2010

DISSECTING THE BIOSYNTHESES OF GILVOCARCINS AND RAVIDOMYCINS

Madan Kumar Kharel

University of Kentucky, kharel_madan@yahoo.com

Click here to let us know how access to this document benefits you.

ABSTRACT OF DISSERTATION

Madan Kumar Kharel

The Graduate School

University of Kentucky

2010

DISSECTING THE BIOSYNTHESES OF GILVOCARCINS AND RAVIDOMYCINS

_____

ABSTRACT OF DISSERTATION

_____

A dissertation submitted in partial fulfillment of the requirement
for the degree of Doctor of Philosophy in the College of Pharmacy
at the University of Kentucky

By
Madan Kumar Kharel

Jhapa, Nepal

Director: Dr. Jürgen Rohr
Professor of Pharmaceutical Science

Lexington, Kentucky

2010

ABSTRACT OF DISSERTATION

DISSECTING THE BIOSYNTHESES OF GILVOCARCINS AND RAVIDOMYCINS

Gilvocarcin V (GV) and ravidomycin (RMV) exhibit excellent antitumor activities in the presence of near-UV light at low concentration maintaining a low *in vivo* cytotoxicity. Although, the exact molecular mechanism for *in vivo* actions of these antibiotics has yet to be determined, a [2+2] cycloaddition reaction of the vinyl side chain with DNA thymidine residues in addition to the inhibition of topoisomerase II and DNA-histone H3 cross-linking are reported for the GV's mechanism of action. Such activities have made these molecules interesting candidates for the biosynthetic investigation to generate analogues with improved activity/solubility. Previous biosynthetic studies have suggested that the GV biosynthetic pathway involves a number of synchronously occurring transformations leading to the oxidative C-C bond cleavage and other intriguing biosynthetic reactions, such as the vinyl side chain formation, methylations, *C*-glycosylation and dehydrogenation. Although gene inactivation results identified many candidate genes whose corresponding enzymes are involved in these biochemical transformations, their exact functional roles and the identity of their natural substrates remained elusive. To provide more insights into these complex biochemical tranfromations, three specific aims were set up.

Specific aim 1 was to clone and characterize the RMV biosynthetic gene cluster. Through the comparison of GV cluster with the RMV cluster, the genes encoding the biosynthesis of sugar and tetracyclic aromatic moieties were identified. RavGT, the sole glycosyltransferase of the RMV cluster has demonstrated to have unprecedented sugar donor substrate flexibility, transferring an amino-pyranose sugar as well as a neutral furanose sugar.

Specific aim 2 was to characterize all of the TDP-D-ravidosamine biosynthetic enzymes. The aim also included to a one-pot enzymatic synthetic protocol for the routine production of TDP-D-ravidosamine.

Specific aim 3 focussed on a total enzymatic synthesis of defucogilvocarcin M (defucoGM), the polyketide-derived core of GV and RMV. This aim clearly identified the minimal enzymes required to biosynthesize the complex architecture of defucoGM from the simple building blocks acetate and malonate. In addition, the GV-pathway enzyme GilR was fully characterized. Through *in vitro* studies, GilR was shown to catalyze the dehydrogenation of hemiacetal moiety of the penultimate intermediate pregilvocarcin V to the lactone moiety of GV at the last step.

Keywords: Gilvocarcin V, Ravidomycin V, enzymatic total synthesis, deoxysugar biosynthesis, dehydrogenase

_____

Student's Signature

_____

Date

DISSECTING THE BIOSYNTHESES OF GILVOCARCINS AND RAVIDOMYCINS

By

Madan Kumar Kharel

Jürgen Rohr, Ph.D.

Director of Dissertation

Robert A. Yokel, Ph.D.

Director of Graduate Studies

Date

RULES FOR THE USE OF DISSERTATIONS

Unpublished dissertations submitted for the Doctor's degree and deposited in the University of Kentucky library are as rule open for inspection, but are to be used only with due regard to the right of authors. Bibliographical references may be noted, but quotations or summaries of parts may be published only with permission of the authors, and with usual scholarly acknowledgment.

Extensive copying or publication of the dissertation in whole or in part requires also the consent of the Dean of the Graduate School of the University of Kentucky.

A library which borrows this dissertation for use by its patrons is expected to secure the signature of each user

Name   Date

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

DISSERTATION

Madan Kumar Kharel

The Graduate School

University of Kentucky

2010

DISSECTING THE BIOSYNTHESES OF GILVOCARCINS AND RAVIDOMYCINS

_____

DISSERTATION

_____

A dissertation submitted in partial fulfillment of the requirement

for the degree of Doctor of Philosophy in the College of Pharmacy

at the University of Kentucky

By

Madan Kumar Kharel

Jhapa, Nepal

Director: Dr Jürgen Rohr, Professor of Pharmarceutical Sciences

Lexington, Kentucky

2010

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SCHEMES

LIST OF ABBREVIATIONS

aa	amino acid
ACP	acyl carrier protein
APCI	atmospheric pressure chemical ionization
ARO	aromatase
AT	acyl transferase
ATP	adenosine triphosphate
BLAST	basic local alignment search tool
BVMO	Baeyer-Villiger monooxygenase
bp	base pairs
CLF	chain length factor (KS_β)
CoA	coenzyme A
COSY	correlation spectroscopy
CYC	cyclase
CYP450	cytochrome P-450 monooxygenases
DEBS	6-deoxyerythronolide B synthase
DH	dehydratase
DNA	deoxyribonucleic acid
DMSO	dimethylsulfoxide
E. coli	*Escherichia coli*
EI-HR-MS	Electron impact high resolution mass spectrometry

ER	Enoyl reductase
ESI	electrospray ionization
FAD	flavin adenine dinucleotide
G-1-P	glucose-1-phosphate
GNAT	GCN-5 related *N*-acetyltransferase
GT	glycosyltransferse
GV	gilvocarcin V
HPLC	high performance liquid chromatography
IMAC	Immobilized metal affinity chromatography
IPTG	isopropyl-*β*-D-thiogalactoside
KS	ketoacyl synthase
KR	keto reductase
LB	Luria broth
MALDI-TOF	Matrix Assisted Laser Desorption /Ionization- Time Of Flight
MCAT	malonyl CoA acyl carrier protein transacylase
MT	methyltransferase
MW	Molecular weight
NAD^+	nicotinamide adenine dinucleotide, oxidized form
NAD(P)H	nicotinamide adenine dinucleotide phosphate, reduced form
NMR	nuclear magnetic resonance
ORF	open reading frame

PCP   peptidyl carrier protein

PCR   polymerase chain reaction

PKS   polyketide synthase

R_t HPLC retention time

SAR   structure activity relationship

SAM   *S*-adenosyl methionine

TE   thioesterase

UV   ultraviolet radiation

Section One: Background

Polyketides

Polyketides represent one of the major classes of natural products. The polyketide class covers a heterogenous group of compounds such as macrolides, enedyines, anthracyclines, angucyclines, polyenes, macrolactams and polyethers (*1*). They are ubiquitous in nature spanning from prokaryotes such as bacteria to eukaryotes such as plants, mollusks and insects. Their exact roles in the context of their original biological system are not fully known yet, however, it is believed that these secondary metabolites serve as pigments, virulence factors, infochemicals or as defense weapons. Many of these compounds or their derivatives have become clinically useful pharmaceutics such as antibiotics, immunosuppressants, antiparasites, cholesterol-lowering agents, and antitumor agents, while others are infamous toxins or virulence factors (*2*). Antibacterial agents, such as oxytetracycline (**1**), erythromycin A (**5**), and the rifamycin derivative rifampicin (**9**), monensin (**6**), the antifungal agent amphotericin B (**7**), the cholesterol lowering agent lovastatin (**3**), the immunosuppressant FK506 (**4**), the antitumor agents doxorubicin (**2**), and mithramycin (**8**) etc. represent some of the clinically used polyketides (**Figure 1**.**1**). Because of their tremendous structural diversities and pharamacological relevance, much attention has been paid in the last few decades to access synthetic routes for these natural products or derivatives thereof. Despite remarkable progress in chemical synthetic techniques over the past decades, the total synthesis of many polyketides is still very challenging due to their vast structural complexities and diversities, although they are assembled from the simplest building blocks acetate or propionate (*1*). This led to the emergence of a new research area called "Combinatorial Biosynthesis" which utilizes the synthesis of natural products or derivatives thereof modifying nature's original synthetic strategies and tools. Different types of polyketides and their biosynthetic routes are briefly discussed below.
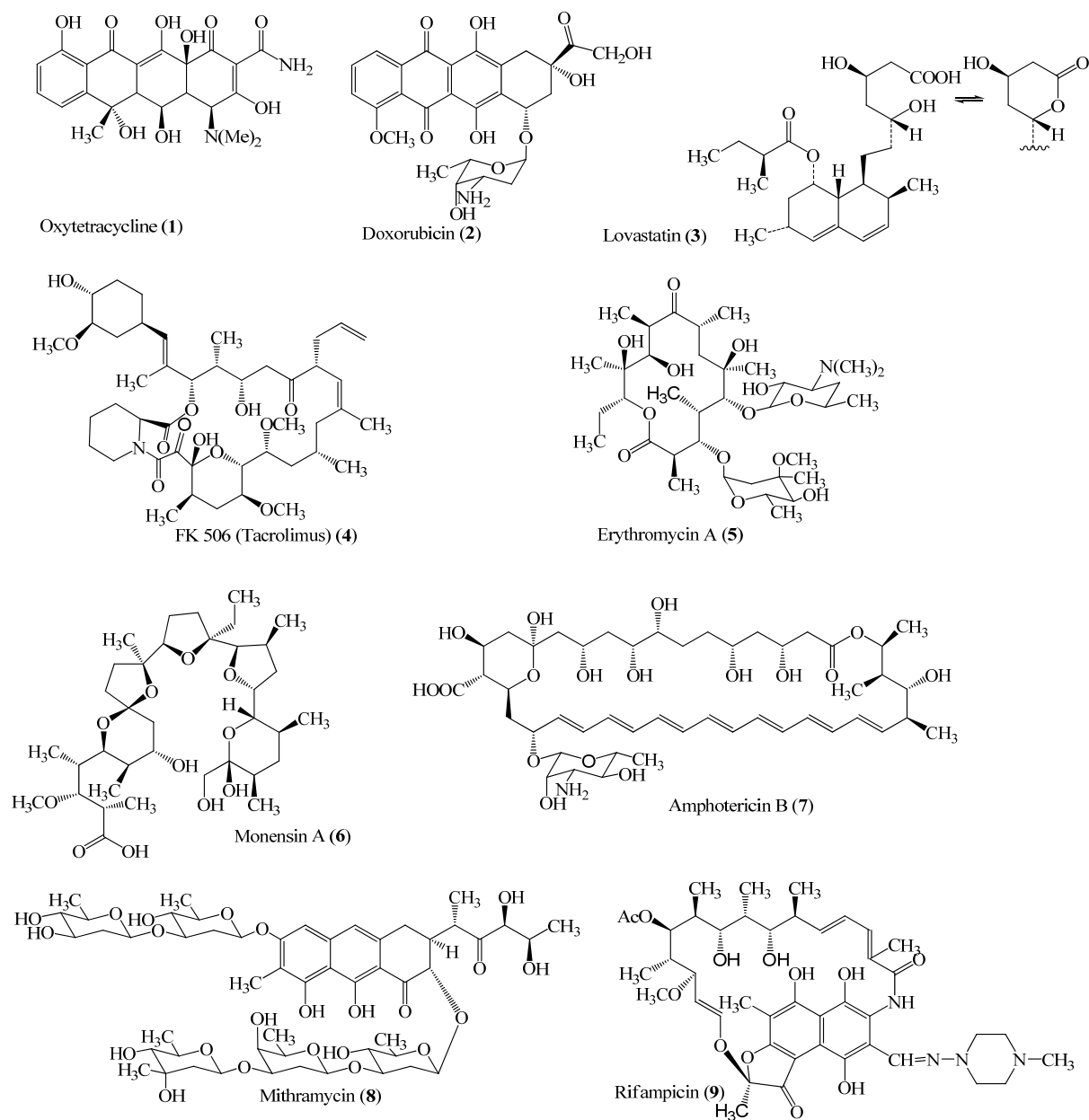
Figure 1.1 Representative examples of polyketide drugs.

Biosynthesis of polyketides

Polyketides biosyntheses are catalyzed by complexes of mono-or multi-functional enzymes called polyketide synthases (PKSs). The genes encoding these enzymes are usually organized in clusters, presumably to facilitate a coordinated regulation of expression of multiple

enzymes required for the many steps in these specialized biosynthetic pathways (*3*). The biosynthetic genes are also associated with resistance genes and export genes to prevent the producer organisms from self-intoxication of their own products. The chemistry associated with PKSs is closely related to the well studied fatty acid synthases (FASs). Both pathways do not only share a common biosynthetic logic, but they also share common building blocks such as acetyl-coenzyme A (acetyl-CoA) and malonyl-CoA (*4*). Decarboxylative Claisen thioester condensations of an activated starter unit with malonyl-CoA or malonyl-CoA-derived extender units generate *β*-ketoester intermediates. This process requires *β*-ketoacyl synthase (KS), acylcarrier protein (ACP) and an optional malonyl/acyltransferase (MAT/AT). During fatty acid biosynthesis, every elongation step follows *β*-ketoreduction, dehydration and enoyl reduction catalyzed by ketoreductase (KR), dehydratase (DH) and enoyl reductase (ER), respectively, to yield a fully saturated acyl backbone. In contrast, each polyketide chain extension cycle is followed by full, partial or no reduction steps giving rise to a complex pattern of functionalization (**Scheme 1.1**). Unlike FASs, PKSs are more promiscuous, use a variety of building blocks and can also generate products with various chain lengths. In both pathways, the carbon backbone continues to grow until a defined length is achieved, and the fully grown products cleave off from the thioester-bound substrates of the enzyme complexes.

Based on the assembly line enzymes, their mode of actions and architectures, PKSs are divided into three major categories known as type I, type II and type III (*5*). Type I PKSs consist of large multifunctional enzyme complexes where each enzyme has a catalytic domain and a carrier protein domain organized like beads on a string (*6*). Such assembly lines are further divided into two different systems, such as modular and iterative. In the modular system, each module consists of optional *β*-keto processing domains (KR, DH, ER) in addition to a set of KS, AT and ACP domains. These catalytic and carrier domains are used only once as the growing chain passes from *N*-terminal to *C*-terminal. The number of modules thus reflects the number of cycles catalyzed by a particular type I PKS and the presence of KR, DH and ER domains determines the degree of *β*-keto reduction in each module. In general, the number of modules and the number of catalytic domains in each module corresponds to the overall architecture of the PKSs product, thus making it possible to predict the product structure through the analysis of catalytic domains of PKSs or vice versa (*1*). The PKS of the erythromycin biosynthesis represents the proto-typical PKS (**Scheme 1.2**).
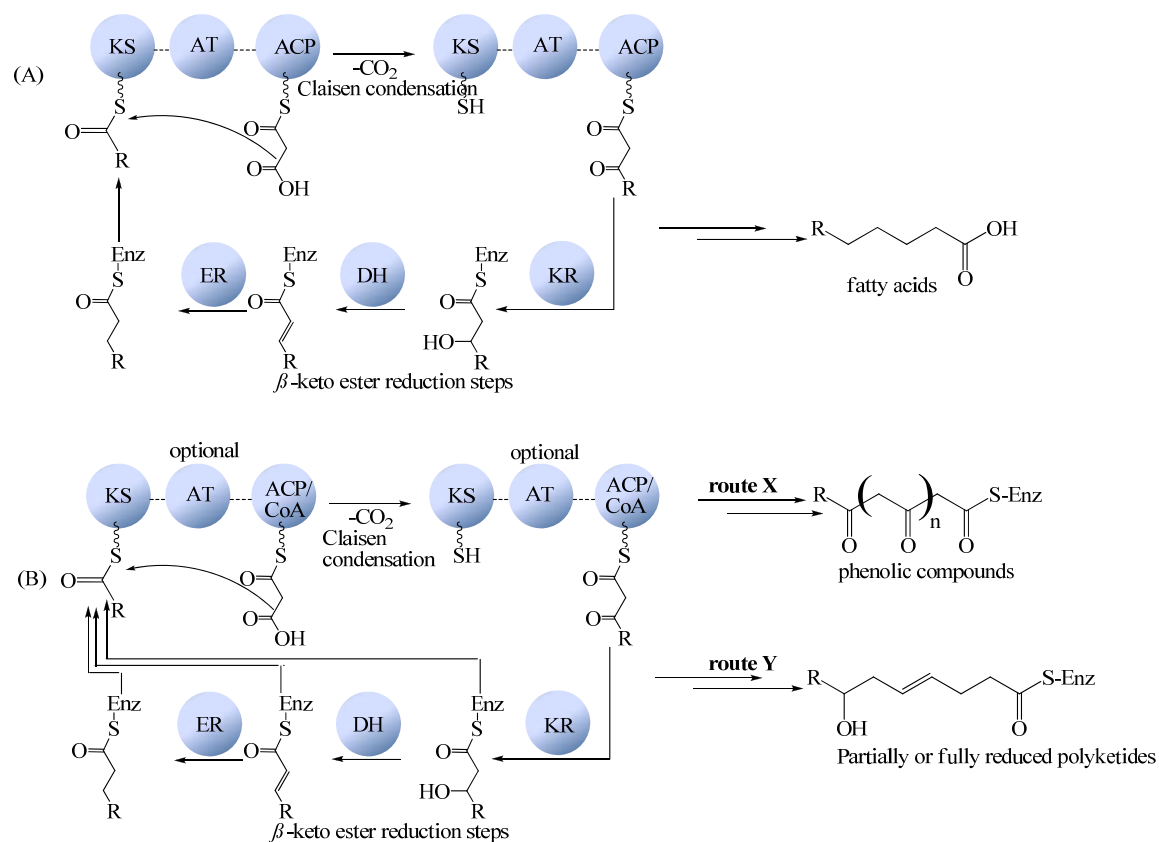
Scheme 1.1 Fatty acids (A) and polyketide (B) biosyntheses. Route X and route Y represent biosynthesis of unreduced and fully/partially reduced polyketides.
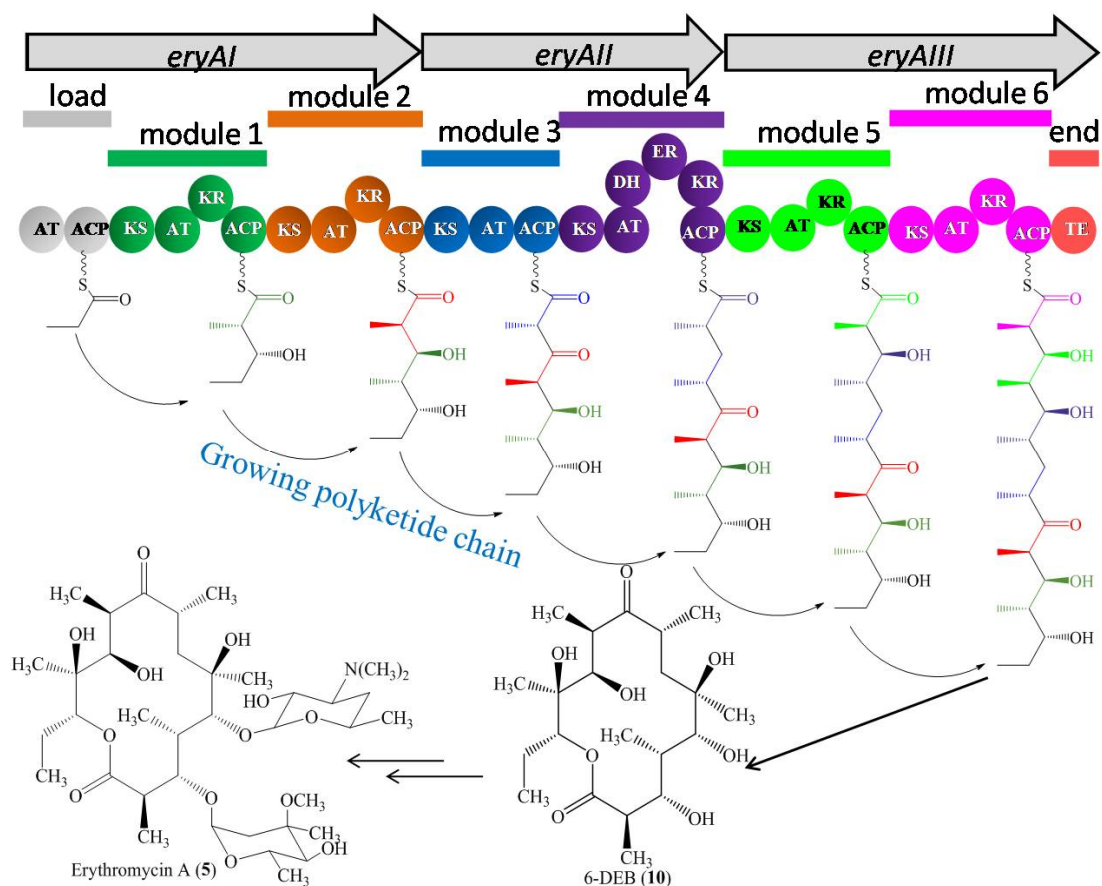
Scheme 1.2 Erythromycin A biosynthesis, 6-DEB: 6-deoxyerythronolide B.

