5-11 B). Based on the retention times of molecular mass standards within the size exclusion column, CetB eluted as a 48 KDa protein (Figure 5-11 D). This result confirmed that CetB forms a

homodimeric structure.

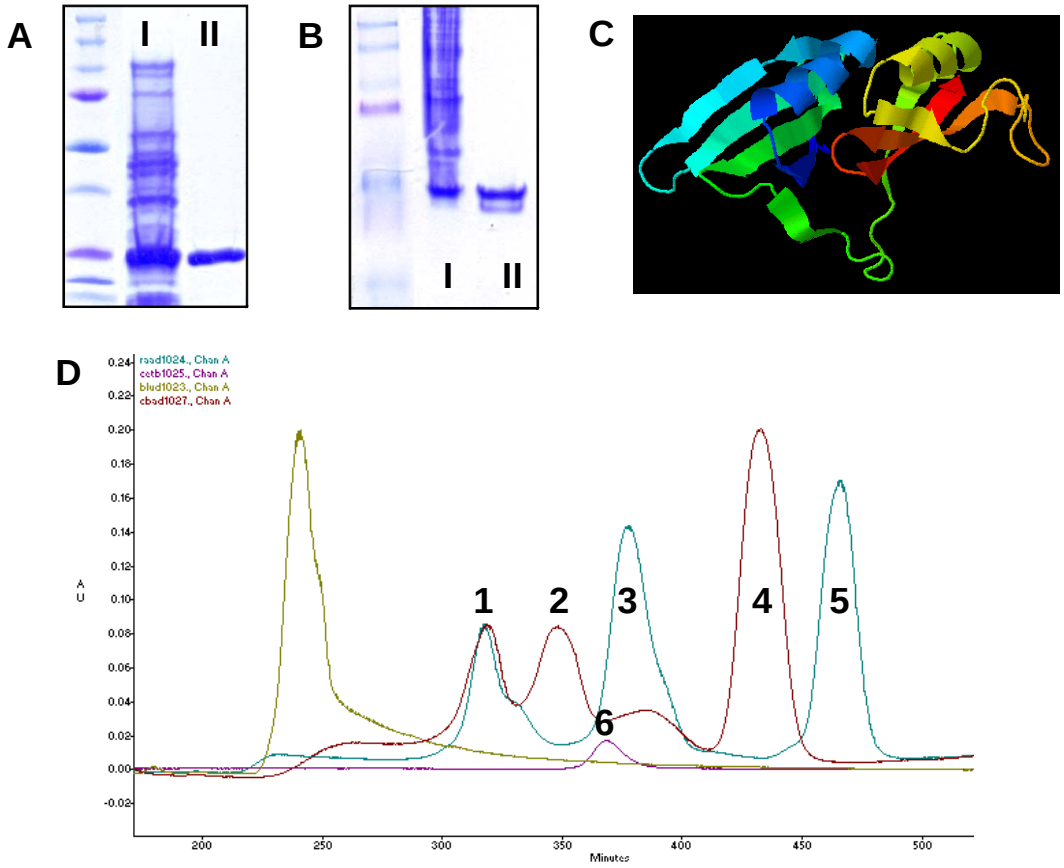


Figure 5-11. CetB modeling and dimerization analysis. (A) SDS-PAGE of CetB. (B) Native-PAGE of CetB. Lane I = soluble fraction of CetB, lane II = Ni-NTA agarose purified CetB. (C) Modeling of CetB using Phyre program. (D) High performance gel filtration chromatography of CetB. 1 = Alcohol dehydrogenase (150 KDa), 2 = BSA (66 KDa), 3 = Ovalbumin (45 KDa), 4 = Chymotrypsinogen A (25 KDa), 5 = Ribonuclease A (13.7 KDa), 6. CetB (48 KDa) (dimer).

The comparison of three-dimensional structure of MMCE, GLO, BRP, and several DHBD reveals that they are not apparently related to each other at the level of the amino acid sequence. However, comparisons against the current structural database, using DALI, showed that although these proteins have diverse functions,

they have a common fold in which βαβββ modules are combined in several different ways (Figure 5-2). Many, but not all, are enzymes in which the fold provides a platform for binding an essential metal ion.[173] Structural comparisons, as well as phylogenetic analyses, strongly indicate that the modern family (VOC family) of proteins represented by these enzymes arose through a rich evolutionary history that includes multiple gene duplication and fusion events. A significant early event is proposed to be the establishment of metal binding in an oligomeric ancestor prior to the first gene fusion.[186] With the characterization of CetB, here we first report that the 2-*epi*-5-*epi*-valiolone epimerase with Ni^{2+} binding site should be listed as the sixth class of the VOC family in addition to BRP, DHBD, GLO, FOS, and MMCE.

Chapter 6

Conclusion

The work presented in this thesis dissertation deals with the identification, sequencing, and functional analysis of the biosynthetic gene cluster of the antitumor agent cetoniacytone A. The results provide important insight into the mode of formation of this unique aminocyclitol natural product, and will contribute to future studies that aim to create new aminocyclitol analogs. In addition, a comparative analysis of CetA, a cyclase that catalyzes the first reaction step in cetoniacytone biosynthesis, with other related sugar phosphate cyclases (SPCs) revealed that this large family of cyclases can be divided into distinct subclasses that have unique and predictable activities that correspond to their primary amino acid sequences. This data has been used to further identify a new unknown SPC subfamily, which is widely distributed in fungi and cyanobacteria. The work also revealed that CetB is a new member of the VOC family with a 2-*epi*-5-*epi*-valiolone epimerase activity.

6.1. Identification of the cetoniacytone biosynthetic gene cluster

This part of the dissertation identified the cetoniacytone A biosynthetic gene cluster and proposed the biosynthetic pathway to cetoniacytone A. Analysis of a 26 Kb genome DNA of *Actinomyces* sp. strain Lu 9419 chromosome revealed 13

potential structural genes, which were predicted to be involved in the biosynthesis of cetoniacytone A, and four unrelated genes. They include genes demonstrated to encode 2-*epi*-5-*epi*-valiolone synthase (CetA), 2-*epi*-5-*epi*-valiolone epimerase (CetB), and predicted to encode arylamine *N*-acetyltransferase (CetD), drug resistance transporter (CetE), FAD/FMN-containing dehydrogenases (CetF), aminotransferase (CetH), pyranose oxidase (CetL), and aminotransferase (CetM). Deduced products of *cetC, cetI, cetJ,* and *cetK* are hypothetical cupin proteins with unknown function. However, these proteins may be involved in the epoxidation and dehydration steps required for the biosynthesis of cetoniacytone A as proposed in Figure 3-4.

Based on the gene cluster and the structure of cetoniacytone A, the biosynthetic pathway to cetoniacytone A was proposed as follows. The first step is a cyclization reaction by CetA in which sedoheptulose 7-phosphate is converted to 2-*epi*-5-*epi*-valiolone. CetB is then required to epimerize 2-*epi*-5-*epi* valiolone to 5-*epi*-valiolone. Further downstream the pathway may involve oxidation, amination, dehydration, and epoxidation steps. Finally, the transfer of the acetyl functionality would yield cetoniacytone A.

During the BLAST analysis of the cetoniacytone A biosynthetic gene cluster, a number of homologous genes from an unknown biosynthetic gene cluster in *Frankia alni* ACN14a genome were revealed. Detailed sequence analyses of the

genes around this putative gene cluster suggest that this cluster may be responsible for the biosynthesis of a compound related to cetoniacytone.

The significance of this part of work is two-fold: 1) These results revealed the biosynthetic gene cluster and proposed the biosynthetic pathway of cetoniacytone A, a novel potential antitumor agent. 2) During a comparison of the gene cluster of cetoniacytone A with the genome data base, a putative biosynthetic gene cluster for a compound related to cetoniacytone A has been predicted in *Frankia alni* ACN14a genome.