The cluster is composed of three open reading frames (ORFs), namely DEBS1 (*eryAI*), DEBS2 (*eryAII*) and DEBS3 (*eryAIII*) with a total of six modules and a loading domain (*3, 7, 8*). The biosynthesis starts with the loading of a starter unit propionyl-CoA to the holo-ACP domain. The KS domain of module 1 self acylates with the extender unit (methylmalonyl-CoA). Decarboxylative condensation between the ACP-tethered starter unit and the extender unit results in the formation of a diketide. The reduction of the keto group by the KR domain results in the formation of a hydroxyl group at the C-3 position. The AT domain of module 1 catalyzes the transfer of the growing chain from the loading ACP to the ACP of module 1. The process continues until the designated chain length is achieved. The final task of the PKS is to disconnect the mature PKS product from the ACP in module 6. This is achieved by thiosterase domain of *eryAIII*, which catalyzes an intramolecular lactone formation, where the hydroxyl group at C3 serves as a nucleophile to the carbonyl carbon of the thioester resulting in the formation of 6-erythronolide B (**10**). However, there are several other type I PKSs, where a module is used more

than once, some domains are skipped or the modules lack AT domains. In such AT deficient PKSs, standalone ATs load starter or extender units to ACPs in *trans* fashion (*9, 10*). Such iterative type I PKSs are usually found in fungi. The lovastatin PKS is a typical example (*11, 12*). In these PKSs, KR, DH and ER domains are optionally used in every round of chain elongation as in lovastatin biosynthesis (**Scheme**). How this is controlled is not yet fully understood.
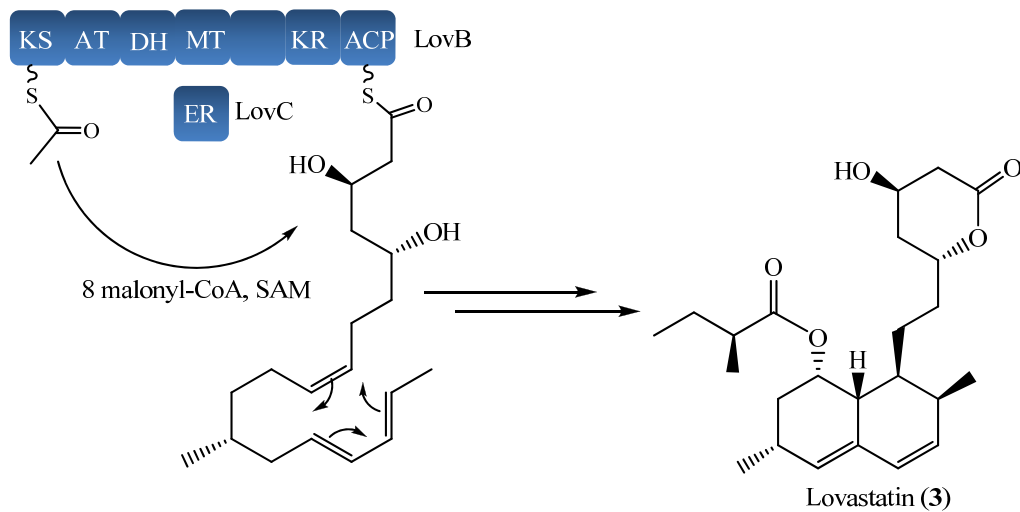
Scheme 1.3 Lovastatin biosynthesis. The intermediate undergoes intramolecular Diels-Alder reaction to establish a decalin moiety of lovastatin.

Iteratively operating type II PKSs are more common in prokaryotes. Unlike type I PKSs, a minimal set of enzymes are used iteratively to generate poly β-keto thioester. Ketoacyl synthases [KS_α and KS_β (chain length factor, CLF)] and ACP constitute the minimal PKS (**Scheme 1.4**) (*13*). However, other enzymes, such as, ketoreductases (KRs) and cyclases (CYCs) and aromatases (AROs) work with the minimal PKS in an enzymatic complex to control the folding, cyclization and aromatization pattern of the nascent polyketide chain. The type II PKS system responsible for the actinorhodin (**11**) biosynthesis was first reported from *Streptomyces coelicolor* (*14*). Doxorubicin (**2**) (**Scheme 1.**), gilvocarcin V (GV, **12**), landomycin A (**13**) and tetracycline (**14**) are other examples of type II PKS-derived antibiotics (**Figure**). Acetyl-CoA or propionyl-CoA are used as starter units, whereas malonyl-CoA is used for chain elongation. Type II PKSs are usually found in actinomycetes with an exception of a few other Gram-negative bacteria (*15, 16*).
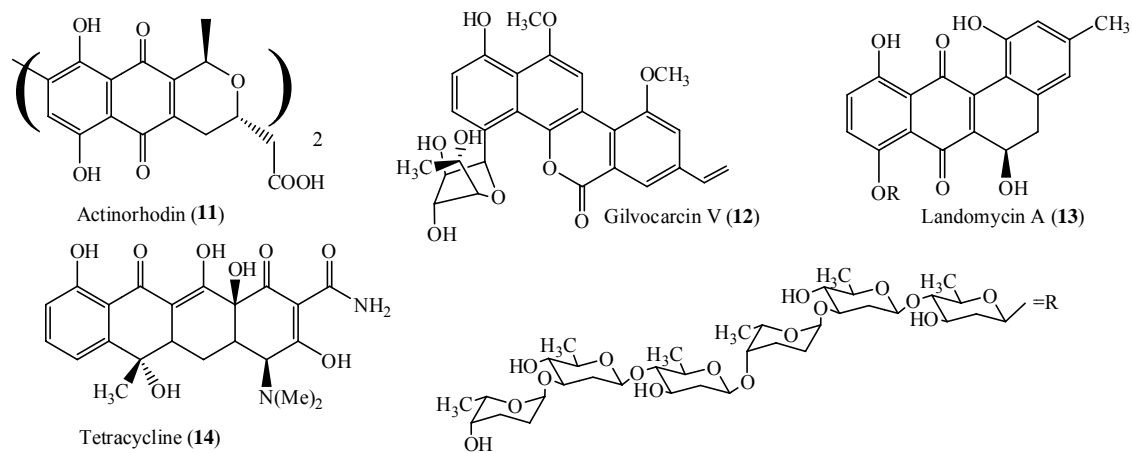
Figure 1.2 Some examples of type II PKS derived antibiotics.
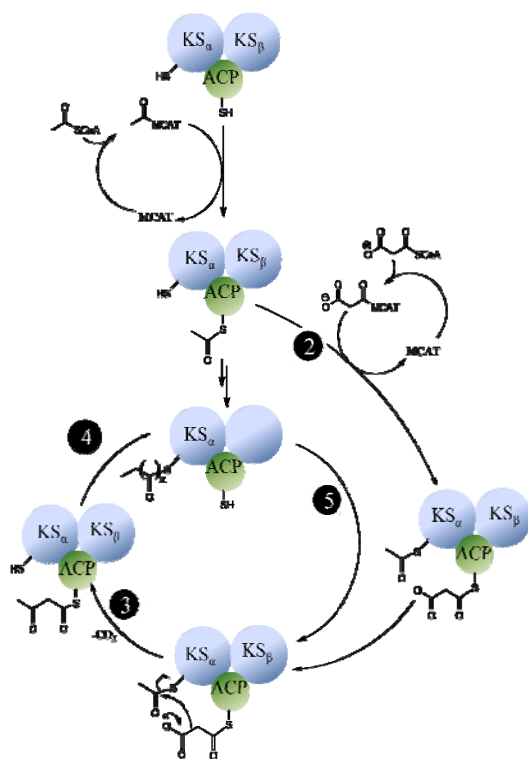
Scheme 1.4 Polyketide chain formation/elongation by type II PKSs.
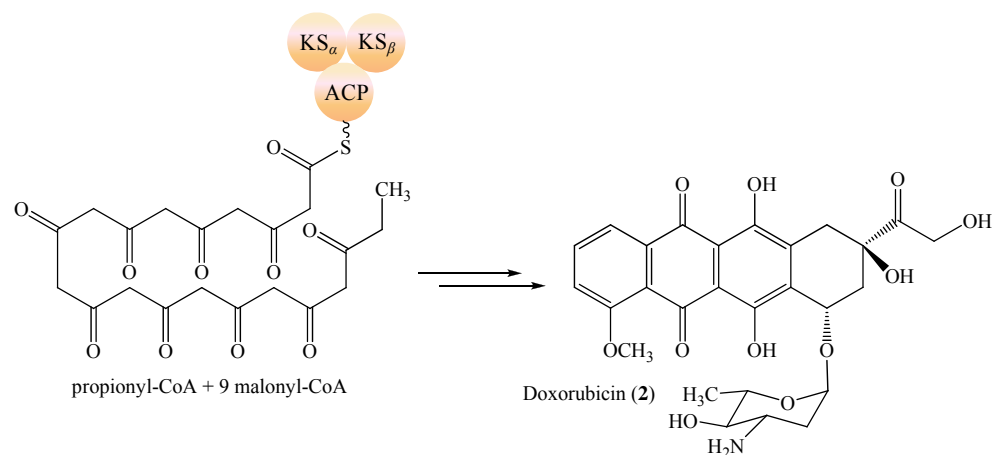
Scheme 1.5 Type II PKS-catalyzed biosynthesis of doxorubicin.

Iterative type III PKSs do not require any ACP-bound extender units for chain elongation and utilize acyl-CoA substrates directly (**Scheme 1.**). Biosynthesis of the plant product naringenin chalcone (**15**) through a sequential addition of 3 molecules of malonyl-CoA to a p-coumaryl-CoA starter unit represents a typical type III PKS catalysis. These enzymes are known as stilbene/chalcone synthases, and are found in plants and bacteria as well as in fungi (*17-19*). In addition to these three general types of polyketides, there are mixed type polyketide synthases, such as type I-type II, type I-type III, FAS-PKS hybrid or PKS-NRPS (non-ribosomal peptide synthases) (*1*).
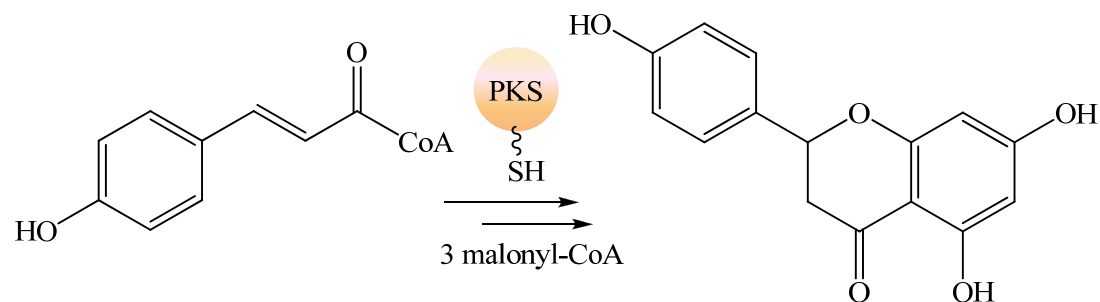
Scheme 1.6 Type III PKS-mediated biosynthesis of naringenin chalcone

8

Post-PKS tailoring reactions

The products of PKSs or PKS-conjugates usually undergo a series of tailoring events such as glycosylation, methylation, oxidation, reduction, prenylation and halogenations to generate biologically active compounds. Such tailorings cause tremendous change to the pharmacological property of the parent polyketides, such as solubility or binding ability of the drug to a receptor. Erythromycin A represents such a typical example where the PKS-product (6-DEB) has no activity, but oxygenation and glycosylations reactions change it into a powerful antibacterial agent (*20*). On the other hand, post-PKS tailoring reactions provide a unique platform to design natural product derivatives through pathway engineering, also known as a type of "combinatorial biosynthesis" (*21*). This typically involves targeted gene inactivation, gene replacement, homologous/heterologous expression of the genes or site-directed mutagenesis techniques, or combinations. Preparation of mithramycin analogues with altered sugar pattern through the expression of various deoxysugar biosynthetic constructs, preparation of ketopremithramycins and combinatorial biosynthesis of urdamycins are some of several examples of such a combinatorial biosynthetic approach (*22-27*). Some common post-PKS tailoring reactions, which are observed in various natural product biosynthetic pathways, and are relevant to this work, are briefly discussed in the following section.

Glycosylation

Glycosylation is known to be one of the most important post-PKS tailoring reactions because it not only modulates biological activity of the molecule, but can also provide self-resistance to the producer organism. Activation of 6-DEB by glycosylation, and inactivation of oleandomycin or erythromycin A by addition of glucose to their desosamine moieties are typical examples (*28*). Glycosylation usually occurs in the later stages of a biosynthetic pathway and the reaction typically involves an enzyme or rarely two enzymes known as glycosyltransferases (GTs) which use NDP-activated sugars as donor substrates and an acceptor substrate. The transfer of the sugar to the acceptor substrate yields *O-*, *N-*, *S-* or *C-* glycosidic bonds. These glycosyltransferases are grouped into ~90 sequence-based families (*29*). However, three dimensional structures of all of the reported GTs fall into two major folds, GT-A and GT-B (*21, 29*). Glycosylation chemistry involves a GT-mediated nucleophilic attack of the acceptor substrate at the anomeric carbon of the donor substrate, with the simultaneous release of the

nucleotide diphosphate, either in an S_N2-fashion or via an oxocarbenium-ion-intermediate (if the activating group leaves prior to the glycosidic bond formation) (**Scheme 1.7**) (*30*). Although glycosylations have long been believed to be unidirectional, Thorson et. al. have recently demonstrated that some natural product *O*-glycosylation reactions in fact are reversible (*31*).
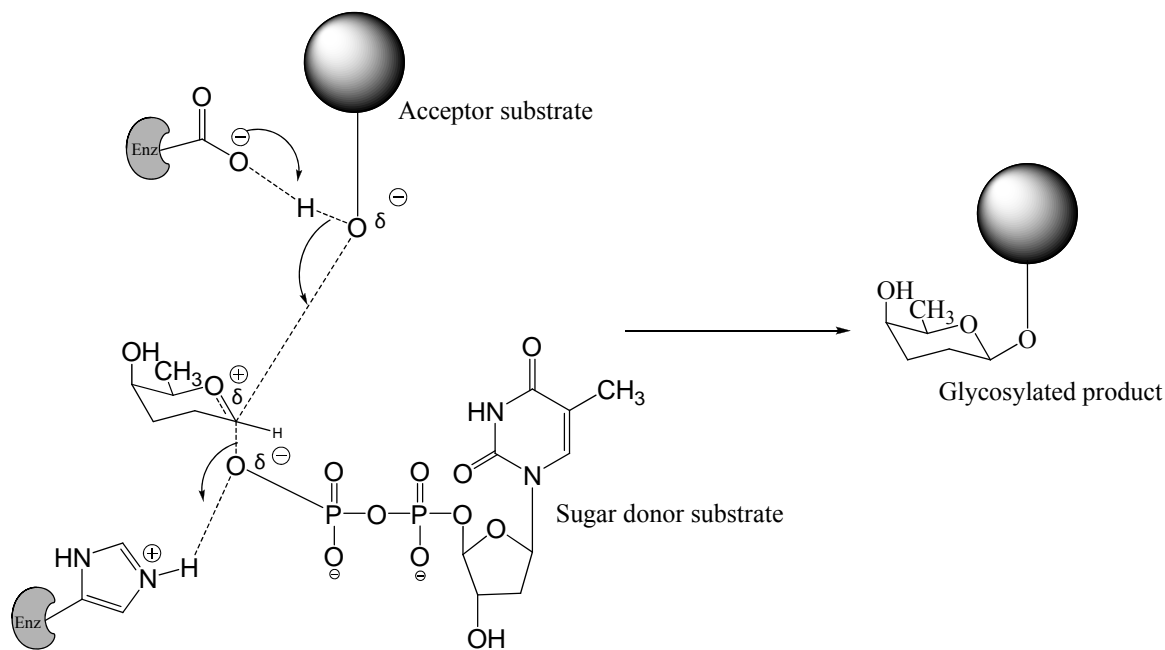
Scheme 1.7 Proposed mechanism for inverting type *O*-GTs. The figure is adapted from Rohr et al 2002 (*30*).

Most of the *C*-glycosides from microbial secondary metabolites have the sugar attached ortho or para to an electron rich functional groups (e.g. -OH) of the aromatic moiety. Such functional groups create higher electron density at the ortho- or para- position, making it possible to attack at the anomeric centre of the sugar (**Scheme 1.**) as in the case of urdamycin A and GV (*32, 33*). However, exact mechanisms for *C*- or *N*-glycosylations are still obscure. Based on the stereochemical change at the anomeric carbon of the sugar, GTs are further classified into two groups; inverting GTs and retaining GTs. Most of the secondary metabolite producer actinomycete GTs fall into the inverting type of GT-1 family.
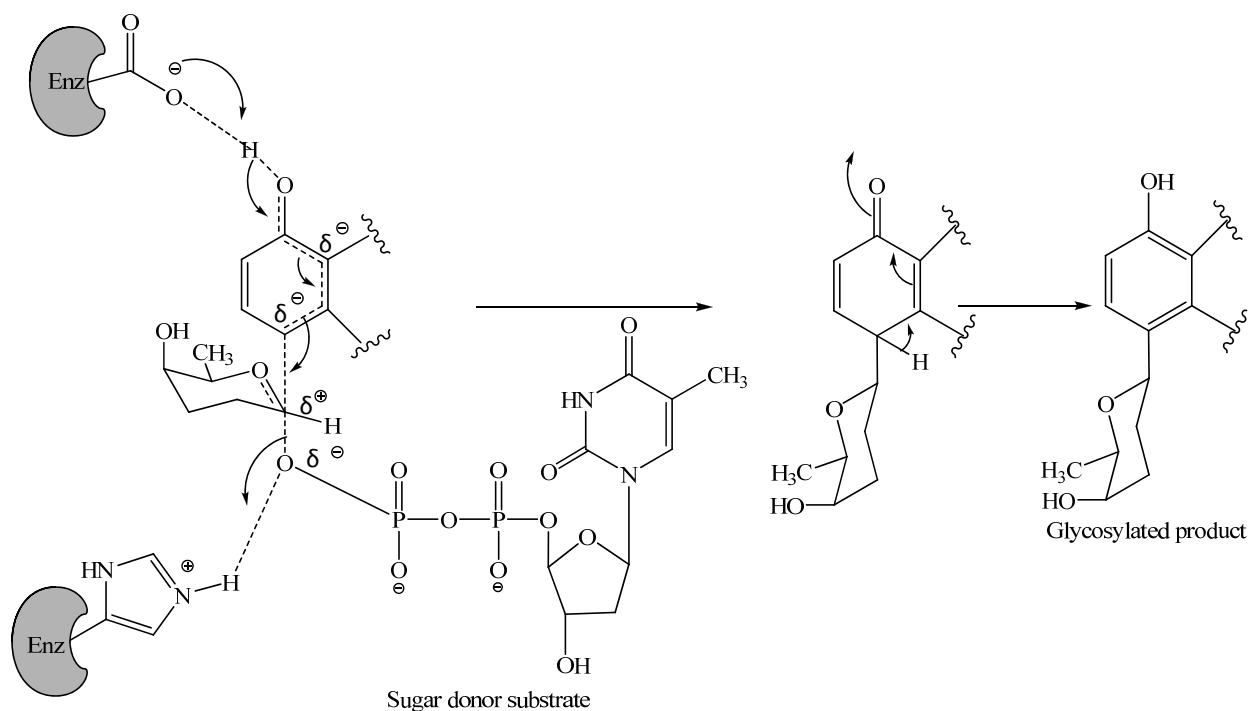
Scheme 1.8 A hypothetical mechanism for *C*-glycosylation.

In the past few decades, a variety of GTs have been identified from various secondary metabolite biosynthetic pathways. Most of these GTs have relatively strict substrate specificities. However, some others appeared to operate on a variety of sugar donor substrates. Those GTs with inherent broad substrate specificities are of special interest as they can be utilized to generate a variety of unnatural derivatives of natural products with potentially better therapeutic indices. Generation of an array of tetracenomycins, steffimycins and aranciamycins using a variety of sugar donor substrates with promiscuous ElmGT, StfGT and AraGT, respectively, highlights the potential application of such GTs in combinatorial biosynthesis (*34-40*). In addition, point mutation or domain swapping or saturation mutagenesis can be carried out to improve promiscuity of the GTs of interest. Generation of various unnatural glycosides through the modification of OleD (glycosyltransferase from oleandomycin biosynthetic pathway)(*41-43*) and generation of prejadomycin *C*-glycosides through the expression of UrdGT2 into LndE minus mutant are some representatives of such endeavors (*44*). These results clearly show the tremendous application of GTs to rationally design natural product glycosides through combinatorial biosynthesis.

Deoxysugar biosynthesis

Sugar moieties of most of the microbial natural products are often deoxygenated to various extents prior to their attachment to the aglycones. Deoxygenation of hexose sugars usually begins with conversion hexose-1-phosphate into its nucleotide diphosphate (NDP) derivative by NDP-hexose-nucleotidyl transferases (*45*). The dedicated enzymes utilize NTP (ATP, CTP, GTP, UTP or TTP) as a source of NDP. While TDP-activations are the most dominant in microbial deoxysugar biosynthetic pathways, GDP- and UDP-activations were also reported. The aminopentose moiety of calicheamicin and L-lyxose-derived moiety of avilamycin are examples of UDP-α-D-glucose-derived deoxysugars (*46, 47*). Similarly, deoxysugar moieties of nystatin and amphotericin B are derived from GDP-mannose (*48*). NDP-hexose-4,6-dehydratase catalyzes a redox reaction that deoxygenates the C-6 carbon and oxidizes C-4 carbon of the activated sugar to generate NDP-4-keto-6- deoxy-D-hexose (**18**). The product serves as a common intermediate for all of 6-deoxysugar pathways. Formation of various TDP-deoxysugars through **18** is shown in the **Scheme 1.9** (*49*).
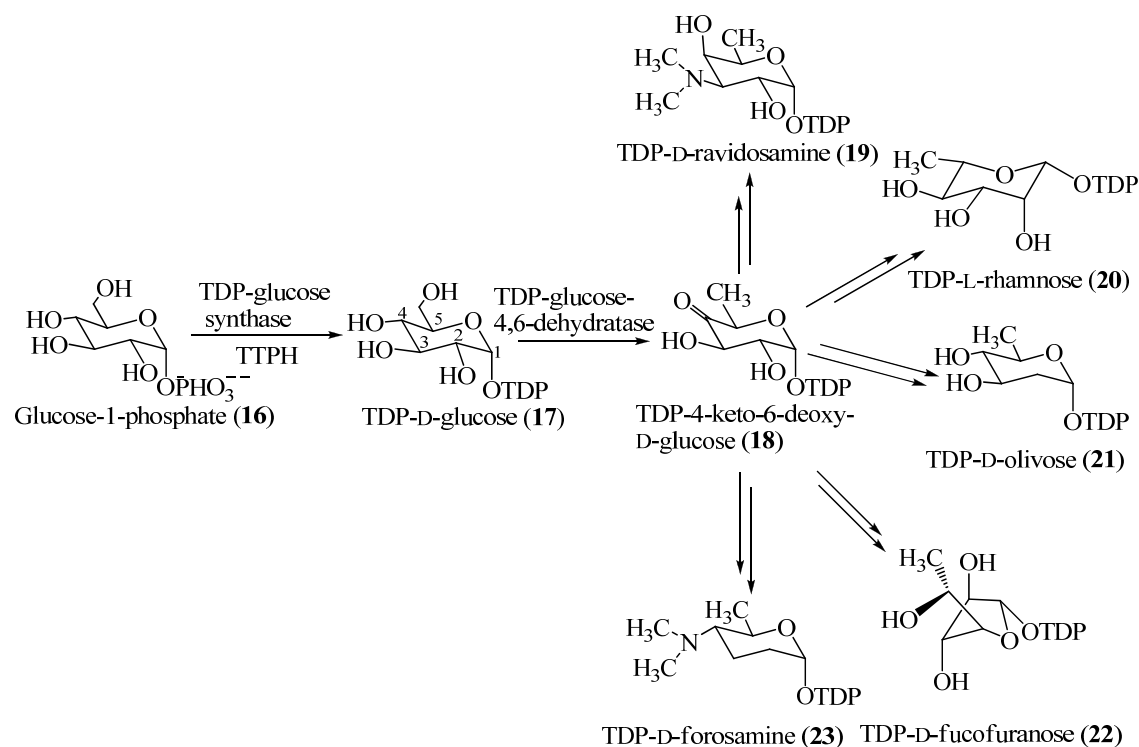
Scheme 1.9 Biosyntheses of 6-deoxysugars through the common intermediate (**18**)

Oxygenation

The products of PKSs often undergo a series of oxidations such as hydroxylations, epoxidations, anthrone oxidations, desaturations, Baeyer-Villiger monooxygenation, etc. which are usually needed for their biological activities (*21*). Typically, these reactions are catalyzed by a large family of enzymes, oxygenases. Based on the number of oxygen atoms incorporated into the substrate, they are further classified as monooxygenases (extracts one O-atom from an O_2-molecule) and dioxygenases (utilizes both O-atoms of an O_2-molecule). Most of the monooxygenases utilize reduced NADPH as reducing equivalents and catalyze hydroxylations, epoxidations or Baeyer-Villiger oxidations, whereas dioxygenases result in the production of peroxides or dioxetanes, which further rearrange or react to yield diols (**Scheme 1.10**) (*21*).
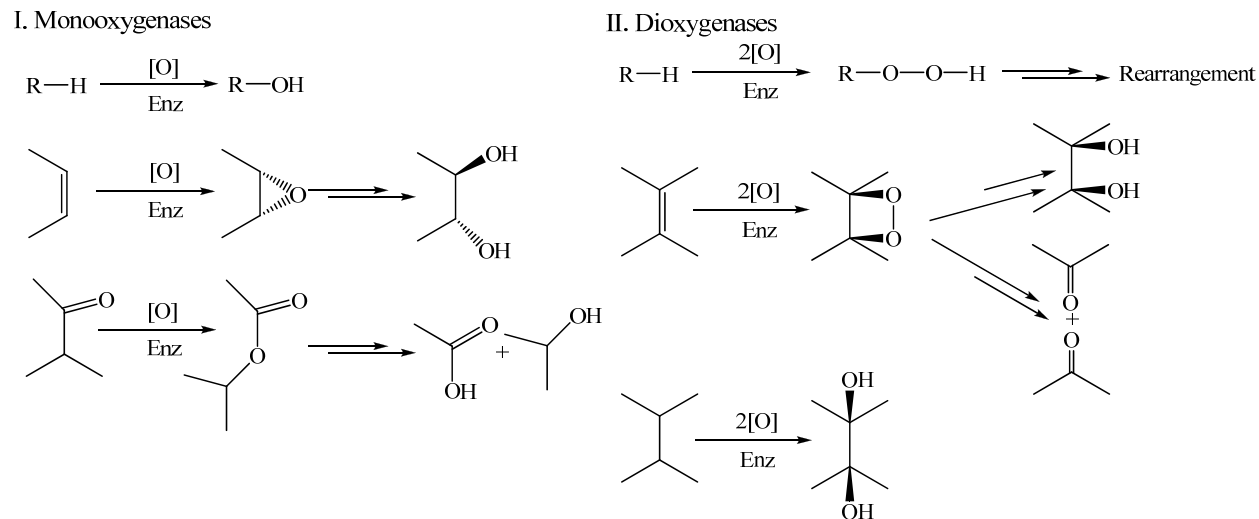
Scheme 1.10 Structural diversity introduced by oxygenases (adapted from Rohr et al. 2002) (*21*).

Cytochrome P-450 monooxygenases (CYP450s) constitute a very large and diverse family of hemoproteins. They can be one-component (reductase activity associated with the enzyme) or two-component (requires another enzyme called ferredoxin reductase) enzymes. Stereospecific hydroxylation and epoxidation are the most common CYP450 catalyzed reactions observed in post-PKS modifications. PikC-mediated hydroxylation of narbomycin (**24**) in the pikromycin (**25**) biosynthetic pathway is a good example of an CYP450 (**Scheme 1.**) (*50*). Like GTs, broader substrate specificity of CYP450s can be exploited for the generation of novel natural product analogues. This approach was successfully used to generate several analogues of pikromycin and erythromycin (*51-53*).

13
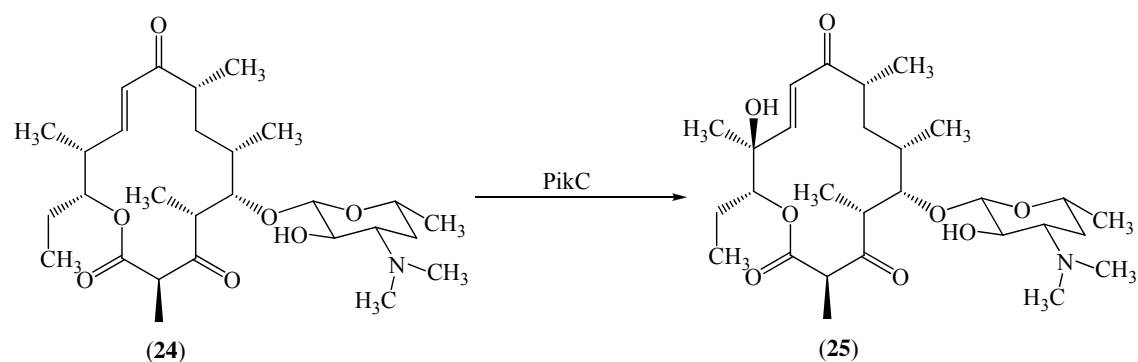
Scheme 1.11 Hydroxylation of narbomycin by PikC.

Baeyer-Villiger monooxygenases (BVMOs) involved in secondary metabolite biosynthetic pathways are rapidly growing. MtmOIV from mithramycin, (*54*) CmmOIV from chromomycin, (*55*) JadH from jadomycin, (*56*) GilOI from gilvocarcin (*57*) are proved/proposed examples of such enzymes. BVMOs are flavin-dependent and catalyze the typical Baeyer-Villiger oxidation reaction, i.e. the insertion of an oxygen atom in cyclic ketone to form a lactone. The reaction requires NADPH for reducing molecular oxygen. Despite the hundreds of reports on BVMOs and their synthetic applications, MtmOIV is the only BVMO whose natural substrate and mechanism of catalysis is characterized. MtmOV catalyzes the conversion of premithramycin B (**26**) to premithramycin B-lactone (**27**), which eventually opens up to yield mithramycin DK, the penultimate intermediate of the mithramycin biosynthetic pathway (**Scheme 1.**) (*54, 58*). Mechanistically, two electrons and two protons are transferred from NADPH to FAD generating FADH transiently. The reduced flavin immediately receives an atmospheric oxygen molecule to generate a reactive peroxyflavin ion. Nucleophilic attack of the ion on the ketone carbon followed by the migration of the C-C bond generates a lactone. Other flavin-dependent monooxygenases are PgaE, PgaM (*59*), and even cofactor-free anthrone oxygenases such as TcmH (*21*), AknX (*60*), JadG (*61*) are known. Most of these catalyze simply hydroxylations and their number is steadily growing.
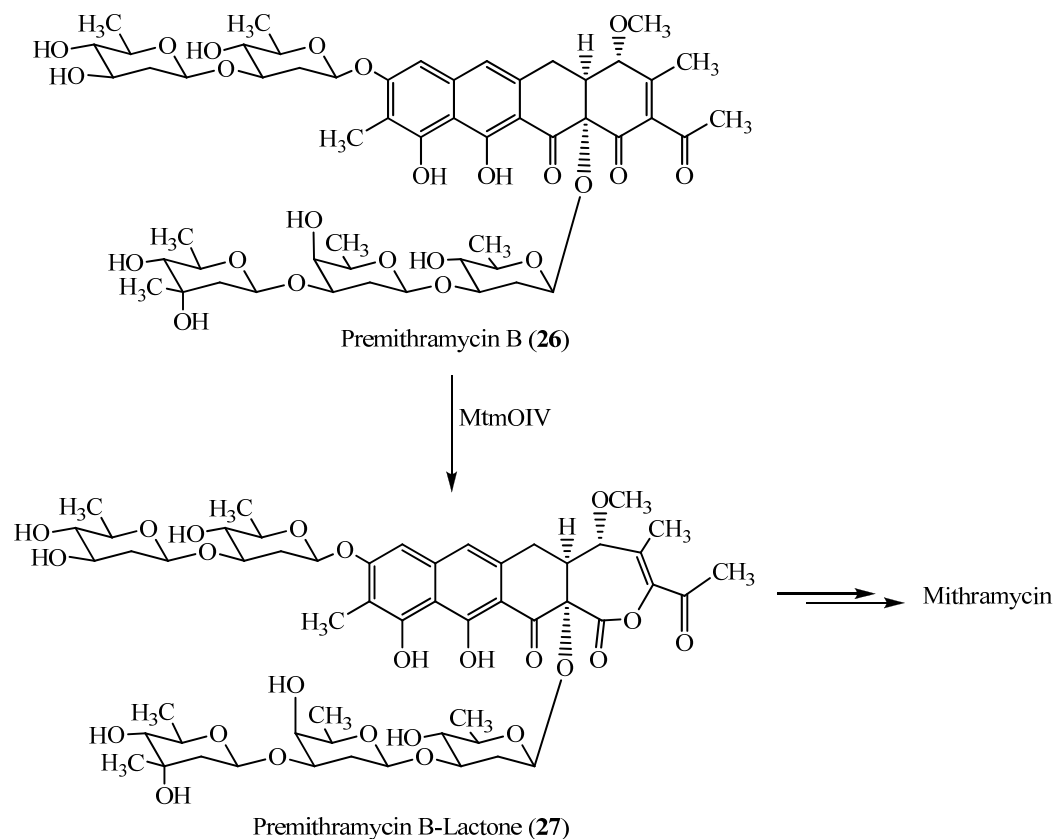
Scheme 1.12 MtmOIV-catalyzed Baeyer-Villiger oxidation of premithramycin B.
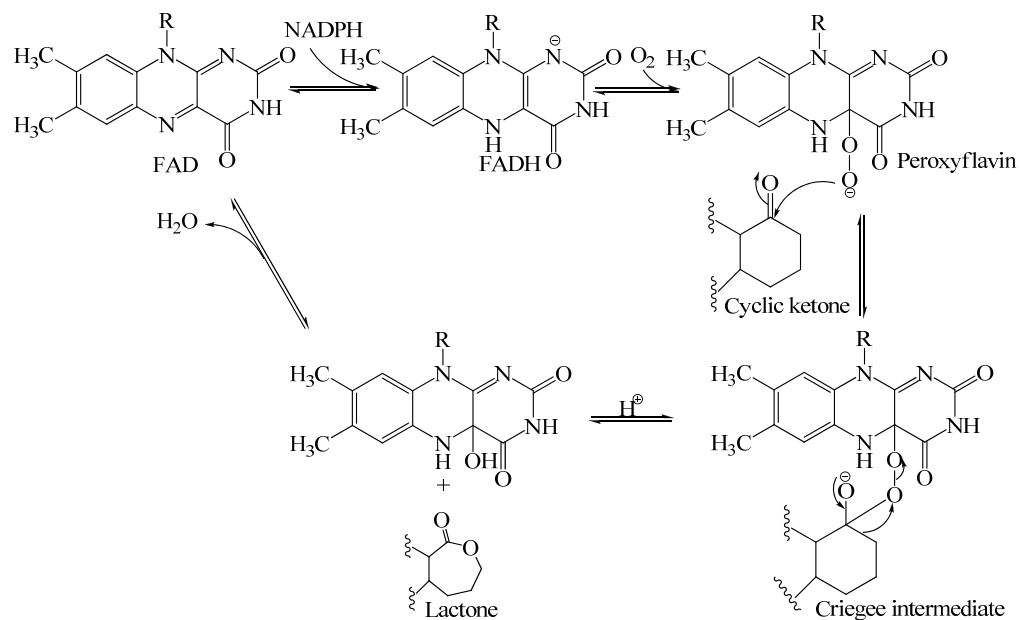
Scheme 1.13 Mechanism of BVMO catalysis.