6.2. Enzymatic analysis of 2-epi-*5*-epi-*valione synthase and a comparative analysis of SPCs*

The second part of the dissertation takes a comparative approach to analyze the SPCs superfamily. SPCs from both primary and secondary metabolic pathways show high sequence similarity. Results from these studies demonstrate that the primary sequence of SPCs represents a distinct fingerprint that is characteristic of the enzyme's biological activity and can be used as a tool for directed isolation of biosynthetic gene clusters and in evaluating the metabolic potential of previously uncharacterized organisms.

CetA is highly similar to known 2-*epi*-5-*epi*-valiolone synthase, AcbC and ValA. This study confirmed that the function of CetA as a 2-*epi*-5-*epi*-valiolone

synthase and shows high substrate specificity to sedoheptulose 7-phosphate. Signature sequences from the 2-*epi*-5-*epi*-valiolone synthase class of SPCs were used to isolate an SPCs homolog, *prlA*, and finally isolate the pyralomicin biosynthetic gene cluster from the pyralomicin producer, *Nonomuraea spiralis*. The same strategy was also used by the Hong group in Korea to isolate the salQ gene from the salbostatin biosynthetic gene cluster. The gene product was also identified as a 2-*epi*-5-*epi*-valiolone synthase in this study.

Further phylogenetic analysis of SPC sequences previously annotated as DHQ synthases from genome sequencing projects has identified a new and unique clade of SPCs that may regulate the biosynthesis of a novel set of secondary metabolites. In addition, a putative 2-*epi*-5-*epi*-valiolone synthase gene from the BE-40644 biosynthetic gene cluster was identified and the biochemical activity of the gene product, BE-Orf9, was confirmed. Thus, the detailed comparisons of primary sequence of SPCs provide a valuable genetic resource that can be used to quickly identify and clone biosynthetic gene clusters from known cyclitol-producing organisms.

6.3. Enzymatic analysis of CetB

Enzymatic characterization of CetB confirmed that CetB catalyzes the second metabolic step in cetoniacytone biosynthesis mediating the epimerization of 2-*epi*-5-*epi*-valiolone to 5-*epi*-valiolone. Sequence analysis showed that CetB is a

new member of the VOC superfamily with high similarity to GLO. The data from high performance gel filtration chromatography support that, during evolution, CetB occurs as a two modules protein homodimer. Site directed mutagenesis and metal binding analysis showed that CetB is a Ni^{2+}-dependent protein with four metal binding sites. With the enzyme characterization of CetB, here we first report that the 2-*epi*-5-*epi*-valiolone epimerase should be listed as the sixth class of the VOC family in addition to BRP, DHBD, GLO, FOS, and MMCE.

Bibliography

1. Kittendorf, J. D.; Sherman, D. H., Developing tools for engineering hybrid polyketide synthetic pathways. *Curr Opin Biotechnol* **2006,** 17, (6), 597-605.

2. Floss, H. G., Combinatorial biosynthesis--potential and problems. *J Biotechnol* **2006,** 124, (1), 242-57.

3. Walsh, C. T., Combinatorial biosynthesis of antibiotics: challenges and opportunities. *Chembiochem* **2002,** 3, (2-3), 125-34.

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

5. Marsden, A. F.; Wilkinson, B.; Cortes, J.; Dunster, N. J.; Staunton, J.; Leadlay, P. F., Engineering broader specificity into an antibiotic-producing polyketide synthase. *Science* **1998,** 279, (5348), 199-202.

6. 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, (5), 1846-51.

7. Stutzman-Engwall, K.; Conlon, S.; Fedechko, R.; Kaczmarek, F.; McArthur, H.; Krebber, A.; Chen, Y.; Minshull, J.; Raillard, S. A.; Gustafsson, C., Engineering the aveC gene to enhance the ratio of doramectin to its CHC-B2 analogue produced in Streptomyces avermitilis. *Biotechnol Bioeng* **2003,** 82, (3), 359-69.

8. Arslanian, R. L.; Tang, L.; Blough, S.; Ma, W.; Qiu, R. G.; Katz, L.; Carney, J. R., A new cytotoxic epothilone from modified polyketide synthases heterologously expressed in Myxococcus xanthus. *J Nat Prod* **2002,** 65, (7), 1061-4.

9. Bohnert, H. J.; Nelson, D. E.; Jensen, R. G., Adaptations to Environmental Stresses. *Plant Cell* **1995,** 7, (7), 1099-1111.

10. Majumder, A. L.; Chatterjee, A.; Ghosh Dastidar, K.; Majee, M., Diversification and evolution of L-myo-inositol 1-phosphate synthase. *FEBS Lett* **2003,** 553, (1-2), 3-10.

11. Majumder, A. L.; Johnson, M. D.; Henry, S. A., 1L-myo-inositol-1-phosphate synthase. *Biochim Biophys Acta* **1997,** 1348, (1-2), 245-56.

12. Stein, A. J.; Geiger, J. H., The crystal structure and mechanism of 1-L-myo-inositol- 1-phosphate synthase. *J Biol Chem* **2002,** 277, (11), 9484-91.

13. Loewus, M. W.; Loewus, F. A.; Brillinger, G. U.; Otsuka, H.; Floss, H. G., Stereochemistry of the myo-inositol-1-phosphate synthase reaction. *J Biol Chem* **1980,** 255, (24), 11710-2.

14. Herrmann, K. M.; Weaver, L. M., The Shikimate Pathway. *Annu Rev Plant Physiol Plant Mol Biol* **1999,** 50, 473-503.

15. Dewick, P. M., The biosynthesis of shikimate metabolites. *Nat Prod Rep* **1990,** 7, (3), 165-89.

16. Herrmann, K. M., The Shikimate Pathway: Early Steps in the Biosynthesis of Aromatic Compounds. *Plant Cell* **1995,** 7, (7), 907-919.

17. Knaggs, A. R., The biosynthesis of shikimate metabolites. *Nat Prod Rep* **2001,** 18, (3), 334-55.

18. Floss, H. G., Natural products derived from unusual variants of the shikimate pathway. *Nat Prod Rep* **1997,** 14, (5), 433-52.

19. Hawkins, A. R., The complex Arom locus of Aspergillus nidulans. Evidence for multiple gene fusions and convergent evolution. *Curr Genet* **1987,** 11, (6-7), 491-8.

20. Charles, I. G.; Keyte, J. W.; Brammar, W. J.; Smith, M.; Hawkins, A. R., The isolation and nucleotide sequence of the complex AROM locus of Aspergillus nidulans. *Nucleic Acids Res* **1986,** 14, (5), 2201-13.

21. Moore, J. D.; Coggins, J. R.; Virden, R.; Hawkins, A. R., Efficient independent activity of a monomeric, monofunctional dehydroquinate synthase derived from the N-terminus of the pentafunctional AROM protein of Aspergillus nidulans. *Biochem J* **1994,** 301 (Pt 1), 297-304.

22. Hawkins, A. R.; Smith, M., Domain structure and interaction within the pentafunctional arom polypeptide. *Eur J Biochem* **1991,** 196, (3), 717-24.

23. Duncan, K.; Edwards, R. M.; Coggins, J. R., The pentafunctional arom enzyme of Saccharomyces cerevisiae is a mosaic of monofunctional domains. *Biochem J* **1987,** 246, (2), 375-86.

24. Graham, L. D.; Gillies, F. M.; Coggins, J. R., Over-expression of the yeast multifunctional arom protein. *Biochim Biophys Acta* **1993,** 1216, (3), 417-24.

25. Lambert, J. M.; Boocock, M. R.; Coggins, J. R., The 3-dehydroquinate synthase activity of the pentafunctional arom enzyme complex of Neurospora crassa is Zn2+-dependent. *Biochem J* **1985,** 226, (3), 817-29.

26. Banerji, S.; Wakefield, A. E.; Allen, A. G.; Maskell, D. J.; Peters, S. E.; Hopkin, J. M., The cloning and characterization of the arom gene of Pneumocystis carinii. *J Gen Microbiol* **1993,** 139, (12), 2901-14.

27. Nichols, C. E.; Hawkins, A. R.; Stammers, D. K., Structure of the 'open' form of Aspergillus nidulans 3-dehydroquinate synthase at 1.7 A resolution from crystals grown following enzyme turnover. *Acta Crystallogr D Biol Crystallogr* **2004,** 60, (Pt 5), 971-3.