15

Methylation

Methyltransfer reactions are among the most common post-PKS tailoring reactions, observed in the biosynthesis of a variety of bioactive microbial secondary metabolites. *N-*, *C-* or *O*-methylations of PKS products occur through the activity of corresponding methyltransferases (MTs) in a regioselective manner. These enzymes require *S*-adenosyl methionine-(SAM) as the methyl donor source, and function in various quaternary structural forms. For example, while carminomycin 4-*O*-MT (DauK) from *Streptomyces* sp. strain C5, which methylates carminomycin to daunorubicin, functions in the tetrameric form (*62, 63*), emodin *O*-MT (*64*) and tetracenomycin-*O*-MT (TcmO) (*65*) work in homohexameric and homodimeric forms, respectively. MTs are widely used for combinatorial biosynthetic approaches. Methylation of anthracyclines rhodomycin D and 13-carboxy-13-deoxycarminomycin by DauC are representative of such example (*66*). Methylations are also involved in the biosynthesis of deoxysugars such as L-mycarose, D-desosamine, etc. (*36, 67, 68*). The identification of the bifunctional methyltransferase/cyclase TcmN from the tetracenomycin biosynthetic pathway shows the evolutionary diversity of this family of enzymes (*69*).

Antitumor antibiotics gilvocarcin V and ravidomycin V

GV (**12**) produced by *Streptomyces griseoflavus* Gö 3592 represents the most important member of benzo[*d*]naptho[1,2-*b*]pyran-6-one anticancer antibiotics (*70, 71*). The compound was first isolated by Nakano et al. in 1981 from *Streptomyces gilvotanareus* (*72-74*). It exhibits high potency against tumor cell lines at low concentration with a low *in vivo* toxicity. In addition, GV serves as an antiviral and an antibacterial agent (*71,75,76*). These properties made this molecule particularly interesting for the biosynthetic investigation and for the development of its analogues with improved activity through combinatorial biosynthesis. It has been reported that GV promotes a [2+2] cycloaddition reaction of the C-8-vinyl side chain with DNA thymidine residues in addition to the inhibition of topoisomerase II (*77-79*). A recent report of GV-promoted cross-linking between DNA and histone H3 is particularly remarkable (*80*). However, the exact molecular mechanism for the *in vivo* action of GV is yet to be determined. Despite an early success on total synthesis of gilvocarcins (V, M and E), none of their analogues have ever been synthesized (*81-84*). The structural complexity of GV and a costly and laborious synthetic process could be understood as a part of the hurdle for generating new GV derivatives.

The GV producer *Streptomyces griseoflavus* also produces minor amounts of gilvocarcin M (GM, **12-b**). Both compounds have an identical tetracyclic aromatic moiety linked to D-fucofuranose through C-4 glycosidic linkage, but GM has a methyl group at C-8 position instead of the vinyl group of GV. This answers the question why GM is far less photoactive as compared to the GV indicating the crucial role of the vinyl side chain in conferring biological activity (*77*). Several other analogues of GV, such as the ravidomycins (RMs, **28, 29**), chrysomycins (CMs, **32**, **32-a**), FE35A (**30**) and B (**31**) (*85*), BE 12406 A (**38**) and B (**39**) (*86, 87*), Mer 1020 dA-D (**33-36**) (*88*) and polycarcin V (**37**) have been isolated from various actinomycete species (**Figure 1.**) (*89*). They are different from GV in their deoxysugar moieties, but members of the so-called gilvocarcin class compounds (*71*). Unlike GV (**12**) or most of the gilvocarcin class compounds, **38** and **39** have *O*-glycosidically linked sugars.

RMV (**28**) is a potent antitumor and antibacterial antibiotic isolated from the culture broth of *Streptomyces ravidus* (*90*). RMV shows a broad range of biological activities (*91*). It is about two fold more active than GV against varieties of gram-positive (MICs = ≤ 0.06 µg mL^{-1} and $0.12 - 4$ µg mL^{-1} in the presence and in the absence of visible light, respectively) and gram negative (MICs = 64 - >128 µg mL^{-1} and ≥ 0.5 µg mL^{-1} in the presence and in the absence of visible light, respectively) bacteria (*92*). Such activities could be attributed to RMV's role in inhibiting DNA and RNA synthesis. RM shows remarkable activity against P38 leukemia cell lines both *in vitro* and *in vivo* in mice (*85, 93*).
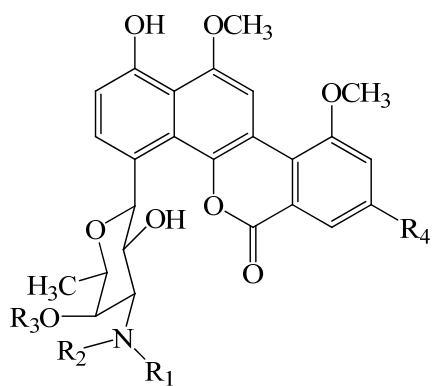
28 : $R_1=CH_3$, $R_2=CH_3$, $R_3=COCH_3$, $R_4=CH=CH_2$
29 : $R_1=CH_3$, $R_2=CH_3$, $R_3=H$, $R_4=CH=CH_2$
29-a: $R_1=CH_3$, $R_2=CH_3$, $R_3=H$, $R_4=CH_3$
30 : $R_1=CH_3$, $R_2=COCH_3$, $R_3=H$, $R_4=CH=CH_2$
31 : $R_1=H$, $R_2 = COCH_3$, $R_3=COCH_3$, $R_4=CH=CH_2$
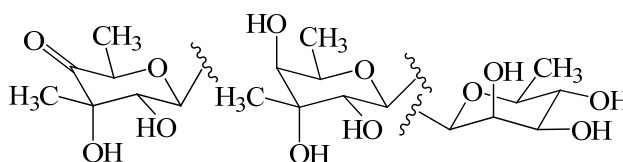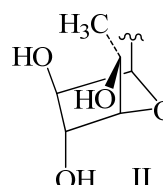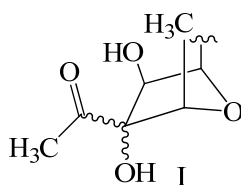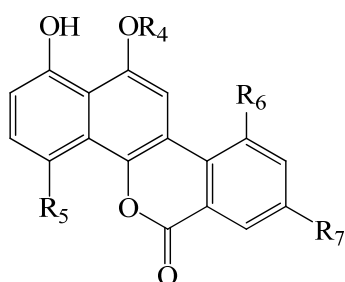
12-b : $R_7=R_4=CH_3$, $R_5=$**II**, $R_6=OCH_3$
12-a : $R_7=CH_2$-CH_3, $R_4=CH_3$, $R_5=$**II**, $R_6=OCH_3$
12 : $R_7=CH=CH_2$, $R_4=CH_3$, $R_5=$**II**, $R_6=OCH_3$
32-a : $R_7=R_4=CH_3$, $R_5=$ **V**, $R_6=OCH_3$
32 : $R_7 = CH=CH_2$, $R_4=CH_3$, $R_5=$ **V**, $R_6=OCH_3$
33 : $R_7=R_4=CH_3$, $R_5=$**I**, $R_6=OCH_3$

34 : $R_7=R_4=CH_3$, $R_5=$**III**, $R_6=OCH_3$
35 : $R_7=CH=CH_2$, $R_4=CH_3$, $R_5=$**III** $R_6=OCH_3$
36 : $R_7=CH=CH_2$, $R_4=CH_3$, $R_5=$**I**, $R_6=OCH_3$
37 : $R_7=CH=CH_2$, $R_4=CH_3$, $R_5=$**VI**, $R_6=OCH_3$
38 : $R_7=CH_3$, $R_4=$**VI**, $R_5=H$, $R_6=OCH_3$
39 : $R_7=CH_3$, $R_4=$**VI**, $R_5=H$, $R_6=OH$

Figure 1.3 Gilvocarcins, ravidomycins and related antibiotics.

However, cytotoxicity of RMV (**28**) was reported for human colon carcinoma cell lines (*92*). In addition, RMV (**28**) has been reported as a potent inhibitor of human topoisomerase II (*94*). Like GV, the exact mechanism of action of RMV at the molecular level is still unclear. However, photoactivated antitumor and antibacterial activities of GV (**12**) and RMV (**28**), and their common coumarin-based backbone indicate their similar mechanism of actions. Although, there is no direct experimental report on activity comparison of RMV with the other analogs, reports in the literature indicated RMV to be the most active compound against tumor cell lines (*93*). This is a clear indication that the more hydrophilic amino sugar moiety has a crucial role in the biological activities. Such amino compounds also provide better formulation and delivery

options in their salt forms (*95*). Structure activity relationship (SAR) of the RMV analogues generated through chemical modification suggested that 2′- and 4′-hydroxyl groups of the D-ravidosamine subunit are essential for the biological activities. However, its inherent cytotoxicity hindered its further development as an anticancer drug necessitating the generation of new analogues with better pharmacological profiles.

Biosynthetic investigations of gilvocarcin V and ravidomycin V

Incorporation studies of gilvocarcins, ravidomycins and chrysomycins showed that the unique backbone originates from acetate and propionate, and the pathway involves a C-C bond cleavage and an oxidative rearrangement of an angucyclinone (*96-98*). However, the exact intermediates involved, mode of oxidative rearrangement and other post-PKS tailoring reactions were largely unclear. Recently, Rohr et al. have isolated the complete gene cluster for GV biosynthesis in a single cosmid, namely cosG9B3. Heterologous production of GV in *S. lividans* TK24 clearly showed that the cosmid harbors the entire GV biosynthetic locus (*71*). The sequencing of the cluster revealed the presence of anticipated PKS genes (g*ilABCFKPQ*), four putative oxygenases (encoded by *gilOI, OII, OIII, OIV*), a *C*-glycosylransferase (encoded by *gilGT*), a methyltransferase gene (*gilMT*), an oxidoraductase gene (*gilR*) and several other genes of unknown function such as *gilNLVM* besides putative repressor and resistance genes (**Figure 1.**).
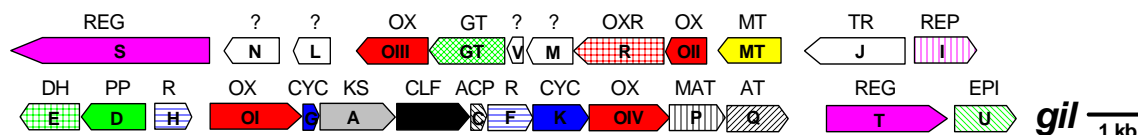
Figure 1.4 GV biosynthetic gene cluster (the figure was taken directly from Rohr et al. 2003) (*71*).

In an effort to figure out the mode of the oxidative C-C bond cleavage, all of the pathway oxygenases were inactivated in cosG9B3 individually or sequentially. Thus, the generated mutant cosmids were heterologously expressed in *S. lividans* TK24. Homo-rabelomycin (**48**) and 2,3-dehydro-homo-UWM6 (**41**) were identified as major metabolites from the culture broth of ΔGilOIV or ΔGilOIVΔGilOII and ΔGilOIΔGilOIIΔGilOIV or ΔGilOI mutants, respectively (**Scheme 1.**) (*56,57*). Similarly dehydro-homo-rabelomycin (**49**) was isolated from a ΔGilOII mutant. In none of these compounds, a C-C bond cleavage occurred,

which indicated that all of the oxygenases are necessary for the C5-C6 bond cleavage of the hypothetical intermediate **42**. Identification of similar compounds through the inactivation of homologous oxygenases of the jadomycin pathway showed that both jadomycin A (**44-c**) and GV (**12**) biosynthetic pathways share common intermediates prior to their C-C-bond cleavages (**Scheme 1.**) (*30, 61*). Interestingly, a cross complementation of the ΔGilOI mutant strain with JadH and of the Δ*gilOIV* mutant with JadF restored the production of GV. However, GV production was not restored when *jadG* was expressed into the ΔGilOII strain (*56*). This indicated that both pathways share at least two functionally identical oxygenases and might have identical modes of oxidative C-C bond cleavage. GilOII and JadG might have functions unique to the GV and jadomycin pathways, respectively. Identification of a GV biosynthetic intermediate, pregilvocarcin V (**47**), from the ΔGilR mutant provided indirect proof that the GV biosynthesis involves an acid-aldehyde intermediate (**44**). While non-enzymatic incorporation of L-isoleucine to this intermediate, followed by spontaneous decarboxylation and oxidation resulted in the production of jadomycin A (**Scheme 1.**) (*99, 100*), C-C bond rotation, *O*-methylations, vinyl side chain formation, glycosylation, and oxidation were necessary to generate GV (**Scheme 1.**). However, the exact mechanism of such oxidative ring opening for both pathways was remain undetermined.

Rohr et al. further inactivated several other genes and analyzed the metabolites accumulated by the resulting mutants. Inactivation of GilGT completely halted the production of GV and accumulated defuco-GV (**50**) (*33*). The production of GV was restored by complementing the mutant with a GilGT-expression plasmid. Similarly, gilvocarcin E (**12-a**) was accumulated as the major metabolite of the ΔGilOIII mutant. These results clearly showed that GilGT and GilOIII were involved in *C*-glycosylation and C-8-vinyl side chain formation, respectively.
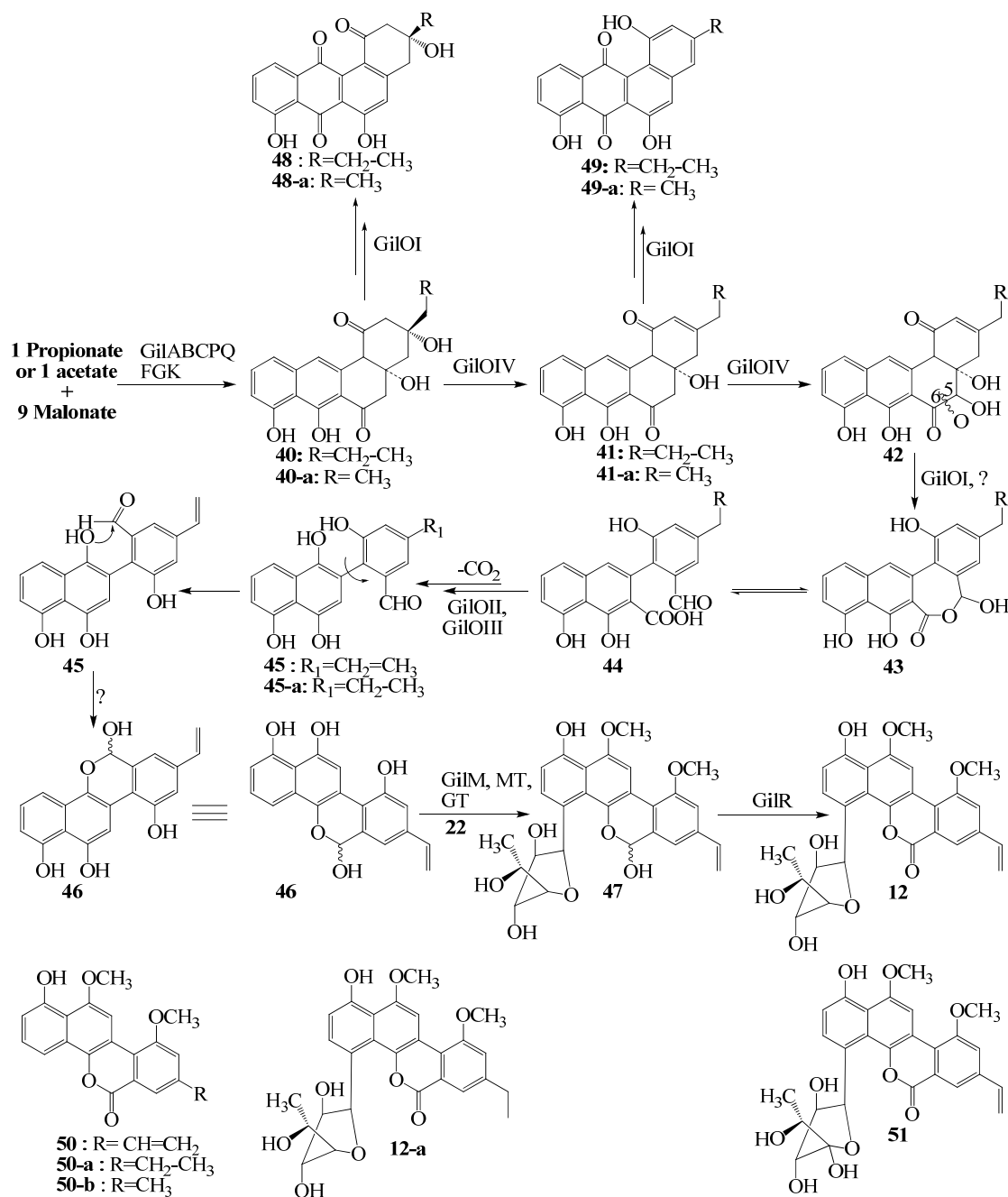
Scheme 1.14 Proposed pathway for GV biosynthesis.
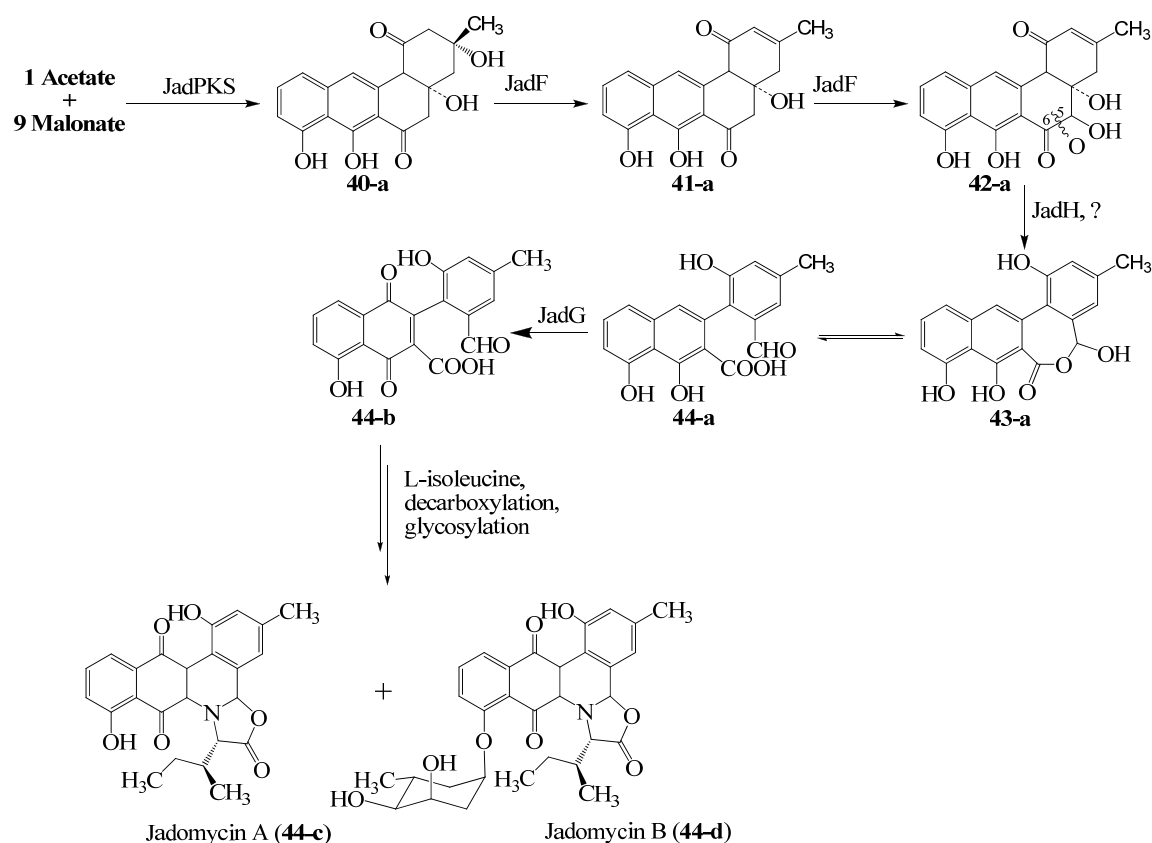
Scheme 1.15 Biosynthetic pathway of jadomycins A and B.

None of the compounds isolated from the mutant strains were converted into GV when fed to a culture of the ΔGilOIV mutant. This indicated that the accumulated compounds were biosynthetic shunt products, thus, making the substrate of GilGT and GilOIII elusive. Gilvocarcin V production was completely abolished when *gilM* (unknown gene) was inactivated and the production was restored when the mutant was complemented with a *gilM*-expression construct (*101*). The mutant accumulated the shunt product, homo-rabelomycin (**48**), as the major metabolite and also produced an other yet uncharacterized minor compound. This clearly showed that GilM was a crucial enzyme for the GV biosynthesis, although its exact functional role was not clear. Expression of *gilMT*-deleted cosmid in *S. lividans* TK24 did not accumulate any metabolites. However, GilM and GilMT were proposed to catalyze C-C bond rotation and methylation of phenolic hydroxyl groups (**45**→**47**), respectively.

D-fucofuranose is rarely found in glycosylated natural products. Its biosynthesis was of particular interest for its unique ring-contraction chemistry as this sugar residue was crucial for

the biological activity of the GV. Inactivation of one of the potential deoxysugar biosynthetic gene *gilU* (ketoreductase) resulted in the accumulation of a new GV derivative, 4′-hydroxy-GV (**51**), with an improved biological activity (*102*). The results provided a clear proof that GilU stereospecifically reduced the 4-keto group of **18** (**Scheme 1.9**) to generate TDP-D-fucose, which further undergoes yet unknown ring-contraction and glycosylation steps to yield the D-fucofuranose moiety of GV.

Summary

Polyketides constitute one of the largest group of natural products with diverse structures and biological activities. Structural and functional knowledge gained about their biosynthetic machineries at the genetic and molecular levels over the past few decades has made the nucleotide sequence based rational design of new polyketides possible. In this context, generation of thousands of new natural products through combinatorial biosynthesis has largely been appreciated, which is otherwise not possible through traditional/modern chemical synthesis. With an aim to develop a photo-activated antitumor drug, our lab has been focusing on the biosynthesis of GV and combinatorial biosynthetic approach to generate new GV derivatives for structure activity relationships (SAR) studies. In this context, GV biosynthesis has been extensively studied through gene inactivation experiments. Analyses of metabolites from various mutants produced several new gilvocarcin derivatives/intermediates. This helped outline the functional roles of the individual gene products. However, cross feeding experiments showed that none of the isolated compounds appeared to be GV biosynthetic intermediates, except pre-GV (**47**, the principal metabolite of the GilR-minus mutant) thus, making the exact biosynthetic intermediates and the sequence of biosynthetic events elusive. Given the fact that combinatorial biosynthesis of new derivatives requires an in depth understanding of the involved biosynthetic enzymes and their orchestrations, this study was focused on characterizing individual GV biosynthetic enzymes *in vitro*. The results would provide more insights into the intriguing chemistry involved in GV biosynthesis, especially in the post-PKS tailoring events, such as oxidative rearrangement, *C*-glycosylation, *O*-methylation, lactone formation etc. As discussed earlier, glycodiversification provides a unique means to generate GV/RMV analogues, but this requires flexible glycosyltransferases (GTs) and a library of NDP-sugars (*42,43*). Unfortunately, there was only a handful of *C*-GTs known that have marginal substrate promiscuity (*32,44,102*)

and no other *C*-GT that was capable of transferring amino or branched sugars were reported so far, except the Med-ORF8 which utilizes TDP-angolosamine as the donor substrate (*103*). In this context, cloning, sequencing and characterization of the ravidomycin biosynthetic gene cluster could provide an additional opportunity to expand the combinatorial biosynthetic toolbox to further investigate or to design/generate GV/RMV analogues through combinatorial biosynthesis. Comparison of the two (GV/RMV) gene clusters would also help sort out the genes, whose products catalyze the chemistry unique to each pathway, for instance, the genes involved in contracting the D-furopyranose to D-fucofuranose moiety of GV.

Specific aims

The objectives of this study were to elucidate the functional role of post-PKS tailoring enzymes involved in the biosynthesis of GV (**12**) and RMV (**28**) and to evaluate their substrate specificities; a prerequisite for the successful utilization of a combinatorial biosynthesis approach towards generating new GV analogues. To achieve this goal, the following three specific aims were addressed:

Specific aim 1: Clone and characterize RMV biosynthetic gene cluster.

Through the comparison of the GV and RMV gene cluster, genes responsible for the biosynthesis of the sugar and the tetracyclic aromatic moieties can be identified for both compounds. In addition, ravidomycin biosynthetic enzymes could provide additional tools to generate GV/RMV analogues through combinatorial biosynthesis.

Specific aim 2: Characterize all of the 4-*O*-acetyl-TDP-D-ravidosamine biosynthetic enzymes.

Characterization of the putative 4-*O*-acetyl-TDP-D-ravidosamine biosynthetic genes can provide valuable insights into the biosynthetic pathway of the sugar moiety of RMV. Successful reconstitution of the sugar pathway *in vitro* could provide a means to generate the natural sugar donor substrate of the RMV *C*-GT for *in vitro* evaluation. In addition, the intermediates of the 4-*O*-acetyl-TDP-D-ravidosamine biosynthetic pathway can be tested as sugar donor substrates for RavGT or GilGT.

Specific aim 3. Demonstrate a total enzymatic synthesis of defucogilvocarcin M (defucoGM), the common polyketide-derived core of GV and RMV.

This aim can identify a minimal set of enzymes required to biosynthesize defucoGM from the anticipated pathway intermediate prejadomycin. Alternatively, a successful reconstitution of the GilPKS and post-PKS enzymes could demonstrate the biosynthesis of relatively complex defucoGM from simple building blocks, such as acetate and malonate. The results would provide more insights into the GV post-PKS tailoring reactions and will also provide a framework for the further study of substrate specificities of the involved enzymes.

Section Two: Cloning, sequencing and characterization of ravidomycin biosynthetic gene cluster

Results

Cloning, sequence analysis and organization of the ravidomycin gene cluster

A library screening strategy was taken to clone the ravidomycin (RMV) biosynthetic gene cluster. A genomic library of *Streptomyces ravidus* (*S. ravidus*) was constructed in the pOJ446 *E.coli-Streptomyces* shuttle vector (*104*). Like gilvocarcin V (GV), biosyntheses of the polyketide core and the sugar moiety of RMV (**28**) required type II PKS enzyme keto-acyl synthase α (KS$_α$) and a deoxysugar biosynthetic enzyme NDP-sugar 4,6-dehydratase, respectively. Therefore, their corresponding genes were used to probe the genomic library. The primers based on conserved residues of *gilA* (encodes KS$_α$ in GV pathway) (*71*), and its homologues from the National Center for Biotechnology Information (NCBI) nucleotide sequence database resulted in the expected amplifications from *S. ravidus* genomic DNA. Similarly, a part of NDP-glucose-4,6-dehydratase gene was amplified using the previously reported forward and a newly generated reversese primers (see appendices) (*105*). The cosmid library was screened for positive colonies using a digoxigenin (DIG)-labelled KS$_α$ gene probe. The positive colonies were then subjected to a second round of hybridization with DIG-labeled TDP-glucose 4,6-dehydratase gene probe (**Figure 2.1**). The hybridization was further confirmed through southern blot analysis (**Figure 2.2**). Out of approximately 2000 colonies screened, four different colonies from the *S. ravidus* library hybridized with both probes. However, restriction analyses of the cosmid DNA from two colonies revealed identical digestion pattern. One of the positive cosmids namely cosRav32 was sequenced via a shotgun approach. Bioinformatics analyses of the cosRav32 sequence in public databases revealed a cluster of 35 open reading frames (ORFs). The ravidomycin gene cluster was flanked by ORFs encoding functions obviously irrelevant for RMV (**28**) biosynthesis. The organization of the cluster and the putative function of the genes are summarized in **Figure 2.3** and **Table 2.1**, respectively. The nucleotide sequence data have been deposited in the EMBL nucleotide database under the accession number FN565485.

ketoacyl synthase gene probe   4,6-dehydratase gene probe   ketoacyl synthase gene probe   4,6-dehydratase gene probe
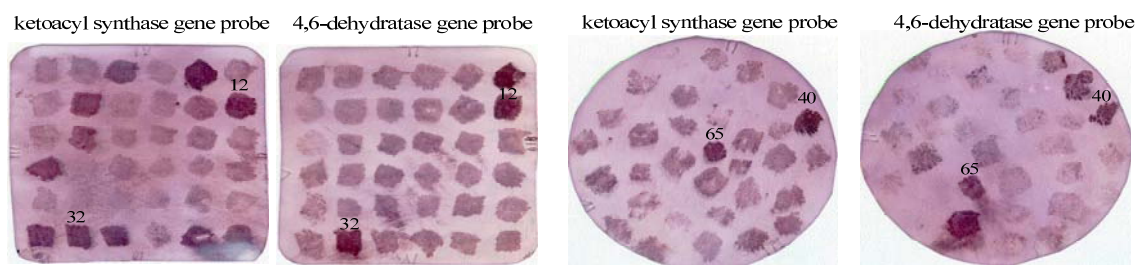
Figure 2.1 Colony hybridization of *S. ravidus* genomic library using KS_α and TDP-glucose-4,6-dehydratase gene probes. Twelve colonies including no. 12, 32, 65 and 40 were hybridized with the both probes.

I. 4,6-dehydratase probe

A   12 32 40 65 60 62 63 67 164 4 18 44

B   12 32 40 65 60 62 63 67 164 4 18 44

II. Ketoacyl synthase probe

C   12 32 40 65

D   12 32 40 65
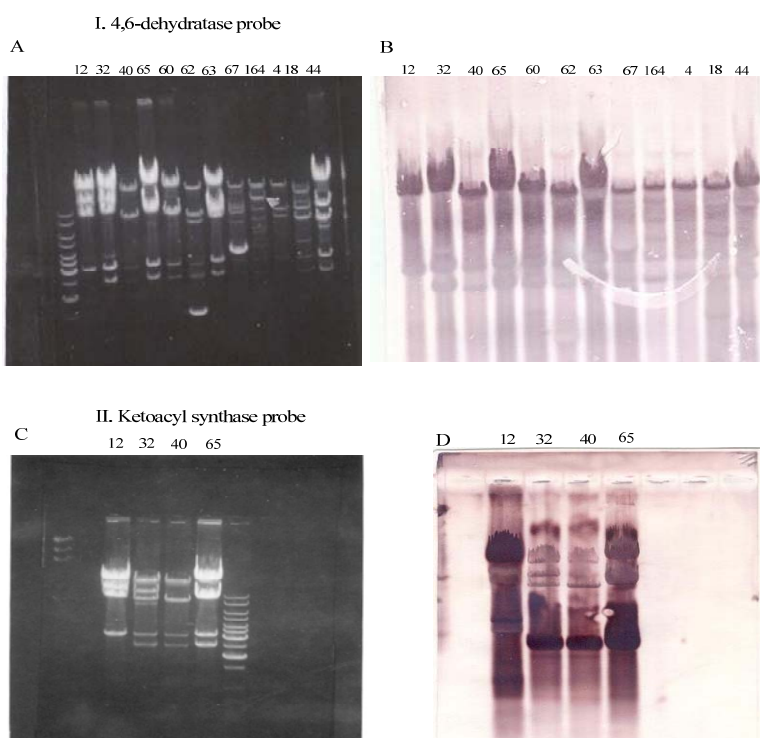
Figure 2.2 Southern blot of hybridization of the isolated cosmids with KS_α and 4,6-dehydratase gene probes. Numbers on the top of the figures represents the DNA sample isolated from the corresponding no. of positive *E. coli* colony. A: hybridization of *Bam*HI digested cosmid DNA with 4,6-dehydratase gene probes. B: hybridization of *Bam*HI digested cosmid DNA with 4,6-dehydratase gene probes. DNA samples from colony 12, 32, 40 and 65 were selected for the hybridization with KS_α gene probe.

27

Biosynthesis of the polyketide backbone

Nine genes in the RMV cluster were identified to be involved in the biosynthesis of the polyketide-derived backbone. The encoded proteins of *ravA*, *ravB* and *ravC* displayed typical homologies to the various minimal PKS enzymes [ketoacyl synthases (KS$_\alpha$), chain length determinants (KS$_\beta$ or CLF) and the acyl carrier protein (ACP), respectively], typical of a type II PKS system and were flanked by two oxygenase genes (*ravOIII* and *ravOIV*). RavA and RavB exhibited 79% and 78% identity to UrdA (KS$_\alpha$) and JadA (KS$_\beta$) from the urdamycin and jadomycin biosynthetic pathways, respectively (*106, 107*).
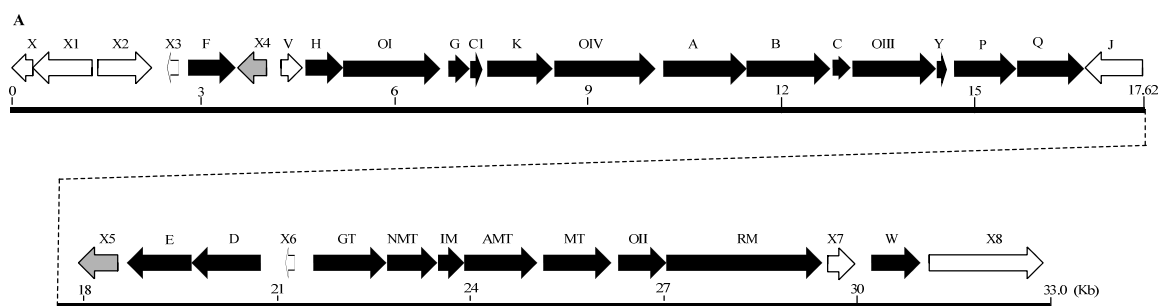
Figure 2.3 A gene cluster for ravidomycin biosynthesis: the black filled arrows indicate biosynthetic genes; gray arrows represent regulatory genes and blank arrows represent genes of enzymes with unknown functions.

Like other Ks$_\alpha$/Ks$_\beta$ heterodimers, RavA contained a conserved amino acid motif STG**C**TSGLD including the active site cysteine (*108*), but RavB had a glycine residue in place of the cysteine. The product of *ravC* displayed 86% identity to JadC from the jadomycin pathway (*107*). Interestingly, an additional ACP encoded by (*ravC1*) was also found in between two cyclase genes (*ravG* and *ravK*). The encoded amino acid sequence of *ravC1* exhibited 68% and 61% identities to KinC (AH012623) and JadC from the kinamycin and jadomycin pathways, respectively. Both ACPs had a conserved LG(X)D**S** motif including the serine residue for attachment of the 4′-phosphopanthetheine arm (*109*). Despite their close homologies, these ACPs might be specific for the selection and priming of a starter unit, either acetate or propionate. In contrast, only one ACP has been reported for many other well characterized type II PKS clusters including those of aclacinomycin A (*110*), doxorubicin (*111*) and even the most

28

closely related GV (*71*). Unlike other typical type II PKS clusters, the putative ketoreductase gene *ravF* was stand alone at the end of the RMV cluster. The deduced amino acids of *ravF* exhibited strong similarities to various angucycline PKS associated ketoreductases such as SimA6 (77% identity), UrdD (72%) and AknA (72%) of the simocyclinone (*112*), urdamycin A (*106*) and aklavinone (*113*) biosynthetic pathways, respectively. Therefore, we hypothesized that RavF reduces the 9-keto position of the growing polyketide chain to generate a secondary alcohol which later dehydrates during cyclization and aromatization of the first ring. Two putative cyclase genes, *ravK* and *ravG*, were separated by *ravC1*. The product of *ravG* displayed 76% amino acid identitiy to the cyclase (Schp9) of Sch 47554 biosynthetic pathway (*114*), whereas the product of *ravK* showed 61%, 60% and 58% identities to the bifunctional dehydratases/cyclases from the oviedomycin (OvmA) (*115*), Sch 47554 (SchP4) (*114*) and griseusin (Orf4) (*116*) biosynthetic pathways, respectively. The homologues of *ravG* and *ravK* were also found in the GV cluster. Based on the amino acid similarities, we hypothesized that RavG and RavK are involved in several cyclizations and dehydrations steps required to install an aromatic angucyclinone intermediate (**40**). Located downstream of *ravY*, genes *ravP* and *ravQ* were translationally coupled and their encoded proteins exhibited similarity to malonyl CoA:ACP transacylase (MAT) and acyltransferase (AT), respectively. Given the fact that Type II PKS often recruits MAT from the fatty acid pathway (*117*), the presence of two such enzymes in the RMV gene cluster was unique. However, there was such precedence in the GV biosynthetic cluster as well. Homology search showed that RavP was similar to various type II PKS-associated MATs such as PgaH (*118*) (51% identity), Aur1M (*119*) (49%) and GilP (*71*) (45%). These homologous enzymes assist loading of the starter/extender unit, preferentially malonyl-CoA to the ACP. However, the product of *ravQ* resembled various other ATs such as AknF (*110*) (45% identity), FrnK (*120*) (46%) and EncL (*121*) (40%) DpsD (*122*) and GilQ (39%), which were involved in the biosyntheses of aromatic polyketides with the starter units other than acetate. Therefore, we favored RavQ to be responsible for the selection of propionate as a starter unit, which turns into the vinyl side chain of **28**. GilP might self load and then transfer acetyl-CoA to the ACPs RavC/RavC1.