28. Nichols, C. E.; Ren, J.; Leslie, K.; Dhaliwal, B.; Lockyer, M.; Charles, I.; Hawkins, A. R.; Stammers, D. K., Comparison of ligand-induced conformational changes and domain closure mechanisms, between prokaryotic and eukaryotic dehydroquinate synthases. *J Mol Biol* **2004,** 343, (3), 533-46.

29. Carpenter, E. P.; Hawkins, A. R.; Frost, J. W.; Brown, K. A., Structure of dehydroquinate synthase reveals an active site capable of multistep catalysis. *Nature* **1998,** 394, (6690), 299-302.

30. Nichols, C. E.; Ren, J.; Lamb, H. K.; Hawkins, A. R.; Stammers, D. K., Ligand-induced conformational changes and a mechanism for domain closure in Aspergillus nidulans dehydroquinate synthase. *J Mol Biol* **2003,** 327, (1), 129-44.

31. Nichols, C. E.; Ren, J.; Lamb, H.; Haldane, F.; Hawkins, A. R.; Stammers, D. K., Identification of many crystal forms of Aspergillus nidulans dehydroquinate synthase. *Acta Crystallogr D Biol Crystallogr* **2001,** 57, (Pt 2), 306-9.

32. Millar, G.; Coggins, J. R., The complete amino acid sequence of 3-dehydroquinate synthase of Escherichia coli K12. *FEBS Lett* **1986,** 200, (1), 11-7.

33. Frost, J. W.; Bender, J. L.; Kadonaga, J. T.; Knowles, J. R., Dehydroquinate synthase from Escherichia coli: purification, cloning, and construction of overproducers of the enzyme. *Biochemistry* **1984,** 23, (19), 4470-5.

34. Han, M. A.; Lee, H. S.; Cheon, C. I.; Min, K. H.; Lee, M. S., Cloning and analysis of the aroB gene encoding dehydroquinate synthase from Corynebacterium glutamicum. *Can J Microbiol* **1999,** 45, (10), 885-90.

35. Barten, R.; Meyer, T. F., Cloning and characterisation of the Neisseria gonorrhoeae aroB gene. *Mol Gen Genet* **1998,** 258, (1-2), 34-44.

36. Bereswill, S.; Fassbinder, F.; Volzing, C.; Haas, R.; Reuter, K.; Ficner, R.; Kist, M., Cloning and functional characterization of the genes encoding 3-dehydroquinate synthase (aroB) and tRNA-guanine transglycosylase (tgt) from Helicobacter pylori. *Med Microbiol Immunol* **1997,** 186, (2-3), 125-34.

37. Hasan, N.; Nester, E. W., Dehydroquinate synthase in Bacillus subtilis. An enzyme associated with chorismate synthase and flavin reductase. *J Biol Chem* **1978,** 253, (14), 4999-5004.

38. Bender, S. L.; Mehdi, S.; Knowles, J. R., Dehydroquinate synthase: the role of divalent metal cations and of nicotinamide adenine dinucleotide in catalysis. *Biochemistry* **1989,** 28, (19), 7555-60.

39. Skinner, M. A.; Gunel-Ozcan, A.; Moore, J.; Hawkins, A. R.; Brown, K. A., Dehydroquinate synthase binds divalent and trivalent cations: role of metal binding in catalysis. *Biochem Soc Trans* **1997,** 25, (4), S609.

40. Park, A.; Lamb, H. K.; Nichols, C.; Moore, J. D.; Brown, K. A.; Cooper, A.; Charles, I. G.; Stammers, D. K.; Hawkins, A. R., Biophysical and kinetic analysis of wild-type and site-directed mutants of the isolated and native dehydroquinate synthase domain of the AROM protein. *Protein Sci* **2004,** 13, (8), 2108-19.

41. Kharel, M. K.; Subba, B.; Basnet, D. B.; Woo, J. S.; Lee, H. C.; Liou, K.; Sohng, J. K., A gene cluster for biosynthesis of kanamycin from Streptomyces kanamyceticus: comparison with gentamicin biosynthetic gene cluster. *Arch Biochem Biophys* **2004,** 429, (2), 204-14.

42. Kharel, M. K.; Basnet, D. B.; Lee, H. C.; Liou, K.; Woo, J. S.; Kim, B. G.; Sohng, J. K., Isolation and characterization of the tobramycin biosynthetic gene

cluster from Streptomyces tenebrarius. *FEMS Microbiol Lett* **2004,** 230, (2), 185-90.

43. Busscher, G. F.; Rutjes, F. P.; van Delft, F. L., 2-Deoxystreptamine: central scaffold of aminoglycoside antibiotics. *Chem Rev* **2005,** 105, (3), 775-91.

44. Fourmy, D.; Recht, M. I.; Blanchard, S. C.; Puglisi, J. D., Structure of the A site of Escherichia coli 16S ribosomal RNA complexed with an aminoglycoside antibiotic. *Science* **1996,** 274, (5291), 1367-71.

45. Moazed, D.; Noller, H. F., Interaction of antibiotics with functional sites in 16S ribosomal RNA. *Nature* **1987,** 327, (6121), 389-94.

46. Schroeder, R.; Waldsich, C.; Wank, H., Modulation of RNA function by aminoglycoside antibiotics. *Embo J* **2000,** 19, (1), 1-9.

47. Huang, F.; Li, Y.; Yu, J.; Spencer, J. B., Biosynthesis of aminoglycoside antibiotics: cloning, expression and characterisation of an aminotransferase involved in the pathway to 2-deoxystreptamine. *Chem Commun (Camb)* **2002**, (23), 2860-1.

48. Yokoyama, K.; Kudo, F.; Kuwahara, M.; Inomata, K.; Tamegai, H.; Eguchi, T.; Kakinuma, K., Stereochemical recognition of doubly functional aminotransferase in 2-deoxystreptamine biosynthesis. *J Am Chem Soc* **2005,** 127, (16), 5869-74.

49. Kudo, F.; Hosomi, Y.; Tamegai, H.; Kakinuma, K., Purification and characterization of 2-deoxy-scyllo-inososе synthase derived from Bacillus circulans. A crucial carbocyclization enzyme in the biosynthesis of 2-deoxystreptamine-containing aminoglycoside antibiotics. *J Antibiot (Tokyo)* **1999,** 52, (2), 81-8.

50. Kudo, F.; Tamegai, H.; Fujiwara, T.; Tagami, U.; Hirayama, K.; Kakinuma, K., Molecular cloning of the gene for the key carbocycle-forming enzyme in the biosynthesis of 2-deoxystreptamine-containing aminocyclitol antibiotics and its comparison with dehydroquinate synthase. *J Antibiot (Tokyo)* **1999,** 52, (6), 559-71.

51. Kudo, F.; Yamauchi, N.; Suzuki, R.; Kakinuma, K., Kinetic isotope effect and reaction mechanism of 2-deoxy-scyllo-inososе synthase derived from butirosin-producing Bacillus circulans. *J Antibiot (Tokyo)* **1997,** 50, (5), 424-8.

52. Tamegai, H.; Nango, E.; Koike-Takeshita, A.; Kudo, F.; Kakinuma, K., Significance of the 20-kDa subunit of heterodimeric 2-deoxy-scyllo-inososе synthase for the biosynthesis of butirosin antibiotics in Bacillus circulans. *Biosci Biotechnol Biochem* **2002,** 66, (7), 1538-45.

53. Subba, B.; Kharel, M. K.; Lee, H. C.; Liou, K.; Kim, B. G.; Sohng, J. K., The ribostamycin biosynthetic gene cluster in Streptomyces ribosidificus: comparison with butirosin biosynthesis. *Mol Cells* **2005,** 20, (1), 90-6.

54. Huang, F.; Haydock, S. F.; Mironenko, T.; Spiteller, D.; Li, Y.; Spencer, J. B., The neomycin biosynthetic gene cluster of Streptomyces fradiae NCIMB 8233: characterisation of an aminotransferase involved in the formation of 2-deoxystreptamine. *Org Biomol Chem* **2005,** 3, (8), 1410-8.

55. Thuy, M. L.; Kharel, M. K.; Lamichhane, R.; Lee, H. C.; Suh, J. W.; Liou, K.; Sohng, J. K., Expression of 2-deoxy-scyllo-inososе synthase (kanA) from kanamycin gene cluster in Streptomyces lividans. *Biotechnol Lett* **2005,** 27, (7), 465-70.

56. Yanai, K.; Murakami, T., The kanamycin biosynthetic gene cluster from Streptomyces kanamyceticus. *J Antibiot (Tokyo)* **2004,** 57, (5), 351-4.

57. Yanai, K.; Murakami, T.; Bibb, M., Amplification of the entire kanamycin biosynthetic gene cluster during empirical strain improvement of Streptomyces kanamyceticus. *Proc Natl Acad Sci U S A* **2006,** 103, (25), 9661-6.

58. Kharel, M. K.; Subba, B.; Lee, H. C.; Liou, K.; Woo, J. S.; Sohng, J. K., An approach for cloning biosynthetic genes of 2-deoxystreptamine-containing aminocyclitol antibiotics: isolation of a biosynthetic gene cluster of tobramycin from Streptomyces tenebrarius. *Biotechnol Lett* **2003,** 25, (24), 2041-7.

59. Kharel, M. K.; Basnet, D. B.; Lee, H. C.; Liou, K.; Moon, Y. H.; Kim, J. J.; Woo, J. S.; Sohng, J. K., Molecular cloning and characterization of a 2-deoxystreptamine biosynthetic gene cluster in gentamicin-producing Micromonospora echinospora ATCC15835. *Mol Cells* **2004,** 18, (1), 71-8.

60. Nango, E.; Kumasaka, T.; Sato, T.; Tanaka, N.; Kakinuma, K.; Eguchi, T., Crystallization and X-ray analysis of 2-deoxy-scyllo-inososе synthase, the key enzyme in the biosynthesis of 2-deoxystreptamine-containing aminoglycoside antibiotics. *Acta Crystallograph Sect F Struct Biol Cryst Commun* **2005,** 61, (Pt 7), 709-11.

61. Flatt, P. M.; Mahmud, T., Biosynthesis of aminocyclitol-aminoglycoside antibiotics and related compounds. *Nat Prod Rep* **2007,** 24, (2), 358-92.

62. Oppolzer, W.; Prelog, V., [The constitution and configuration of rifamycins B, O, S and SV]. *Helv Chim Acta* **1973,** 56, (7), 2287-314.

63. Brufani, M.; Fedeli, W.; Giacomello, G.; Vaciago, A., The x-ray analysis of the structure of rifamycin B. *Experientia* **1964,** 20, (6), 339-42.

64. Kaur, H.; Cortes, J.; Leadlay, P.; Lal, R., Cloning and partial characterization of the putative rifamycin biosynthetic gene cluster from the actinomycete Amycolatopsis mediterranei DSM 46095. *Microbiol Res* **2001,** 156, (3), 239-46.

65. DeBoer, C.; Meulman, P. A.; Wnuk, R. J.; Peterson, D. H., Geldanamycin, a new antibiotic. *J Antibiot (Tokyo)* **1970,** 23, (9), 442-7.

66. Mao, Y.; Varoglu, M.; Sherman, D. H., Molecular characterization and analysis of the biosynthetic gene cluster for the antitumor antibiotic mitomycin C from Streptomyces lavendulae NRRL 2564. *Chem Biol* **1999,** 6, (4), 251-63.

67. 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, (11), 6030-40.

68. Rascher, A.; Hu, Z.; Buchanan, G. O.; Reid, R.; Hutchinson, C. R., Insights into the biosynthesis of the benzoquinone ansamycins geldanamycin and herbimycin, obtained by gene sequencing and disruption. *Appl Environ Microbiol* **2005,** 71, (8), 4862-71.

69. He, W.; Wu, L.; Gao, Q.; Du, Y.; Wang, Y., Identification of AHBA biosynthetic genes related to geldanamycin biosynthesis in Streptomyces hygroscopicus 17997. *Curr Microbiol* **2006,** 52, (3), 197-203.

70. Staley, A. L.; Rinehart, K. L., Biosynthesis of the streptovaricins: 3-amino-5-hydroxybenzoic acid as a precursor to the meta-C7N unit. *J Antibiot (Tokyo)* **1991,** 44, (2), 218-24.

71. Rude, M. A.; Khosla, C., Production of ansamycin polyketide precursors in Escherichia coli. *J Antibiot (Tokyo)* **2006,** 59, (8), 464-70.

72. Watanabe, K.; Rude, M. A.; Walsh, C. T.; Khosla, C., Engineered biosynthesis of an ansamycin polyketide precursor in Escherichia coli. *Proc Natl Acad Sci U S A* **2003,** 100, (17), 9774-8.

73. 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(7)N units in ansamycin antibiotics. *J. Am. Chem. Soc.* **1996,** 118, (32), 7486-7491.

74. 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, (16), 12546-55.

75. Chen, S.; von Bamberg, D.; Hale, V.; Breuer, M.; Hardt, B.; Muller, R.; Floss, H. G.; Reynolds, K. A.; Leistner, E., Biosynthesis of ansatrienin (mycotrienin) and naphthomycin. Identification and analysis of two separate biosynthetic gene clusters in Streptomyces collinus Tu 1892. *Eur J Biochem* **1999,** 261, (1), 98-107.

76. Moore, B. S.; Hertweck, C., Biosynthesis and attachment of novel bacterial polyketide synthase starter units. *Nat Prod Rep* **2002,** 19, (1), 70-99.

77. 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, (36), 10644-5.

78. Hornemann, U.; Eggert, J. H.; Honor, D. P., Role of D- [4-^{14}C]Erythrose and [3-^{14}C] Pyruvate in the Biosynthesis of the meta-C-C_6-N Unit of the Mitomycin Antibiotics in *Streptomyces verticillutus*. *J. Chem. Soc. Chem. Commun.* **1980,** 11.

79. 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, (2), 69-79.

80. Rascher, A.; Hu, Z.; Viswanathan, N.; Schirmer, A.; Reid, R.; Nierman, W. C.; Lewis, M.; Hutchinson, C. R., Cloning and characterization of a gene cluster for geldanamycin production in Streptomyces hygroscopicus NRRL 3602. *FEMS Microbiol Lett* **2003,** 218, (2), 223-30.

81. Yu, T. W.; Bai, L.; Clade, D.; Hoffmann, D.; Toelzer, S.; Trinh, K. Q.; Xu, J.; Moss, S. J.; Leistner, E.; Floss, H. G., The biosynthetic gene cluster of the maytansinoid antitumor agent ansamitocin from Actinosynnema pretiosum. *Proc Natl Acad Sci U S A* **2002,** 99, (12), 7968-73.

82. Mao, Y.; Varoglu, M.; Sherman, D. H., Genetic localization and molecular characterization of two key genes (mitAB) required for biosynthesis of the antitumor antibiotic mitomycin C. *J Bacteriol* **1999,** 181, (7), 2199-208.

83. Eads, J. C.; Beeby, M.; Scapin, G.; Yu, T. W.; Floss, H. G., Crystal structure of 3-amino-5-hydroxybenzoic acid (AHBA) synthase. *Biochemistry* **1999,** 38, (31), 9840-9.

84. Schmidt, D. D.; Frommer, W.; Junge, B.; Muller, L.; Wingender, W.; Truscheit, E.; Schafer, D., alpha-Glucosidase inhibitors. New complex oligosaccharides of microbial origin. *Naturwissenschaften* **1977,** 64, (10), 535-6.

85. Shibata, M.; Mori, K.; Hamashima, M., Inhibition of hyphal extension factor formation by validamycin in Rhizoctonia solani. *J Antibiot (Tokyo)* **1982,** 35, (10), 1422-3.

86. Asano, N.; Yamaguchi, T.; Kameda, Y.; Matsui, K., Effect of validamycins on glycohydrolases of Rhizoctonia solani. *J Antibiot (Tokyo)* **1987,** 40, (4), 526-32.