Table 2.1 Orfs of the RMV gene cluster and proposed function of their corresponding enzymes

Orfs	Size[a] of corresponding enzyme	ID/SM (%)[b]	Closest homologous enzyme	Nucleotide accession no.	Proposed function of product
ravX	106	94/100	PaaB, *S. clavuligerus*	ZP_03184882	Phenylacetic acid degradation protein
ravX1	325	92/94	PaaA, *S. clavuligerus*	ZP_03184883	Phenylacetate-CoA oxygenase
ravX2	270	65/79	*S. pristinaespiralis*	YP_002198855	Unknown
ravX3	65	-	-	-	Hypothetical protein
ravF	261	77/87	SimA6, *S. antibioticus*	AAK06787	Ketoreductase
ravX4	161	66/79	*S. griseus*	YP_001826336	Regulatory protein
ravV	107	30/48	GilV, *S. griseoflavus*	ABE03981	Hypothetical protein
ravH	189	55/67	GilH, *S. griseoflavus*	AAP69587	NADPH-dependent FMN reductase
ravOI	503	57/67	JadH, *S. venezuelae*	AAV52248	Monooxygenase
ravG	108	76/83	schP9, *S.* sp. SCC 2136	CAH10118	Cyclase
ravC1	82	68/79	KinC, *S. murayamaensis*	AAO65348	Acyl carrier protein
ravK	315		ovmA, *S. antibioticus*	CAG14969	Aromatase/cyclase
ravOIV	520	48/58	lndM2, *S. globisporus*	AAT64432	Monooxygenase
ravA	423	79/86	UrdA, *S. fradiae*	CAA60569	Ketoacyl synthase α (KS α)
ravB	401	71/84	*S. venezuelae*	AAB36563	Ketoacyl synthase *β* (KS *β*)
ravC	86	62/75	*S. venezuelae*	AAB36564	Acyl carrier protein
ravOIII	394	70/80	GilOIII, *S. griseoflavus*	AAP69584	Cytochrome P-450 monooxygenase
ravY	72	43/63	*Nocardia farcinica* IFM 10152	YP_116742	Ferredoxin
ravP	323	51/64	PgaH, *S.* sp. PGA64	AAK57533	Malonyl-CoA acyl carrier protein transacylase

ravQ	330	45/55	AknF, *S.galilaeus*	BAB72049	Acyltransferase
ravJ	289	42/53	SchP11, *S.* sp. SCC 2136	CAH10120	Unknown
ravX5	208	56/74	*S. violaceoruber*	CAA09631	Regulatory protein
ravE	328	77/85	*S. sviceus*	YP_002205670	NDP-glucose 4,6-dehydratase
ravD	355	79/88	*S. sviceus*	YP_002205669	NDP-glucose synthase
ravX6	55	-	-	-	Hypothetical protein
ravGT	379	46/66	GilGT, *S. griseoflavus*	AAP69578	*C*-glycosyltransferase
ravNMT	265	56/71	SnogX, *S. nogalater*	CAA12020	*N,N*-dimethyltransferase
ravIM	156	47/69	*Burkholderia phymatum*	YP_001858530	NDP-6-deoxy-3, 4-keto-hexulose isomerase
ravAMT	369	60/75	*Saccharopolyspora erythraea*	YP_001103001	Aminotransferase
ravMT	339	51/66	*S. tubercidicus*	AAT45298	*O*-methyltransferase
ravOII	231	68/77	GilOII, *S. griseoflavus*	AAP69583	Anthrone Monooxygenase
ravRM	776				
	N-terminal 535 aa,	54/67;	*S. galilaeus*	ABI15166	Dehydrogenase
	C-terminal, 241 aa	41/58	GilM, *S. griseoflavus*	AAP69591	Unknown
ravX7	126	-	-	-	Hypothetical protein
ravW	247	37/46	*Arthrobacter aurescens* TC1	YP_950015	*N*-acetyltransferases
ravX8	564	59/74	*S. griseus*	YP_001827711	Hypothetical protein

[a]: number of amino acids; [b]: identity/similarity; aa: amino acids
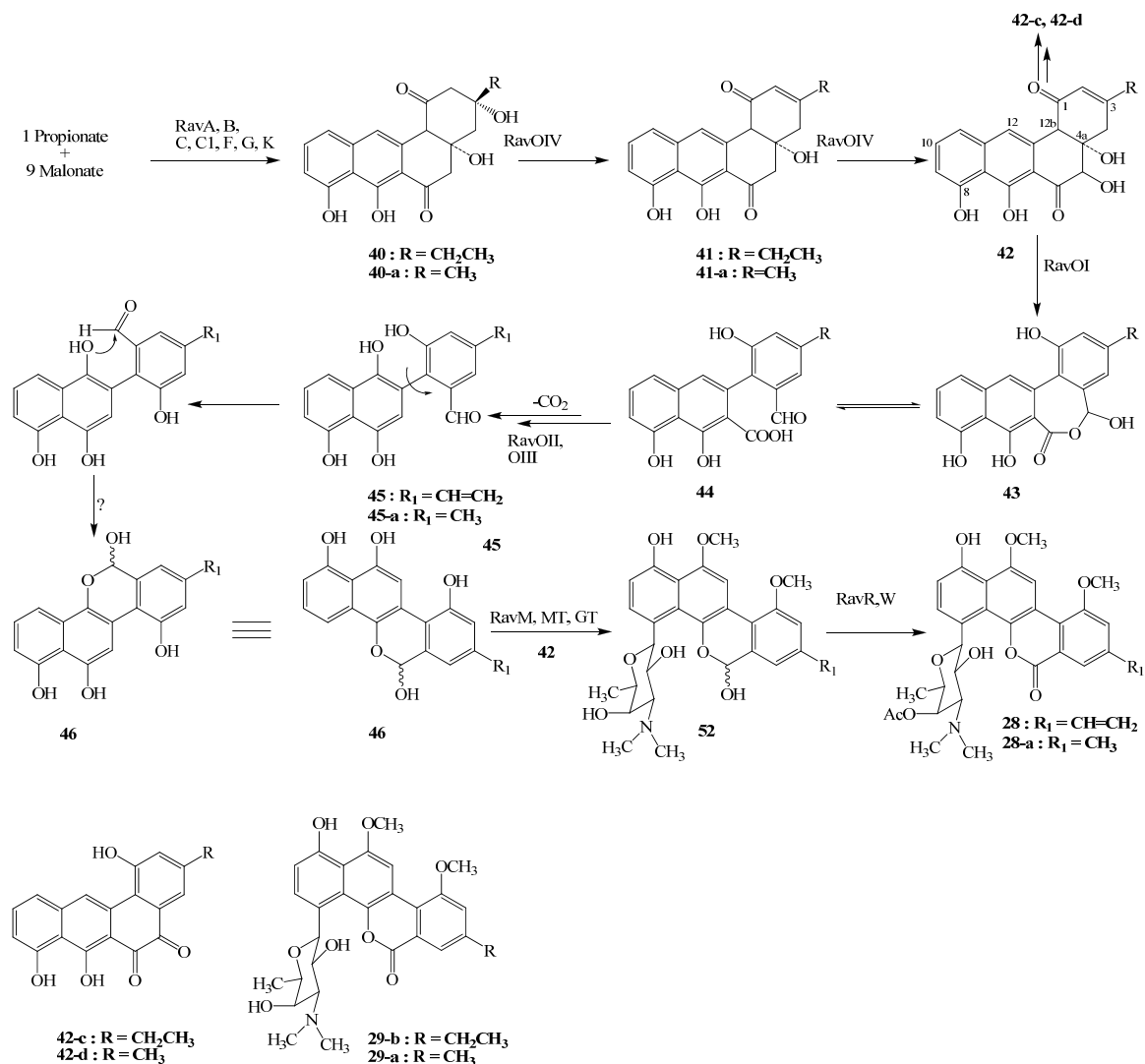
Scheme 2.1 Proposed pathway for the biosynthesis of RMV (**28**), involving enzymes encoded by the cloned RMV gene cluster

Post-PKS tailoring enzymes

As expected, the post-PKS tailoring enzymes encoded by the RMV cluster were similar to those of the GV counterpart. Four different oxygenases (encoded by *ravOI, OII, OIII* and *OIV*) were scattered throughout the cluster. Homology searches in the databases showed that the products of *ravOI* and *ravOIV* resembled the flavin adenine dinucleotide (FAD)-dependent monooxygenases. RavOIV displayed its identity/similarity to LndM2 (48%/58%) from the landomycin, OvmOII (47%/56%) from the oviedomycin,(*123*) PgaM (46%/58%) from the

gaudimycin,(*59*) JadF (46%/55%) and JadH (42%/54%) from the jadomycin(*61*) and GilOI (35%/47%) and GilOIV (29.8%) from the GV pathways (*57*). Similarly, SimA7 (*124*) (61%/73%), OvmOI (*115*) (60%/72%), PgaE (*125*) (58%/71%), JadH (*61*) (57%/67%), GilOI (*71*) (42%/55%) and GilOIV (30% identity) were the closer homologues of RavOI. Here, it was important to note that most of these homologues were proposed or characterized to have dual oxygenation and dehydration activities, and they all act on angucyclinones. Recent results showed that both JadF and GilOIV catalyze identical reactions, i.e. 2,3-dehydration of **40** and hydroxylation at the C-5 position of **41** despite the very different pathway end products (**scheme 2.1**) (*56*). Similarly JadH and GilOI were also shown to have identical activities: both catalyze 4a,12b dehydration of **42**, and may also be involved in C-C bond cleavage presumably through Baeyer-Villiger monooxygenation of the C-6 carbonyl. Based on these similarity studies, we proposed that RavOIV and RavOI were functionally identical to GilOIV and GilOI of the GV pathway, respectively. The start codon of *ravOI* was transcriptionally coupled with *ravH* that encodes a 189 aa-protein showing its similarity (55%-61% identity) to the various NADPH-dependent FAD/FMN reductases including the recently characterized GV pathway counterpart GilH (*102*). We assumed that RavH might generate the reduced coenzyme (FADH) from NADPH and FAD to provide reducing equivalents for the oxidative cascade catalyzed by RavOI or RavOIV. The encoded sequence of *ravOII* showed its closest homology to GilOII (68%/77%) of the GV pathway besides other putative anthrone oxygenases. Like GilOII in the GV pathway, RavOII might introduce a hydroxyl group at the naphthalene moiety of **44** which then would react with the aldehyde to generate a hemiacetal **46**. The encoded product of *ravOIII* exhibited similarity to various cytochrome-P450 monooxygenases with GilOIII of the GV pathway being the closest homologues (70%/80%). Through inactivation of the corresponding gene, we previously showed that GilOIII is involved in the vinyl side chain formation of GV (*33*). Therefore, a similar role of RavOIII in RMV biosynthesis could be safely assigned. Careful sequence analyses revealed a small ORF *ravY* whose start codon was translationally coupled to the stop codon of *ravOIII*. The product of *ravY* shared moderate identity (40%-50%) to various putative ferredoxin proteins in the database. Since most of the cytochrome-P450 monooxygenases require ferredoxin as an electron carrier to fulfill their oxidative activities, we propose that the product of *ravY* might serve as an electron carrier for RavOIII. The lack of such

a gene in the GV cluster was not surprising as there were many such precedents, where ferredoxin proteins are located outside the secondary metabolite biosynthetic gene clusters.

Unlike stand alone *gilM* and *gilR* of the GV cluster, the homologous genes in the RMV cluster were fused to a larger ORF namely *ravRM*, whose *N*-terminal protein sequence showed similarity to GilR and whose *C*-terminal protein sequence was similar to GilM. Through both *in vivo* and *in vitro* studies, GilR was shown have a dehydrogenase activity which converts the hemiacetal moiety of preGV (**47**) to the lactone moiety of GV (**12**) (see section 4) (*56, 126*). Gene deletion experiments showed that GilM was necessary for GV biosynthesis.(*101*) However, its exact role was not fully known. Homology searches of GilM revealed a SAM-binding domain and showed its similarity to SAM-depedent methyltransferases from various organisms. This indicated that GilM might be a methyltransferase responsible for methylating one of the phenolic hydroxyl groups (C10-OH or C12-OH) of **46** during GV biosynthesis. In this context, we propose RavRM to be a bifunctional enzyme capable of catalyzing identical reactions of GilR and GilM. Like GilMT, the product of *ravMT* displayed similarity to various *O*-methyltransferases indicating its potential role in the methylation of the remaining hydroxyl group of **46**.

Genes for deoxysugar biosynthesis and attachment

A total of six genes encoding deoxysugar biosynthetic enzymes were identified in the RMV cluster. *RavD* and *ravE* appear to be translationally coupled. The product of *ravD* and *ravE* resembled the various NDP-glucose synthases and NDP-glucose-4,6-dehydratases from *Streptomyces* sps, respectively. RavD displayed 76%, 75% and 71% identity to MtmD, GilD and StrD of the mithramycin (*127*), GV (*71*) and streptomycin (*128*) deoxysugar biosynthetic pathways, respectively. Similarly, RavE showed 73%, 72% and 69% identity to MtmE, GilE and StrE of these pathways, correspondingly. This suggested that RavD might be a NDP-hexose synthase that catalyzes the first step of the NDP-ravidosamine biosynthesis and RavE catalyzes the follow-up step to generate NDP-4-keto-6-deoxy-D-glucose (**18**, **Scheme 2.2**). Start and stop codons of the rest of the proposed ravidosamine biosynthetic genes *ravNMT*, *ravIM*, *ravAMT* were translationally coupled. *RavIM* represented a unique gene of the RMV cluster which displayed its closest similarity to putative WxcM (bifunctional sugar isomerase/*N*-acetyltransferase) domain-containing proteins from various bacteria such as *Burkholderia*

phymatum, *Silicibacter* sp. TM1040 and *Desulfococcus oleovorans*. Genetic analyses showed that the WxcM protein from *Xanthomonas campestris* pv. *campestris* catalyzes isomerization of TDP-4-keto-6-deoxy-D-glucose (**18**) to TDP-3-keto-6-deoxy-D-galactose and also acetylates the amino group of TDP-3-amino-3,6-dideoxy-D-galactose (**56**) during lipopolysaccharide (LPS) biosynthesis (*129*). RavIM also showed its close similarity to the NDP-hexose-3,4-ketoisomerases such as (ORF2) (*130*) (51%/68%, identity/similarity) and Tyl1a (*131*) (35%/54%) from the mycaminose biosynthetic pathway. Similarly, RavIM also exhibited 39% identity and 60% similarity to the FdtA from *Aneurinibacillus thermoaerophilus*. FdtA has been shown to catalyze TDP-4-keto-6-deoxy-D-glucose (**18**) to TDP-3-keto-6-deoxy-D-galactose (**55**) during the biosynthesis of a surface layer 3-acetamido-3,6-dideoxy-α-D-galactose moiety of the bacterial surface glycan chain (*132*), whereas Tyl1a acts on the same substrate but generates the product TDP-3-keto-6-deoxy-D-glucose (**53**) with an equatorial-facing hydroxyl group at C-4 position instead. We propose that RavIM catalyzes the isomerization reaction identical to that of FdtA (**Scheme 2.2**) generating an axial hydroxyl group at 4-position, which later is acetylated in RMV. If this hypothesis is true, RavIM would be the first enzyme from a secondary metabolite biosynthetic pathway to have such an activity.

The product of *ravAMT* was similar to EryC1(*133*) (60%/75%), OleN2 (*134*) (61%/73%) and DesV (*135*) (59%/71%). All of these homologous enzymes were pyridoxal-5-phosphate (PLP)-dependent and were proposed/characterized to catalyze amino transfer from an amino acid, preferably L-glutamate, to the carbonyl carbon of a deoxysugar intermediate (*135, 136*). Therefore, we propose that RavAMT serves as an aminotransferase which converts **55** to **56** using L-glutamate as an amino donor during TDP-D-ravidosamine biosynthesis. The encoded amino acid sequence of *ravNMT* showed significant similarity to the various deoxysugar *N*-methyltransferases including SnogX (56%/71%) and SnogA (57%/70%) from TDP-L-nogalamine (*137*) and to other deoxysugar *N,N*-dimethyltransferases, e.g. DesVI (55%/71%) and OleM1 (53%/68%) from TDP-D-desosamine pathways, respectively (*68, 136*). Both of these *N,N*-dimethyltransferases transfer a methyl group from *S*-adenosyl methionine (SAM) to the amino nitrogen twice, in an iterative fashion.
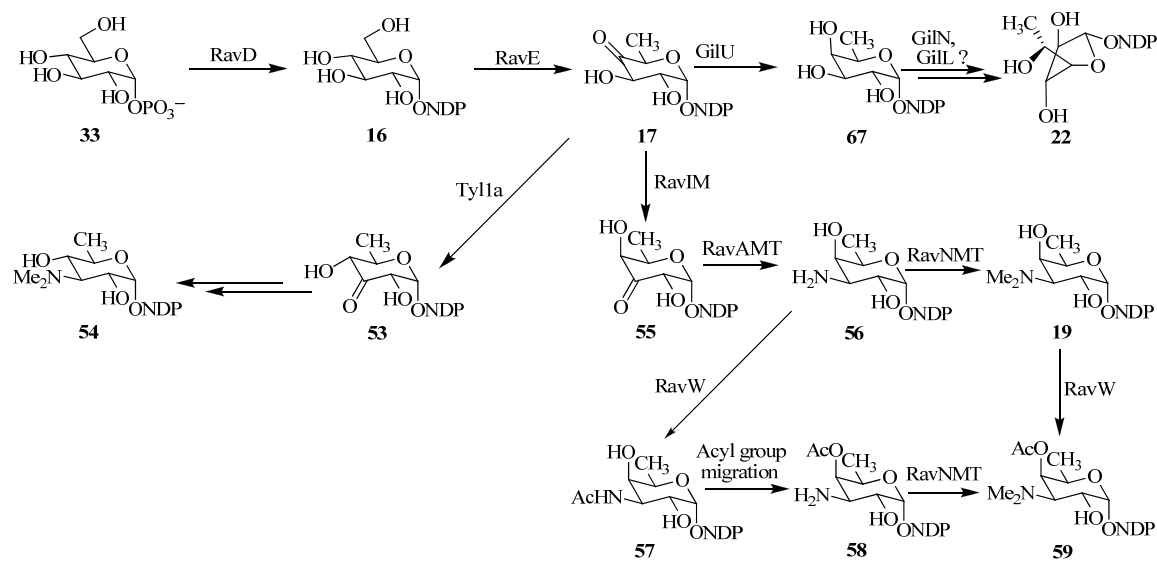
Scheme 2.2 Proposed pathways for the biosyntheses of 4′-*O*-acetyl-NDP-D-ravidosamine, and NDP-D-mycaminose.

In context of RMV biosynthesis, we assumed that RavNMT catalyzes a similar reaction (**56→19**) utilizing the product of RavAMT as the substrate. RavW was the only candidate enzyme that could acetylate the 4-hydroxyl group of **19**. The product of *ravW* shows modest similarity to GCN-5 related *N*-acetyltransferase (GNAT)-family enzymes which catalyze the transfer of an acetyl group from acetyl-CoA to the primary amine of a variety of acceptor substrates and function in diverse pathways of prokaryotes and eukaryotes. Like ATs, PKS-associated GNATs were reported to catalyze the conversion of malonyl-CoA to acetyl-CoA and then transfer the product to the thiol group of *holo*-ACPs directly (*138*). Regarding the functional role of RavW, there could be three possible scenarios: i) RavW functions as AT; ii) RavW catalyzes *N*-acetylation of **56** where the acetyl group then migrates to the 4′-hydroxyl group (**57→58**); iii) RavW might serve as an acetyltransferase which catalyzes direct *O*-acetylation of **19** before or after the establishment of the C-C-glycosidic bond. The presence of two ATs in the RMV cluster and lack of an RavW homologue in the GV cluster made the first possibility less likely. The suspected *N* to *O*-acyl-shift was not observed previously when preparing TDP-3-acetamido-3,6-dideoxy-D-galactose enzymatically (see section 3) (*132*). This made the second option less possible. However, we could not rule out this possibility as the closely related antibiotic FE35B (**31**) contains an 3-acetamido and 4-O-acetyl groups. The presence of significant amount of deacetylravidomycin V (**29**) in the culture broth of *S. ravidus* indicated acetylation of the 4′-OH group as the last step of the deoxysugar biosynthesis prior to or after the

36

attachment of the sugar. Therefore, we favored the third possibility. However, further experiments were necessary to clarify these issues.