87. Kawasaki, T.; Kuzuyama, T.; Furihata, K.; Itoh, N.; Seto, H.; Dairi, T., A relationship between the mevalonate pathway and isoprenoid production in actinomycetes. *J Antibiot (Tokyo)* **2003,** 56, (11), 957-66.

88. Matsumoto, N.; Iinuma, H.; Sawa, T.; Takeuchi, T.; Hirano, S.; Yoshioka, T.; Ishizuka, M., Epoxyquinomicins A, B, C and D, new antibiotics from Amycolatopsis. II. Effect on type II collagen-induced arthritis in mice. *J Antibiot (Tokyo)* **1997,** 50, (11), 906-11.

89. Matsumoto, N.; Tsuchida, T.; Sawa, R.; Iinuma, H.; Nakamura, H.; Naganawa, H.; Sawa, T.; Takeuchi, T., Epoxyquinomicins A, B, C and D, new antibiotics from Amycolatopsis. III. Physico-chemical properties and structure determination. *J Antibiot (Tokyo)* **1997,** 50, (11), 912-5.

90. Matsumoto, N.; Tsuchida, T.; Umekita, M.; Kinoshita, N.; Iinuma, H.; Sawa, T.; Hamada, M.; Takeuchi, T., Epoxyquinomicins A, B, C and D, new antibiotics from Amycolatopsis. I. Taxonomy, fermentation, isolation and antimicrobial activities. *J Antibiot (Tokyo)* **1997,** 50, (11), 900-5.

91. Toyokuni, T.; Jin, W.-Z.; Rinehart, K. L., Biosynthetic Studies on Validamycins: A $C_2 + C_2 + C_3$ Pathway to an Aliphatic C7N Unit. *J. Am. Chem. Soc.* **1987,** 109, 3481-3482.

92. Degwert, U.; van Hulst, R.; Pape, H.; Herrold, R. E.; Beale, J. M.; Keller, P. J.; Lee, J. P.; Floss, H. G., Studies on the biosynthesis of the alpha-glucosidase inhibitor acarbose: valienamine, a m-C7N unit not derived from the shikimate pathway. *J Antibiot (Tokyo)* **1987,** 40, (6), 855-61.

93. Mahmud, T.; Tornus, I.; Egelkrout, E.; Wolf, E.; Uy, C.; Floss, H. G.; Lee, S., Biosynthetic Studies on the R-Glucosidase Inhibitor Acarbose in Actinoplanes sp.: 2-*epi*-5-*epi*-Valiolone Is the Direct Precursor of the Valienamine Moiety. *J. Am. Chem. Soc.* **1999,** 121, (30), 6973-6983.

94. Mahmud, T.; Lee, S.; Floss, H. G., The biosynthesis of acarbose and validamycin. *Chem Rec* **2001,** 1, (4), 300-10.

95. Dong, H.; Mahmud, T.; Tornus, I.; Lee, S.; Floss, H. G., Biosynthesis of the validamycins: identification of intermediates in the biosynthesis of validamycin A by Streptomyces hygroscopicus var. limoneus. *J Am Chem Soc* **2001,** 123, (12), 2733-42.

96. Wu, X.; Flatt, P. M.; Schlorke, O.; Zeeck, A.; Dairi, T.; Mahmud, T., A comparative analysis of the sugar phosphate cyclase superfamily involved in primary and secondary metabolism. *Chembiochem* **2007,** 8, (2), 239-48.

97. Naganawa, H.; Hashizume, H.; Kubota, Y.; Sawa, R.; Takahashi, Y.; Arakawa, K.; Bowers, S. G.; Mahmud, T., Biosynthesis of the cyclitol moiety of pyralomicin 1a in Nonomuraea spiralis MI178-34F18. *J Antibiot (Tokyo)* **2002,** 55, (6), 578-84.

98. Schlorke, O.; Krastel, P.; Muller, I.; Uson, I.; Dettner, K.; Zeeck, A., Structure and biosynthesis of cetoniacytone A, a cytotoxic aminocarba sugar produced by an endosymbiontic Actinomyces. *J Antibiot (Tokyo)* **2002,** 55, (7), 635-42.

99. Stratmann, A.; Mahmud, T.; Lee, S.; Distler, J.; Floss, H. G.; Piepersberg, W., The AcbC protein from Actinoplanes species is a C7-cyclitol synthase related to 3-dehydroquinate synthases and is involved in the biosynthesis of the alpha-glucosidase inhibitor acarbose. *J Biol Chem* **1999,** 274, (16), 10889-96.

100. Yu, Y.; Bai, L.; Minagawa, K.; Jian, X.; Li, L.; Li, J.; Chen, S.; Cao, E.; Mahmud, T.; Floss, H. G.; Zhou, X.; Deng, Z., Gene cluster responsible for validamycin biosynthesis in Streptomyces hygroscopicus subsp. jinggangensis 5008. *Appl Environ Microbiol* **2005,** 71, (9), 5066-76.

101. Bai, L.; Li, L.; Xu, H.; Minagawa, K.; Yu, Y.; Zhang, Y.; Zhou, X.; Floss, H. G.; Mahmud, T.; Deng, Z., Functional analysis of the validamycin biosynthetic gene cluster and engineered production of validoxylamine A. *Chem Biol* **2006,** 13, (4), 387-97.

102. Lee, S.; Egelkrout, E., Biosynthetic studies on the alpha-glucosidase inhibitor acarbose in Actinoplanes sp.: glutamate is the primary source of the nitrogen in acarbose. *J Antibiot (Tokyo)* **1998,** 51, (2), 225-7.

103. Lee, S.; Sauerbrei, B.; Niggemann, J.; Egelkrout, E., Biosynthetic studies on the alpha-glucosidase inhibitor acarbose in Actinoplanes sp.: source of the maltose unit. *J Antibiot (Tokyo)* **1997,** 50, (11), 954-60.

104. Arakawa, K.; Bowers, S. G.; Michels, B.; Trin, V.; Mahmud, T., Biosynthetic studies on the alpha-glucosidase inhibitor acarbose: the chemical synthesis of isotopically labeled 2-epi-5-epi-valiolone analogs. *Carbohydr Res* **2003,** 338, (20), 2075-82.

105. Zhang, C. S.; Stratmann, A.; Block, O.; Bruckner, R.; Podeschwa, M.; Altenbach, H. J.; Wehmeier, U. F.; Piepersberg, W., Biosynthesis of the C(7)-cyclitol moiety of acarbose in Actinoplanes species SE50/110. 7-O-phosphorylation of the initial cyclitol precursor leads to proposal of a new biosynthetic pathway. *J Biol Chem* **2002,** 277, (25), 22853-62.

106. Zhang, C. S.; Podeschwa, M.; Altenbach, H. J.; Piepersberg, W.; Wehmeier, U. F., The acarbose-biosynthetic enzyme AcbO from Actinoplanes sp. SE 50/110 is

a 2-epi-5-epi-valiolone-7-phosphate 2-epimerase. *FEBS Lett* **2003,** 540, (1-3), 47-52.

107. Jian, X.; Pang, X.; Yu, Y.; Zhou, X.; Deng, Z., Identification of genes necessary for jinggangmycin biosynthesis from Streptomyces hygroscopicus 10-22. *Antonie Van Leeuwenhoek* **2006,** 90, (1), 29-39.

108. Singh, D.; Seo, M. J.; Kwon, H. J.; Rajkarnikar, A.; Kim, K. R.; Kim, S. O.; Suh, J. W., Genetic localization and heterologous expression of validamycin biosynthetic gene cluster isolated from Streptomyces hygroscopicus var. limoneus KCCM 11405 (IFO 12704). *Gene* **2006,** 376, (1), 13-23.

109. Kawamura, N.; Sawa, R.; Takahashi, Y.; Isshiki, K.; Sawa, T.; Naganawa, H.; Takeuchi, T., Pyralomicins, novel antibiotics from Microtetraspora spiralis. II. Structure determination. *J Antibiot (Tokyo)* **1996,** 49, (7), 651-6.

110. Kawamura, N.; Sawa, R.; Takahashi, Y.; Issiki, K.; Sawa, T.; Kinoshita, N.; Naganawa, H.; Hamada, M.; Takeuchi, T., Pyralomicins, new antibiotics from Actinomadura spiralis. *J Antibiot (Tokyo)* **1995,** 48, (5), 435-7.

111. Kawamura, N.; Sawa, R.; Takahashi, Y.; Sawa, T.; Naganawa, H.; Takeuchi, T., Pyralomicins, novel antibiotics from Microtetraspora spiralis. III. Biosynthesis of pyralomicin 1a. *J Antibiot (Tokyo)* **1996,** 49, (7), 657-60.

112. Nowak-Thompson, B.; Chaney, N.; Wing, J. S.; Gould, S. J.; Loper, J. E., Characterization of the pyoluteorin biosynthetic gene cluster of Pseudomonas fluorescens Pf-5. *J Bacteriol* **1999,** 181, (7), 2166-74.

113. Torigoe, K.; Wakasugi, N.; Sakaizumi, N.; Ikejima, T.; Suzuki, H.; Kojiri, K.; Suda, H., BE-40644, a new human thioredoxin system inhibitor isolated from Actinoplanes sp. A40644. *J Antibiot (Tokyo)* **1996,** 49, (3), 314-7.

114. Seto, H.; Orihara, N.; Furihata, K., Studies on the Biosynthesis of Terpenoids Produced by Actinomycetes.Part 4. Formation of BE-40644 by the Mevalonate and Nonmevalonate Pathways. *Tetrahedron Letters* **1998,** 39, 9497-9500.

115. Vértesy, L.; Fehlhaber, H.-W.; Schulz, A., The Trehalase Inhibitor Salbostatin, a Novel Metabolite from Streptomyces albus, ATCC21838. *Angewandte Chemie International Edition in English* **1994,** 33, (18), 1844-1846.

116. Choeng, Y.-H.; Yang, J.-Y.; Delcroix, G.; Jung, K. Y.; Chang, Y. K.; Hong, S.-K., Expression and Characterization of Trehalose Biosynthetic Modules in the Adjacent Locus of the Salbostatin Gene Cluster. *J. Microbiol. Biotechnol* **2007,** 17, (10), 1675-1681.

117. Tsuchida, T.; Umekita, M.; Kinoshita, N.; Iinuma, H.; Nakamura, H.; Nakamura, K.; Naganawa, H.; Sawa, T.; Hamada, M.; Takeuchi, T., Epoxyquinomicins A and B, new antibiotics from Amycolatopsis. *J Antibiot (Tokyo)* **1996,** 49, (3), 326-8.

118. Matsumoto, N.; Iinuma, H.; Sawa, T.; Takeuchi, T., Synthesis of anti-rheumatic agent epoxyquinomicin B. *Bioorg Med Chem Lett* **1998,** 8, (21), 2945-8.

119. Matsumoto, N.; Agata, N.; Kuboki, H.; Iinuma, H.; Sawa, T.; Takeuchi, T.; Umezawa, K., Inhibition of rat embryo histidine decarboxylase by epoxyquinomicins. *J Antibiot (Tokyo)* **2000,** 53, (6), 637-9.

120. Umezawa, K.; Ariga, A.; Matsumoto, N., Naturally occurring and synthetic inhibitors of NF-kappaB functions. *Anticancer Drug Des* **2000,** 15, (4), 239-44.

121. Horiguchi, Y.; Kuroda, K.; Nakashima, J.; Murai, M.; Umezawa, K., Antitumor effect of a novel nuclear factor-kappa B activation inhibitor in bladder cancer cells. *Expert Rev Anticancer Ther* **2003,** 3, (6), 793-8.

122. Umezawa, K.; Chaicharoenpong, C., Molecular design and biological activities of NF-kappaB inhibitors. *Mol Cells* **2002,** 14, (2), 163-7.

123. Miyajima, A.; Kosaka, T.; Seta, K.; Asano, T.; Umezawa, K.; Hayakawa, M., Novel nuclear factor kappa B activation inhibitor prevents inflammatory injury in unilateral ureteral obstruction. *J Urol* **2003,** 169, (4), 1559-63.

124. Sato, A.; Oya, M.; Ito, K.; Mizuno, R.; Horiguchi, Y.; Umezawa, K.; Hayakawa, M.; Murai, M., Survivin associates with cell proliferation in renal cancer cells: regulation of survivin expression by insulin-like growth factor-1, interferon-gamma and a novel NF-kappaB inhibitor. *Int J Oncol* **2006,** 28, (4), 841-6.

125. Poma, P.; Notarbartolo, M.; Labbozzetta, M.; Sanguedolce, R.; Alaimo, A.; Carina, V.; Maurici, A.; Cusimano, A.; Cervello, M.; D'Alessandro, N., Antitumor effects of the novel NF-kappaB inhibitor dehydroxymethyl-epoxyquinomicin on human hepatic cancer cells: analysis of synergy with cisplatin and of possible

correlation with inhibition of pro-survival genes and IL-6 production. *Int J Oncol* **2006,** 28, (4), 923-30.

126. Matsumoto, G.; Namekawa, J.; Muta, M.; Nakamura, T.; Bando, H.; Tohyama, K.; Toi, M.; Umezawa, K., Targeting of nuclear factor kappaB Pathways by dehydroxymethylepoxyquinomicin, a novel inhibitor of breast carcinomas: antitumor and antiangiogenic potential in vivo. *Clin Cancer Res* **2005,** 11, (3), 1287-93.

127. Matsumoto, N.; Ariga, A.; To-e, S.; Nakamura, H.; Agata, N.; Hirano, S.; Inoue, J.; Umezawa, K., Synthesis of NF-kappaB activation inhibitors derived from epoxyquinomicin C. *Bioorg Med Chem Lett* **2000,** 10, (9), 865-9.

128. Kieser, T.; Bibb, M. J.; Buttner, M. J.; chater, K. F.; Hopwood, D. A., *practical streptomyces genetics*. 2nd ed.; John Innes: 2000.

129. sambrook, J.; Fritsch, E. F.; Maniatis, T., *Molecular Cloning: A Laboratory Manual*. 2nd ed.; Cold Spring Harbor Laboratory Press: 1989.

130. Kirschning, A.; Bergon, P.; Wang, J. J.; Breazeale, S.; Floss, H. G., Synthesis of 4-amino 3,4-dideoxy-D-arabino-heptulosonic acid 7-phosphate, the biosynthetic precursor of C7N units in ansamycin antibiotics. *Carbohydr Res* **1994,** 256, (2), 245-56.

131. Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J., Basic local alignment search tool. *J Mol Biol* **1990,** 215, (3), 403-10.

132. Bierman, M.; Logan, R.; O'Brien, K.; Seno, E. T.; Rao, R. N.; Schoner, B. E., Plasmid cloning vectors for the conjugal transfer of DNA from Escherichia coli to Streptomyces spp. *Gene* **1992,** 116, (1), 43-9.

133. Zazopoulos, E.; Huang, K.; Staffa, A.; Liu, W.; Bachmann, B. O.; Nonaka, K.; Ahlert, J.; Thorson, J. S.; Shen, B.; Farnet, C. M., A genomics-guided approach for discovering and expressing cryptic metabolic pathways. *Nat Biotechnol* **2003,** 21, (2), 187-90.

134. 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, (6), 1698-706.

135. Kelly, W. L.; Townsend, C. A., Mutational analysis of nocK and nocL in the nocardicin a producer Nocardia uniformis. *J Bacteriol* **2005,** 187, (2), 739-46.

136. 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, (5), 662-73.

137. Methe, B. A.; Nelson, K. E.; Deming, J. W.; Momen, B.; Melamud, E.; Zhang, X.; Moult, J.; Madupu, R.; Nelson, W. C.; Dodson, R. J.; Brinkac, L. M.; Daugherty, S. C.; Durkin, A. S.; DeBoy, R. T.; Kolonay, J. F.; Sullivan, S. A.; Zhou, L.; Davidsen, T. M.; Wu, M.; Huston, A. L.; Lewis, M.; Weaver, B.; Weidman, J. F.; Khouri, H.; Utterback, T. R.; Feldblyum, T. V.; Fraser, C. M., The psychrophilic lifestyle as revealed by the genome sequence of Colwellia psychrerythraea 34H through genomic and proteomic analyses. *Proc Natl Acad Sci U S A* **2005,** 102, (31), 10913-8.