The glycosyltransferase (GT) that establishes *C*-glycosidic linkage between the aromatic backbone and amino sugar of RMV was particularly interesting from the biosynthetic point of view as it could be used for combinatorial biosynthesis. We identified a glycosyltransferase encoding gene (*ravGT*) upstream of *ravNMT*. The deduced amino acid sequence of *ravGT* exhibited similarity to various *O*- and *C*-glycosyltransferases involved in secondary metabolite biosyntheses, such as LanGT2 (47%/67%), GilGT (46%/66%), Med-ORF8 (48%/65%) and UrdGT2 (42%/60%) from GV, landomycin, medermycin and urdamycin biosynthetic pathways, respectively (*103,139,140*). Unlike GilN of the GV cluster, there was no other GT homologue that could potentially be involved in or assist *C*-glycosylation in the RMV pathway. Therefore, we assumed that RavGT might be sufficient to catalyze the attachment of ravidosamine or 4′-*O*-acetyl-ravidosamine to **46**, establishing a C-C bond. Here, it was important to note that UrdGT2, GilGT, SaqGT5 and IroB were the only *C*-glcosyltransferases characterized at a functional level (*33, 141, 142*). The former three enzymes used NDP-deoxysugars as donor substrates, whereas the latter used UDP-glucose. Given the fact that sugar donor substrate flexibility of a GT determines its potential applicability to generate novel molecules through the glycodiversification, we were curious to explore the flexibility of RavGT. The ΔGilGT mutant strain (*S. lividans* TK24/cosG9B3-GilGT⁻), where the *gilGT* gene was deleted in-frame was the ideal host to test the flexibility of RavGT (*33*). This strain harbored the entire GV gene cluster except the glycosyltransferase (encoded by *gilGT*) thus providing a yet unknown sugar acceptor substrate and TDP-D-fucofuranose. Therefore, we attempted cross complementation of the GilGT-minus mutant with a *ravGT* expression construct, (pRavGT). The constitutive expression of the cloned gene was maintained by the presence of erythromycin resistant gene promoter (*ermE*P*). Interestingly, complementation of the mutant with pRavGT restored gilvocarcin E production, however GE (**12-a**) was produced instead of GV (**12**) (**Figure 2.4**). Accumulation of GE instead of GV was not surprising as the activity of GilOIII was suppressed in the GilGT-minus mutant due to a polar effect (*33*).
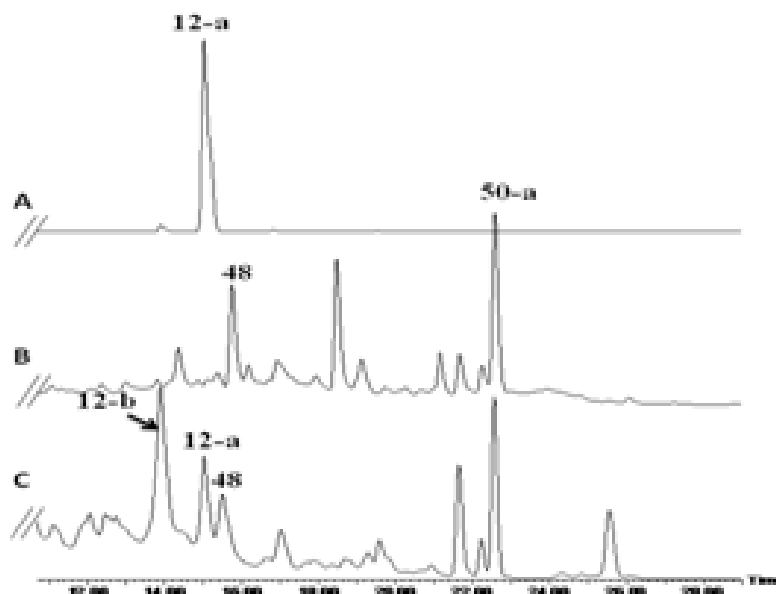
Figure 2.4 HPLC analyses of metabolites generated through the complementation of ΔGilGT strain. Trace A : standard GE (**12-a**); trace B : metabolites [homorabelomycin (**48**); defucoGE (**50-a**)]from ΔGilGT strain; trace C : metabolites [GM (**12-b**); GE (**12-a**); defucoGE (**50-a**)] from ΔGilGT-pRavGT strain.

UrdGT2 was reportedly the only known *C*-GT to have flexibility with regard to both, sugar donor and acceptor substrates (*32, 44*). More recently, the generation of 4′-hydroxygilvocarcins through the inactivation of ketoreductase (GilU) has demonstrated the flexibility of GilGT regarding its donor substrate in the GV pathway. However, GilGT could not accept a variety of amino sugars or branched sugars as donor substrates (unpublished results). Our current results provided the first example of a *C*-glycosyltransferase that could utilize both five-membered TDP-D-fucofuranose (**22**) and the six-membered aminosugar TDP-D-ravidosamine (**19**) as sugar donor substrates.

Genes encoding regulatory and resistance enzymes, or enzymes of unknown function

RavX4 and *ravX5* represent two regulatory genes whose encoded amino acid sequence resembled transctriptional regulatory proteins from different *Streptomyces* sps. The function of six other orfs (*ravX2, X3, J, X6, X7, X8*) could not be assigned through homology studies. The products of *ravX* and *X1* were homolog to the phenylacetic acid degradation pathway enzymes and were not related to RMV biosynthesis. Located at the right and left ends, the products of

38

ravX and *ravX8* showed similarity to hypothetical proteins indicating the end of the cluster at both sides.

Heterologous production of deacetylravidomycins in *S. lividans* TK24

Analysis of the DNA sequence of cosRav32 indicated that the cosmid was harboring all of the necessary genes for the production of RMV. This prompted us to attempt the heterologous expression of the entire RM cluster in *S. lividans* TK24. Such expression would utilize all of the naturally occurring promoters, regulatory, resistance and the RMV biosynthetic genes. The cosmid was introduced in *S. lividans* TK24 through a well-established *E. coli* : *Streptomyces* inter generic conjugation protocol (*143*). The yellowish green exconjugants obtained on the plate were introduced to liquid culture for metabolite production. HPLC analysis showed that there were at least six new metabolite peaks in the culture broth of the *S. lividans* TK64/cosRav32 as compared to the control strain *S. lividans* TK64/pOJ446. UV spectra, retention time in HPLC and ESI-MS spectra comparison revealed that four of the six peaks represent 2,3-dehydroUWM6 (**41-a**, 2 mg L^{-1}), 2,3-dehydro-homo-UWM6 (**41**, 20 mg L^{-1}), preGM-o-quinone (**42-d**, 0.2 mg L^{-1}) and preGV-o-quinone (**42-c**, 1.3 mg L^{-1}) (**Figure 2.5**). The other two peaks (eluted at Rt.= 6.6 min and 7.1 min) exhibit UV spectra typical to that of ravidomycin with a slight hypsochromic shift of one of the maxima from 392 nm to 380 nm. This indicated that the C-8 vinyl side chain of **28** might be replaced by a saturated ethyl or methyl group. Low resolution ESI-MS analyses further showed that the mass of the first compound (MW, $[M+H]^+ = 510$ in positive APCI and $[M-H]^- = 508$ in the negative APCI MS) is 14 amu less than that of the second (MW, $[M+H]^+ = 524$ in positive APCI and $[M-H]^-$ 522 in negative APCI MS) indicating their real molecular weights to be 509 and 523 amu, respectively. This clearly showed that these compounds differ from each other by a single methylene group (-CH_2-) likely at the C-8 position, and unlike **28**, the 4′-*O*-acetyl group may be absent in both compounds. The molecular formula of the latter compound obtained through high resolution EI-MS ($C_{29}H_{33}NO_8$, 523.2207 g mol^{-1}) matched perfectly with the predicted molecular formula of deacetylravidomycin E (**29-b**). However, poor production (< 50 μg L^{-1}) of these metabolites hindered a larger scale isolation for NMR analysis. Given the fact that ancymidol serves as an effective CYP450 inhibitor, and that a homologue of a CYP450 (GilOIII) is involved in the vinyl side chain formation of GV, compound **29-b** could be obtained by inhibiting this enzyme of *S.*

ravidus as well. As expected, the fermentation of *S. ravidus* in the presence of ancymidol revealed a new metabolite in the HPLC-MS, whose retention time and mass were identical to that of **29-b** isolated from *S. lividans* TK64/cosRav32. About 2 mg of the compound was isolated using preparative HPLC and its structure was analyzed through NMR spectroscopy and HR-MS spectrometry. When the ^{1}H NMR of **29-b** was compared with that of **28** (RMV), all signals were observed, except for the vinyl and 4′-*O*-acetyl protons which typically appear at δ 5.5-7.0 ppm and δ 2.0-2.5 ppm, respectively. Instead, an additional triplet (δ 1.32 ppm) and quartet (δ 2.83 ppm) signals were observed. This was a clear indication of the presence of an ethyl side chain at the C-8 position, thereby confirming its structure (see NMR data in the materials and methods section). The structure of the compound was further confirmed by ^{13}C NMR analysis. Based on these observations, we also proposed the structure of the minor compound to be deacetylravidomycin M (**29-a**).

The results clearly show that the isolated cluster covers most of the genes required for the biosynthesis of RMV. However, the marginal accumulation of **29-b** (deacetylRME) instead of **28**, and the accumulation of **41, 41-a, 42-c** and **42-d** in large amounts was surprising. Gene inactivation, complementation and *in vitro* assay results have suggested that **41** and **41-a** are biosynthetic intermediates of the GV and jadomycin pathways, respectively (*30, 56, 57, 61*). This hypothesis might be true for the RMV biosynthesis as well, where RavOI, RavOII and RavOIV could convert **41** to **46** through a series of oxidative reactions including a C-C bond cleavage (**Scheme 2.1**).
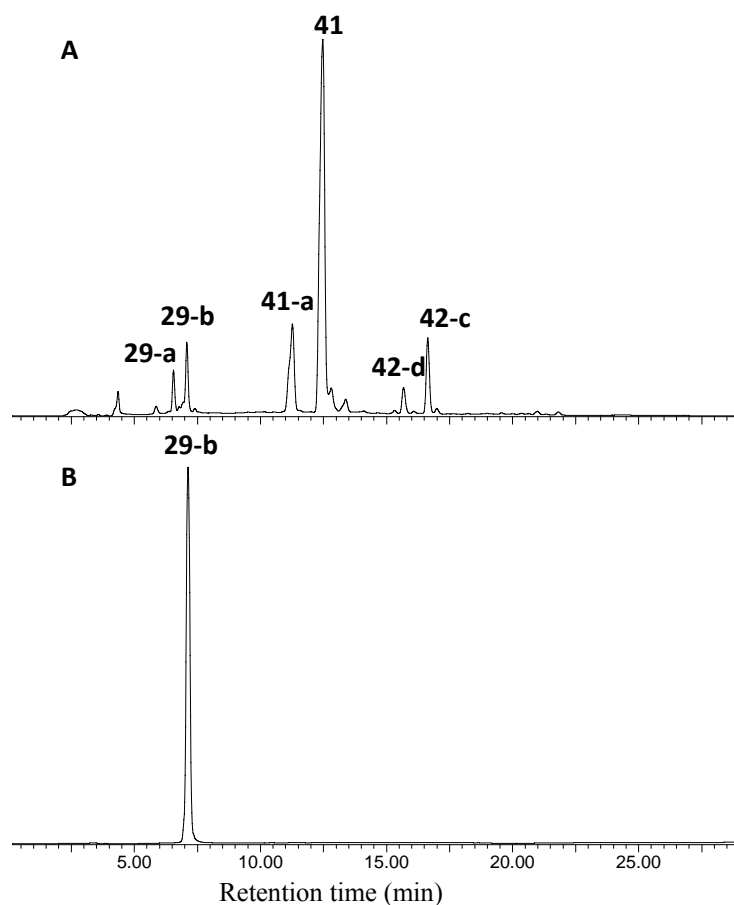
Figure 2.5 HPLC analyses of the metabolites: Trace A; metabolites extracted from *S. lividans* TK24-cosRav32, trace B; standard deacetylravidomycin E (**29-b**); deacetylravidomycin M (29-a); 2,3-dehydro-homo-UWM6 (**41**); 2,3-dehydroUWM6 (**41-a**); pregilvocarcinV-o-quinone (**42-c**); pregilvocarcin M-o-quinone (**42-d**)

The accumulation of these intermediates indicated that there could be some disturbances in the oxygenase activities (particularly, the conversion of **41→45**). However, the lack of an acetyl group at the 4′-position of the sugar residue and the vinyl side chain at the C-8 position in the products suggested that the proposed acetyltransferase (encoded by *ravW*) and CYP450 (encoded by *ravOIII*) are not functioning. This raised two questions: i) were these genes functionally inactive and the real functional genes were somewhere outside of the sequenced region? or ii) was the expression of these genes negatively regulated in the heterologous host? To clarify this issue, a *ravOIII* expression construct (pRavOIII) was generated using a pEM4 vector, into which the gene was expressed under the control of the erythromycin resistance gene promoter (*ermE*P*). This construct was then used to complement the ΔGilOIII strain (S.

41

*lividans*TK24/cosG9B3-GilOIII⁻) that harbors the entire GV cluster except GilOIII, and accumulates GE (**12-a**) and some other biosynthetic shunt products, mainly rabelomycin (**48-a**) and homorabelomycin (**48**). Interestingly, expression of *ravOIII* in the ΔGilOIII mutant resulted in the conversion of GE (**12-a**) to GV (**12**) (**Figure 2.6**). This demonstrated without any ambiguity that *ravOIII* encodes the functionally active enzyme responsible for the vinyl side chain formation of RMV. The incomplete conversion could be due to the imperfect fitting of RavOIII to the other GV biosynthetic enzymes. Therefore, we suspected that expression of *ravOIII* might have been negatively regulated during the heterologous expression of cosRav32. Our attempts to prepare RavW *in vitro* repeatedly failed due to insolubility of the enzyme. Further investigation including an *in vitro* characterization of this enzyme would help assign its exact functional role.

Comparison of the GV and RMV gene clusters

One of the major goals of this study was to identify the genes involved in the biosynthesis of the D-fucofuranose moiety of GV through the comparative analysis of the two clusters. Reflecting on their structural similarities, similar sets of PKS and post-PKS biosynthetic genes were found in both clusters with the exception of an additional ACP (encoded by *ravC1*) and the fused ORF *ravRM* in the RMV cluster. *GilD/ravD* and *gilE/ravE* represented the early deoxysugar biosynthetic genes for both of the D-fucofuranose and 4′-O-acetyl-ravidosamine moieties. As discussed earlier, the products of *RavIM*, *ravAMT*, *ravNMT* and *ravW* could potentially catalyze the follow-up steps to biosynthesize the 4′-O-acetylravidosamine moiety of RMV (see section 3). No homologues of *gilU*, *gilN* and *gilL* were identified in the RMV cluster indicating their unique roles in GV biosynthesis, likely in GV's deoxysugar moiety (D-fucofuranose). While the product of *gilU* was known to catalyze the conversion of TDP-4-keto-6-deoxy-D-glucose (**18**) to TDP-D-fucopyranose (**67**), GilN and GilL are the only candidate enzymes, possibly responsible for the sugar ring contraction that leads to the conversion of **67** to D-fucofuranose (**22**) (**Scheme 2.2**) (*102*). Whether this ring contraction takes place before or after glycosylation remains unclear.
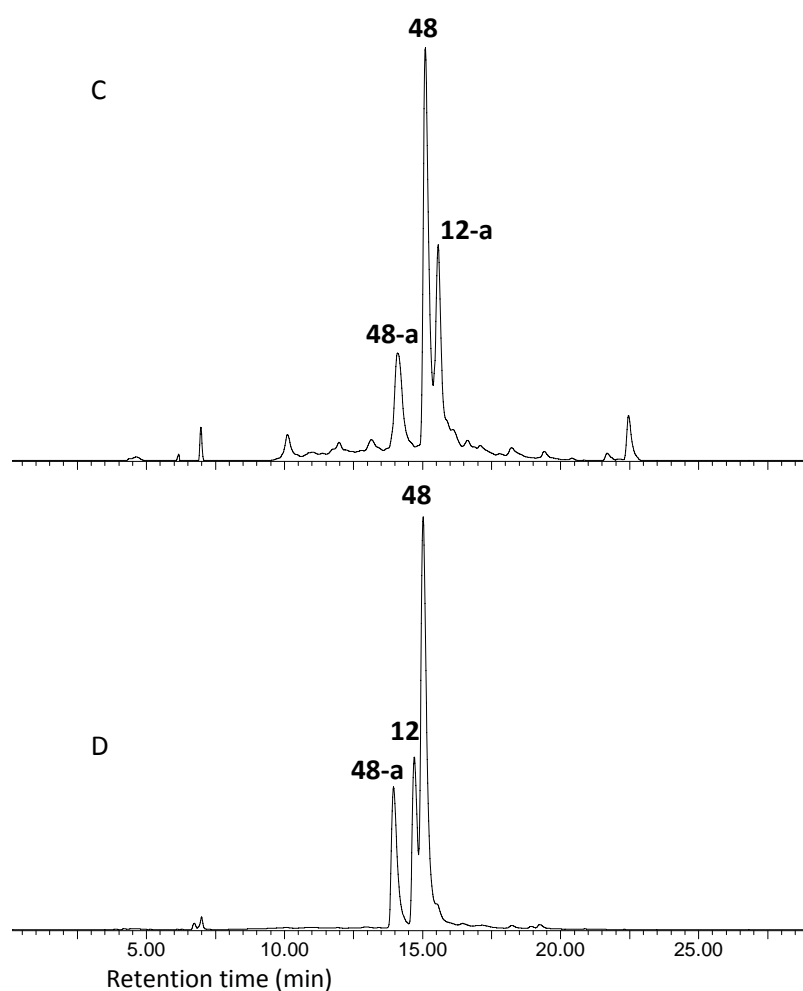
Figure 2.6 HPLC analyses of the metabolites generated through the complementation of ΔGilOIII strain: trace C; from the *S. lividans* TK24-cosG9B3-GilOIII⁻ mutant; trace D: from the *S. lividans* TK24-cosG9B3-GilOIII⁻ mutant strain after complemented with pRavOIII. **12**, **12-a**, **48** and **48-a** represent GV, GE, homo-rabelomycin and rabelomycin, respectively.

Discussion

Glycosylated natural products have diverse biological activities. Because *C*-glycosides are resistant to glycosidase activity, they are more promising for therapeutic use than *O*-glycosides. In the past decades, several dozen secondary metabolite biosynthetic gene clusters from microorganisms have been cloned, sequenced and fully/partially characterized. The majority of these investigations were focused on *O*-glycosylated products. Gilvocarcins, urdamycins, Sch 47554, Sch 47555 (*114*), saquamycin Z, geltamycin B were the only *C*-glycosylated angucyclines whose gene clusters had been cloned. They all harbor neutral

43

deoxysugar moieties. Similarly, the medermycin cluster represented the sole type II-PKS derived secondary metabolite with a *C*-glycosidically linked amino sugar (D-angolosamine). In this context, the identification of the sugar moiety related biosynthetic genes in addition to the defucoGV biosynthetic genes of the RMV cluster was very significant. Production of ravidomycin derivatives in the heterologous host *S. lividans* TK24 clearly showed that the isolated gene cluster is sufficient for the biosynthesis of deacetylravidomycin E (**29-b**). Because RavGT utilizes an amino sugar, presumably TDP-D-ravidosamine as the donor substrate, it could be a valuable tool to generate novel RMV analogues with diverse amino sugars, such as angolosamine, desosamine, or daunosamine. The partial complementation of the GilGT-minus mutant with RavGT was a clear indication of such a possibility. This was the first report of a *C*-glycosyltransferse that utilizes chemically very different deoxysugars (a neutral TDP-D-fucofuranose and a basic TDP-D-ravidosamine). However, further investigations on the sugar donor substrate flexibility of RavGT might be necessary for its successful application in combinatorial biosynthesis. It was well established through inactivation and complementation experiments that GilOIII catalyzes the oxidation of the ethyl side chain to the vinyl group (*33*). Despite the presence of a functionally active GilOIII homologue (RavOIII) in the cluster, it was not clear why deacetylravidomycin E (**29-b**) was observed instead of RMV (**28**) or deacetylRMV (**29**). The results also pose ambiguity over the true function of the potential acetyltransferase RavW.