138. Kaneko, T.; Nakamura, Y.; Sato, S.; Asamizu, E.; Kato, T.; Sasamoto, S.; Watanabe, A.; Idesawa, K.; Ishikawa, A.; Kawashima, K.; Kimura, T.; Kishida, Y.; Kiyokawa, C.; Kohara, M.; Matsumoto, M.; Matsuno, A.; Mochizuki, Y.; Nakayama, S.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takeuchi, C.; Yamada, M.; Tabata, S., Complete genome structure of the nitrogen-fixing symbiotic bacterium Mesorhizobium loti. *DNA Res* **2000,** 7, (6), 331-8.

139. Holton, S. J.; Dairou, J.; Sandy, J.; Rodrigues-Lima, F.; Dupret, J. M.; Noble, M. E.; Sim, E., Structure of Mesorhizobium loti arylamine N-acetyltransferase 1. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **2005,** 61, (Pt 1), 14-6.

140. Paulsen, I. T.; Press, C. M.; Ravel, J.; Kobayashi, D. Y.; Myers, G. S.; Mavrodi, D. V.; DeBoy, R. T.; Seshadri, R.; Ren, Q.; Madupu, R.; Dodson, R. J.; Durkin, A. S.; Brinkac, L. M.; Daugherty, S. C.; Sullivan, S. A.; Rosovitz, M. J.; Gwinn, M. L.; Zhou, L.; Schneider, D. J.; Cartinhour, S. W.; Nelson, W. C.; Weidman, J.; Watkins, K.; Tran, K.; Khouri, H.; Pierson, E. A.; Pierson, L. S., 3rd; Thomashow, L. S.; Loper, J. E., Complete genome sequence of the plant commensal Pseudomonas fluorescens Pf-5. *Nat Biotechnol* **2005,** 23, (7), 873-8.

141. Vardy, E.; Arkin, I. T.; Gottschalk, K. E.; Kaback, H. R.; Schuldiner, S., Structural conservation in the major facilitator superfamily as revealed by comparative modeling. *Protein Sci* **2004,** 13, (7), 1832-40.

142. Raty, K.; Kantola, J.; Hautala, A.; Hakala, J.; Ylihonko, K.; Mantsala, P., Cloning and characterization of Streptomyces galilaeus aclacinomycins polyketide synthase (PKS) cluster. *Gene* **2002,** 293, (1-2), 115-22.

143. 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, (22), 5527-35.

144. Matsunaga, T.; Okamura, Y.; Fukuda, Y.; Wahyudi, A. T.; Murase, Y.; Takeyama, H., Complete genome sequence of the facultative anaerobic magnetotactic bacterium Magnetospirillum sp. strain AMB-1. *DNA Res* **2005,** 12, (3), 157-66.

145. Takakura, Y.; Kuwata, S., Purification, characterization, and molecular cloning of a pyranose oxidase from the fruit body of the basidiomycete, Tricholoma matsutake. *Biosci Biotechnol Biochem* **2003,** 67, (12), 2598-607.

146. Hallberg, B. M.; Leitner, C.; Haltrich, D.; Divne, C., Crystallization and preliminary X-ray diffraction analysis of pyranose 2-oxidase from the white-rot fungus Trametes multicolor. *Acta Crystallogr D Biol Crystallogr* **2004,** 60, (Pt 1), 197-9.

147. Hallberg, B. M.; Leitner, C.; Haltrich, D.; Divne, C., Crystal structure of the 270 kDa homotetrameric lignin-degrading enzyme pyranose 2-oxidase. *J Mol Biol* **2004,** 341, (3), 781-96.

148. Kujawa, M.; Ebner, H.; Leitner, C.; Hallberg, B. M.; Prongjit, M.; Sucharitakul, J.; Ludwig, R.; Rudsander, U.; Peterbauer, C.; Chaiyen, P.; Haltrich, D.; Divne, C., Structural basis for substrate binding and regioselective oxidation of monosaccharides at C3 by pyranose 2-oxidase. *J Biol Chem* **2006,** 281, (46), 35104-15.

149. Ahlert, J.; Distler, J.; Mansouri, K.; Piepersberg, W., Identification of stsC, the gene encoding the L-glutamine:scyllo-inosose aminotransferase from streptomycin-producing Streptomycetes. *Arch Microbiol* **1997,** 168, (2), 102-13.

150. Gonzalez, V.; Santamaria, R. I.; Bustos, P.; Hernandez-Gonzalez, I.; Medrano-Soto, A.; Moreno-Hagelsieb, G.; Janga, S. C.; Ramirez, M. A.; Jimenez-Jacinto, V.; Collado-Vides, J.; Davila, G., The partitioned Rhizobium etli genome:

genetic and metabolic redundancy in seven interacting replicons. *Proc Natl Acad Sci U S A* **2006,** 103, (10), 3834-9.

151. Sohng, J. K.; Oh, T. J.; Lee, J. J.; Kim, C. G., Identification of a gene cluster of biosynthetic genes of rubradirin substructures in S. achromogenes var. rubradiris NRRL3061. *Mol Cells* **1997,** 7, (5), 674-81.

152. Krugel, H.; Schumann, G.; Hanel, F.; Fiedler, G., Nucleotide sequence analysis of five putative Streptomyces griseus genes, one of which complements an early function in daunorubicin biosynthesis that is linked to a putative gene cluster involved in TDP-daunosamine formation. *Mol Gen Genet* **1993,** 241, (1-2), 193-202.

153. Minagawa, K.; Zhang, Y.; Ito, T.; Bai, L.; Deng, Z.; Mahmud, T., ValC, a new type of C7-Cyclitol kinase involved in the biosynthesis of the antifungal agent validamycin A. *Chembiochem* **2007,** 8, (6), 632-41.

154. Hong, S. T.; Carney, J. R.; Gould, S. J., Cloning and heterologous expression of the entire gene clusters for PD 116740 from Streptomyces strain WP 4669 and tetrangulol and tetrangomycin from Streptomyces rimosus NRRL 3016. *J Bacteriol* **1997,** 179, (2), 470-6.

155. Otsuka, M.; Ichinose, K.; Fujii, I.; Ebizuka, Y., Cloning, sequencing, and functional analysis of an iterative type I polyketide synthase gene cluster for biosynthesis of the antitumor chlorinated polyenone neocarzilin in "Streptomyces carzinostaticus". *Antimicrob Agents Chemother* **2004,** 48, (9), 3468-76.

156. Thapa, L. P.; Oh, T. J.; Lee, H. C.; Liou, K.; Park, J. W.; Yoon, Y. J.; Sohng, J. K., Heterologous expression of the kanamycin biosynthetic gene cluster (pSKC2) in Streptomyces venezuelae YJ003. *Appl Microbiol Biotechnol* **2007,** 76, (6), 1357-64.

157. Mehdi, S.; Frost, J. W.; Knowles, J. R., Dehydroquinate synthase from Escherichia coli, and its substrate 3-deoxy-D-arabino-heptulosonic acid 7-phosphate. *Methods Enzymol* **1987,** 142, 306-14.

158. Knaggs, A. R., The biosynthesis of shikimate metabolites. *Nat Prod Rep* **2003,** 20, (1), 119-36.

159. Charles, I. G.; Keyte, J. W.; Brammar, W. J.; Hawkins, A. R., Nucleotide sequence encoding the biosynthetic dehydroquinase function of the penta-

functional arom locus of Aspergillus nidulans. *Nucleic Acids Res* **1985,** 13, (22), 8119-28.

160. Meeks, J. C.; Elhai, J.; Thiel, T.; Potts, M.; Larimer, F.; Lamerdin, J.; Predki, P.; Atlas, R., An overview of the genome of Nostoc punctiforme, a multicellular, symbiotic cyanobacterium. *Photosynth Res* **2001,** 70, (1), 85-106.

161. Kaneko, T.; Nakamura, Y.; Wolk, C. P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, M.; Ishikawa, A.; Kawashima, K.; Kimura, T.; Kishida, Y.; Kohara, M.; Matsumoto, M.; Matsuno, A.; Muraki, A.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S., Complete genomic sequence of the filamentous nitrogen-fixing cyanobacterium Anabaena sp. strain PCC 7120. *DNA Res* **2001,** 8, (5), 205-13; 227-53.

162. Tamegai, H.; Kuki, T.; Udagawa, Y.; Aoki, R.; Nagaya, A.; Tsukada, S., Exploration of genes that encode a carbocycle-forming enzyme involved in biosynthesis of aminoglycoside antibiotics from the environmental DNA. *Biosci Biotechnol Biochem* **2006,** 70, (7), 1711-6.

163. Vetting, M. W.; Wackett, L. P.; Que, L., Jr.; Lipscomb, J. D.; Ohlendorf, D. H., Crystallographic comparison of manganese- and iron-dependent homoprotocatechuate 2,3-dioxygenases. *J Bacteriol* **2004,** 186, (7), 1945-58.

164. Heiss, G.; Muller, C.; Altenbuchner, J.; Stolz, A., Analysis of a new dimeric extradiol dioxygenase from a naphthalenesulfonate-degrading sphingomonad. *Microbiology* **1997,** 143 (Pt 5), 1691-9.

165. Andujar, E.; Santero, E., Site-directed mutagenesis of an extradiol dioxygenase involved in tetralin biodegradation identifies residues important for activity or substrate specificity. *Microbiology* **2003,** 149, (Pt 6), 1559-67.

166. Eltis, L. D.; Bolin, J. T., Evolutionary relationships among extradiol dioxygenases. *J Bacteriol* **1996,** 178, (20), 5930-7.

167. Happe, B.; Eltis, L. D.; Poth, H.; Hedderich, R.; Timmis, K. N., Characterization of 2,2',3-trihydroxybiphenyl dioxygenase, an extradiol dioxygenase from the dibenzofuran- and dibenzo-p-dioxin-degrading bacterium Sphingomonas sp. strain RW1. *J Bacteriol* **1993,** 175, (22), 7313-20.

168. Kovaleva, E. G.; Lipscomb, J. D., Crystal structures of Fe2+ dioxygenase superoxo, alkylperoxo, and bound product intermediates. *Science* **2007,** 316, (5823), 453-7.

169. Martin, T. W.; Dauter, Z.; Devedjiev, Y.; Sheffield, P.; Jelen, F.; He, M.; Sherman, D. H.; Otlewski, J.; Derewenda, Z. S.; Derewenda, U., Molecular basis of mitomycin C resistance in streptomyces: structure and function of the MRD protein. *Structure* **2002,** 10, (7), 933-42.

170. Dumas, P.; Bergdoll, M.; Cagnon, C.; Masson, J. M., Crystal structure and site-directed mutagenesis of a bleomycin resistance protein and their significance for drug sequestering. *Embo J* **1994,** 13, (11), 2483-92.

171. Vanbelle, C.; Brutscher, B.; Blackledge, M.; Muhle-Goll, C.; Remy, M. H.; Masson, J. M.; Marion, D., NMR study of the interaction between Zn(II) ligated bleomycin and Streptoalloteichus hindustanus bleomycin resistance protein. *Biochemistry* **2003,** 42, (3), 651-63.

172. Maruyama, M.; Kumagai, T.; Matoba, Y.; Hayashida, M.; Fujii, T.; Hata, Y.; Sugiyama, M., Crystal structures of the transposon Tn5-carried bleomycin resistance determinant uncomplexed and complexed with bleomycin. *J Biol Chem* **2001,** 276, (13), 9992-9.

173. McCarthy, A. A.; Baker, H. M.; Shewry, S. C.; Patchett, M. L.; Baker, E. N., Crystal structure of methylmalonyl-coenzyme A epimerase from P. shermanii: a novel enzymatic function on an ancient metal binding scaffold. *Structure* **2001,** 9, (7), 637-46.

174. Bobik, T. A.; Rasche, M. E., HPLC assay for methylmalonyl-CoA epimerase. *Anal Bioanal Chem* **2003,** 375, (3), 344-9.

175. Bobik, T. A.; Rasche, M. E., Purification and partial characterization of the Pyrococcus horikoshii methylmalonyl-CoA epimerase. *Appl Microbiol Biotechnol* **2004,** 63, (6), 682-5.

176. Leadlay, P. F., Purification and characterization of methylmalonyl-CoA epimerase from Propionibacterium shermanii. *Biochem J* **1981,** 197, (2), 413-9.

177. Kuhnl, J.; Bobik, T.; Procter, J. B.; Burmeister, C.; Hoppner, J.; Wilde, I.; Luersen, K.; Torda, A. E.; Walter, R. D.; Liebau, E., Functional analysis of the

methylmalonyl-CoA epimerase from Caenorhabditis elegans. *Febs J* **2005,** 272, (6), 1465-77.

178. Ariza, A.; Vickers, T. J.; Greig, N.; Armour, K. A.; Dixon, M. J.; Eggleston, I. M.; Fairlamb, A. H.; Bond, C. S., Specificity of the trypanothione-dependent Leishmania major glyoxalase I: structure and biochemical comparison with the human enzyme. *Mol Microbiol* **2006,** 59, (4), 1239-48.

179. Thornalley, P. J., Glyoxalase I--structure, function and a critical role in the enzymatic defence against glycation. *Biochem Soc Trans* **2003,** 31, (Pt 6), 1343-8.

180. Richter, U.; Krauss, M., Active site structure and mechanism of human glyoxalase I-an ab initio theoretical study. *J Am Chem Soc* **2001,** 123, (29), 6973-82.

181. Himo, F.; Siegbahn, P. E., Catalytic mechanism of glyoxalase I: a theoretical study. *J Am Chem Soc* **2001,** 123, (42), 10280-9.

182. Cameron, A. D.; Olin, B.; Ridderstrom, M.; Mannervik, B.; Jones, T. A., Crystal structure of human glyoxalase I--evidence for gene duplication and 3D domain swapping. *Embo J* **1997,** 16, (12), 3386-95.

183. Ridderstrom, M.; Cameron, A. D.; Jones, T. A.; Mannervik, B., Involvement of an active-site Zn2+ ligand in the catalytic mechanism of human glyoxalase I. *J Biol Chem* **1998,** 273, (34), 21623-8.

184. Pakhomova, S.; Rife, C. L.; Armstrong, R. N.; Newcomer, M. E., Structure of fosfomycin resistance protein FosA from transposon Tn2921. *Protein Sci* **2004,** 13, (5), 1260-5.

185. Fillgrove, K. L.; Pakhomova, S.; Schaab, M. R.; Newcomer, M. E.; Armstrong, R. N., Structure and mechanism of the genomically encoded fosfomycin resistance protein, FosX, from Listeria monocytogenes. *Biochemistry* **2007,** 46, (27), 8110-20.

186. Bergdoll, M.; Eltis, L. D.; Cameron, A. D.; Dumas, P.; Bolin, J. T., All in the family: structural and evolutionary relationships among three modular proteins with diverse functions and variable assembly. *Protein Sci* **1998,** 7, (8), 1661-70.

187. Frickel, E. M.; Jemth, P.; Widersten, M.; Mannervik, B., Yeast glyoxalase I is a monomeric enzyme with two active sites. *J Biol Chem* **2001,** 276, (3), 1845-9.

188. Bennett-Lovsey, R.; Herbert, A.; Sternberg, M.; Kelley, L., Exploring the extremes of sequence/structure space with ensemble fold recognition in the program Phyre. *Proteins: Structure, Function, Bioinformatics* **2008,** 70, (3).

189. Kelley, L. A.; MacCallum, R. M.; Sternberg, M. J., Enhanced genome annotation using structural profiles in the program 3D-PSSM. *J Mol Biol* **2000,** 299, (2), 499-520.