Comparison of the GV and RMV gene clusters helped to identify the unique genes of each cluster. The comparative analyses outlined enzyme candidate candidate (GilN or GilL) to be involved in ring contraction reaction during the biosynthesis of the D-fucofuranose moiety of GV. We observed several features unique to the RMV cluster. Fusion of two *gilR* and *gilM* homologous genes into the single orf *ravR-M* was particularly interesting. This also indicated that the products of both genes might work very closely during GV biosynthesis. Unlike the GV cluster, two functional ACPs were found in the RMV biosynthetic locus. This raised two possibilities: i) the reported ACP of the GV pathway (GilC) does discriminate between propionate or acetate starter units; or ii) another ACP is somewhere outside the cluster in the genome of the GV producer *S. griseoflavus*. In this context it is reasonable to speculate that the GV pathway could be borrowing the missing ACP from the host to reconstitute GV production in *S. lividans* TK24 while producing GV heterologously through the expression of cosG9B3 (the

cosmid that contains all of the hypothesized GV biosynthetic genes). Acyltransferases (ATs) are often missing in type II PKS gene clusters. In those cases, ATs are borrowed from the fatty acid biosynthetic pathway. Like in the GV cluster, the presence of two RavATs (RavP and RavQ) was unique. We suspected that RavP self-loads acetyl-/malonyl-CoA and transfers it to either RavC or RavC1. Similarly, RavQ might transfer propionyl-CoA to the remaining ACP.

In summary, we have cloned and sequenced the gene cluster responsible for the biosynthesis of ravidomycin. Heterologous production of deacetylRME in *S. lividans* TK24 combined with the results of complementation experiments unambiguously demonstrated the functional role of the cluster for RMV biosynthesis. In addition, we discovered a donor substrate flexible *C*-GT capable of utilizing a 5-membered neutral sugar (D-fucofuranose) as well as a 6-membered basic sugar (D-ravidosamine). The isolated gene cluster could serve as a valuable tool to generate GV/RMV analogues through combinatorial biosynthesis.

Experimental section

Construction of plasmids and DNA manipulation

Oligonucleotide primers, plasmids and bacterial strains used in this study are summarized in **Table 2.2** and **2.3**. Routine DNA manipulations, such as cloning, restriction analysis, ligation, transformation etc. were carried out following the standard protocols for *E. coli* (*144*) and *Streptomyces* (*143*). High fidelity *Pfu* DNA polymerase was used to amplify various DNA fragments throughout these experiments. PCR products were cloned into PCR-Blunt-II-TOPO cloning vector (Invitrogen) and sequenced to confirm that mutation had not been incorporated during PCR amplification. For complementation studies, *ravOIII* was amplified and ligated into the pEM4 vector at *Xba*I and *Eco*RI sites whereas *ravGT* was ligated at *Nde*I and *Eco*RI sites of pUWL201PW (*67*).

Bacterial strains and culture conditions

Various strains of *E. coli* were cultured at 37° C in Luria Bertani (LB) broth or on LB-agar medium supplemented with various antibiotics (kanamycin, 50 μg mL^{-1}; chloramphenicol, 25 μg mL^{-1}; apramycin, 50 μg mL^{-1} and ampicillin, 50 μg mL^{-1}) whenever necessary. *Streptomyces* strains were grown on M2-agar or in Soymeal glucose (SG) liquid medium for the

metabolite analyses as reported previously.(*57*) MS-agar medium was used for the transfer of the cosmids from *E. coli* into *Streptomyces* following a standard protocol (*143*). Antibiotics (apramycin, 50 μg mL^{-1} and thiostrepton 25 μg mL^{-1}) were supplemented in the medium whenever necessary.

Construction and screening of genomic cosmid libraries

Genomic DNA of *S. ravidus* was isolated following a standard protocol (*143*). The genomic DNA was partially digested with *Sau*3A1, dephosphorylated with calf intestinal phosphatase (CIP) and ligated to the pOJ446 vector digested with *Bam*HI and *Hpa*I. The ligation sample was transduced into *E. coli* XL1-Blue MRF using Gigapack III XL packaging extract (Stratagene). More than 2300 colonies were obtained as positive clones. Two sets of degenerate primers (KS-probe-for and KS-probe-rev, and DH-probe-for and DH-probe-rev, **Table 2.2**) were designed to amplify the internal nucleotide sequences of ketoacyl synthase (KS$_\alpha$) and TDP-glucose-4,6-dehdyratase from the genomic DNA of *S. ravidus* (see appendix). The amplified KS$_\alpha$ fragments were labeled with DIG and used to screen the library. The positive colonies were further screened with the TDP-glucose-4,6-dehdyratase probe. The colony hybridization and sourthern blot analyses revealed three different cosmid clones. Cosmid cosRav32 was selected for the sequencing.

Sequencing and annotation of the gene clusters

Sequencing of the cosmids was carried out using a shot-gun approach at the College of Agricultural Sciences (University of Kentucky). Phred/Phrap/Consed software package (http://www.phrap.org.) was used to process and assemble raw sequence data into the larger contigs. The small gaps between contigs were filled up by primer walking. Frame plot (http://www0.nih.go.jp/~jun/cgi-bin/frameplot.pl) and ORF finder (http://www.ncbi.nlm.nih.gov/projects/gorf/) programs were used to assign ORFs. Functional assignment of ORF products was performed through database comparison using various BLAST (Basic Local Alignment Search Tool) search tools on the server of the National Center for Biotechnology Information, Bethesda, Maryland (http://www.ncbi.nlm.nih.gov).

Fermentation and isolation of metabolites

Streptomyces strains were grown on MS agar for 6 days at 30 °C. An agar chunk full of bacterial spores was inoculated into 100 mL SG liquid medium in 250 mL Erlenmeyer flasks and incubated for 3 days at 30 °C and 200 rpm to prepare a seed culture. For analysis of metabolites, 3 mL of the seed culture was used to inoculate 100 mL SG medium liquid culture and grown for 5 days. The culture was harvested through centrifugation to obtain culture broth and pellet fractions. The culture broth was extracted with an equal volume of ethyl acetate two times. The combined organic layers were dried in a rotary evaporator. The crude extract was dissolved in 3 mL of methanol and filtered through a 0.2 μM syringe filter prior to high performance liquid chromatography-mass spectrometry (HPLC-MS) analysis. The pellet fraction was resuspended in 20 mL acetone and sonicated for 10 min. The acetone extract was collected through centrifugation and dried under reduced pressure. The prepared extract was dissolved in methanol and subjected to HPLC-MS analysis.

The production of ravidomycins and gilvocarcins by wild/mutant strains were analyzed by the aforementioned protocol. For the isolation of the compounds, the production scale was increased to 4 L. The culture was harvested through centrifugation. The cell pellets were extracted with acetone as described above and dried off the solvent. The extract was combined with the broth and loaded into the reverse phase silica (RP-C_{18}, 200 g) column (5 × 65 cm), pre-equilibrated with deionized water. The effluent was discarded and the column was washed with 1 L of 10% acetonitrile (in water). Elution was carried out using an increasing gradient of acetonitrile (20%, 40%, 60%, 80% and 100%, 1L each). The fractions containing the target compounds were dried in a lyophilizer. Final purification of the metabolites was carried out through preparative scale HPLC. The purity and authenticity of the compound were confirmed through ^{1}H and ^{13}C NMR analyses. A linear gradient of acetonitrile and acidified water (solvent A = 0.1% formic acid in H_2O; solvent B = acetonitrile; 0-15 min 25% B to 100% B; 16-24 min 100% B; 25-26 min 100% to 25% B; 27-29 min 25% B) was used to separate compounds. SunFire™ prepC$_{18}$ column (19 × 150 mm, 5 μm) and Symmetry C$_{18}$ (4.6 × 250 mm, 5 μm) columns were used for preparative and analytical scale separations, respectively. The flow rate was maintained at 10 mL min^{-1} and 1.0 mL min^{-1} for preparative and analytical scale separations, correspondingly. Micromass ZQ 2000 (Waters corporation) equipped with HPLC (Waters

alliance 2695 model) and photodiode array detector (Waters 2996) were used to analyze the compounds. Atmospheric pressure chemical ionization (APCI) and Electrospray ionization (ESI) probes were used to detect molecular ions. Electron impact-high resolution mass spectrometric (EI-HR-MS) analyses of the compounds were acquired at University of Kentucky mass spectrometry facility.

To isolate deacetyl ravidomycin E, 20 mL of fully grown *S. ravidus* seed culture was grown in SG medium (5 flasks×100 mL) and 100 μl of ancymidol solution (in DMSO) was added into the culture to 2 mM final concentration. Only DMSO was used in the control flask. The fermentation was carried out for 6 days and the metabolites were analyzed by HPLC-MS. Isolation and purification of 4′-*O*-deacetylravidomycin E was carried out according to the aforementioned protocol. The structure of the isolated compound was determined by NMR spectroscopy analyses (See NMR data).

Table 2.2 List of primers used in this study

Name of the primers	Oligonucleotide sequence
RavOIII_for	5'-AAG*TCTAGA*AGGAGACGGAGCACCATGGCG-3'
RavOIII_rev	5'-CG*GAATTC*TCTGCTCGTCGGCCGTGGTCACGT-3'
RavGT_for	5'-AT*CATATG*AAAGTCCTGTTCATCGCAGCGGGA-3'
RavGT_rev	5'-AT*GAATTC*TCATCCGCGCACCAGTCCTTCCAG-3'
KS_probe_for	5'-GTSTCSACSGGSTGYACSTCSGGS-3'
KS_probe_rev	5'-SCCGATSGCSCCSAGSGARTGSCC-3'
DH_probe_for	5'-CSGGSGSSGCSGGSTTCATSGG-3'
DH_probe_rev	5'-CAGTGGTCSACGTGSAGCCACTCSCG-3'

Table 2.3 List of bacterial strains used in this study

Strain/ plasmid	Characteristics and relevance	References
E. coli XL1-Blue-MRF	Host for routine cloning works and for the construction of genomic libraries	Stratagene
E. coli ET12567/pUZ8002	Host for conjugal transfer of plasmid into *Streptomyces*	Danis et al.(*145*)
***Streptomyces* species**		
S. ravidus C23201	Ravidomycin producer-wild strain	Wyeth Co.
S. lividans TK24		
S. lividans TK24/cosG9B3-GT⁻	Produces defucoGE and defucoGM-the strain harbors the GilGT deleted mutant cosmid CosG9B3-GilGT⁻	Rohr et al.(*33*)
S. lividans TK24/cosG9B3-OIII⁻	Produces defucoGE and defucoGM-the strain harbors the GilOIII deleted mutant cosmid CosG9B3-OIII⁻	Rohr et al.(*33*)
S. lividans TK24/cosRav32	Produces2,3-dehydroUWM6, 2,3-dehydro-HomoUWM6, preGV M-o-quinone, preGV-GE-o-quinone, ravidomycin M, 4'-o-deacetylravidomycin E -the strain harbors the cosRav32	This work
S. lividans TK24/cosG9B3-GT⁻/pRavGT	Produces GE and GM-pRavGT was introduced into *S. lividans* TK24/cosG9B3-GT⁻	This work
S. lividans TK24/cosG9B3-OIII⁻/pRavOIII	Produces GV and GM-pRavOIII was introduced into *S. lividans* TK24/cosG9B3-OIII⁻	This work
Plasmids/cosmids		
pOJ446	Vector for cosmid library construction	Schoner et al.(*104*)
cosRav32	Entire ravidomycin gene cluster cloned into pOJ446	This work
pEM4		Salas et al.(*134*)
pUWL201PW	*E. coli-Streptomyces* shuttle plasmid that contains *ermE*p*for the expression of genes in *Streptomyces*	Raynal et al.(*67*)
pRavGT	*ravGT* cloned into pUWL201PW	This work
pRavOIII	*ravOIII* cloned into pEM4	This work
PCR-Blunt-II-TOPO	To clone PCR products	Invitrogen

^{1}H and ^{13}C NMR data of deacetylravidomycin E (29-b): ^{1}H NMR (CDCl$_3$, 300 MHz) δ 9.83 (br, s, 1H, OH 1, D$_2$O exchangeable) 8.51 (s, 1H, H11), 7.97 (d, J = 9.0 Hz, 1H, H3), 7.95 (d, J = 3.3 Hz, 1H, H7), 7.25 (d, J = 3.3 Hz, 1H, H9), 7.08 (d, J = 9.0 Hz, 1H, H2), 6.01 (d, J = 9.0 Hz, 1H, H1′), 4.44 (t, J = 9.0 Hz, 1H, H2′), 4.16 - 4.26 (m, 2H, H4′, H5′), 4.15 (s, 3H, OMe12), 4.12 (s, 3H, OMe 10), 3.35 (dd, J = 9.0, 3.0 Hz, 1H, H3′), 2.95 (s, 6H, N(Me)$_2$3′), 2.83 (q, J = 7.5 Hz, 2H, H1″), 1.32 (t, J = 7.5 Hz, 3H, H2″) 1.31 (d, J = 6.0 Hz, 3H ,Me6′); ^{13}C NMR (CDCl$_3$, 75 MHz) δ 170.2, 161.0, 157.5, 155.0, 152.5, 146.8, 129.8, 125.3, 124.7, 122.4, 122.4, 121.3, 116.0, 117.7, 114.5, 113.0, 102.8, 79.8, 78.5, 71.6, 69.6, 68.8, 56.6, 43.1, 29.2, 16.95, 15.37.

Deacetylravidomycin E **(29-b)**
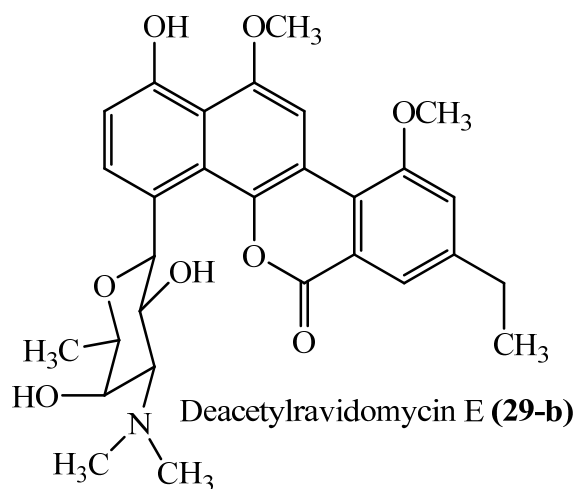
Section Three: Characterization of the TDP-D-ravidosamine biosynthetic pathway

Background

Glycosylated microbial natural products possess a wide range of pharmaceutical properties such as antigenic, chemotherapeutic, and antibiotic (*146*). The sugar moieties of these molecules are often deoxygenated to various extent, and are appended to the backbone through *O*- or *C*-, and rarely through *N*- or *S*-glycosidic linkages. Structural and functional studies have demonstrated that these sugar moieties tremendously contribute to the biological activities of the parent compounds (*147, 148*). Due to the large structural diversity of deoxysugars and their importance for biological activities, much attention has been paid to the elucidation of their biosynthetic pathways (*149, 150*). Recent progress in understanding and manipulation of these pathways has proved the concept that an array of glycosylated natural products can be generated through enzymatic glycodiversification (*151*). However, further application of this concept requires an in-depth knowledge of various deoxysugar biosynthetic pathway enzymes that includes the identification of the preferred substrate, catalytic activity, and their tolerance for other substrates. Thus, the success of generation of new glycosylated natural products largely depends on the sugar donor (NDP-sugar) and acceptor substrate flexibility of the dedicated glycosyltransferases (GTs) (*131*). Out of hundreds of GTs reported thus far, a few have shown flexibility for the donor(*34, 39*) or the acceptor substrates, (*27, 152, 153*) and very rarely for both. The majority of the GTs characterized from the secondary metabolite biosynthetic pathways are *O*-GTs such as AveB1(*31*), GtfD and GtfE (*154*), KanF (*155*), AmphD1 and NysD1(*156*). Unavailability of soluble GTs, their sugar acceptor substrates and NDP-sugar donor substrates contribute to the slow progress of *C*-GTs characterization *in vitro*. Recent advancements in understanding deoxysugar biosynthetic enzymes has made it possible to engineer a pathway to biosynthesize a target NDP-sugar in an efficient way. This approach is far superior than the current chemical synthetic counterpart as it is concerted, less time consuming, stereo- and regioselective with high product yield in addition to the environmentally friendly protocol (*157*).

Ravidomycin (RMV) and gilvocarcin V (GV) differ to each other in their *C*-glycosidically linked deoxysugars moieties. RMV comprises the 4′-*O*-acetyl-D-ravidosamine whereas GV contains D-fucoruranose. Earlier studies of structure-activity relationship (SAR) of

gilvocarcins and ravidomycins have clearly shown that the deoxysugar moieties are crucial for the biological activities (*93, 102*). Although, there is no direct comparison of biological activities among GV class compounds, it has been reported that RMV is the most active compound highlighting the unique role of the amino sugar moiety (*158*). Alteration of the sugar moieties of GV/RMV could be very promising to biosynthesize new analogues with potentially better pharmacological properties. There are many reports on biosynthesis of unnatural *O*-glycosides through combinatorial biosynthesis (*22, 25, 35, 37, 39*). Most of them have utilized 6-membered neutral sugars for the glycoslation. Glycodiversification utilizing basic sugars (aminosugars) or five membered sugars are scarce even though they are highly interesting from a structure/interaction point of view. Characterizations of RavGT/GilGT and their donor and acceptor substrate biosynthetic enzymes would provide a framework to generate GV/RMV analogues with unnatural (for example, a hypothetical 5-membered aminosugar) sugar moieties. Such studies *in vitro* provide a direction to engineer a strain of interest to bioysynthesize the target molecule *in vivo*. In this context, we report the characterization of the TDP-D-ravidosamine biosynthetic pathway through the *in vitro* activity assay of the proposed enzymes. We also report a two-stage, one-pot enzymatic synthesis of TDP-D-ravidosamine through combinatorial biosynthesis, thus providing an efficient route to generate a donor substrate for the RMV *C*-glycosyltransferase RavGT. Furthermore, our efforts to characterize the D-fucofuranose pathway *in vitro* are also discussed in this chapter.

Results

Characterization of RavD and RavE

Bioinformatic analyses of the ravidomycin biosynthetic gene cluster indicated six open reading frames relevant to TDP-D-ravodosamine biosynthesis (**Figure 3.1**). Homology studies indicated that RavD might be an NDP-glucose synthase that could catalyze the first step of the ravidosamine biosynthetic pathway i.e. the attachment of TDP at C-1 of the glucose-1-phosphate (G-1-P)(**16**) (**Scheme 3.1**). Similarly, the encoded product of *ravE* showed similarity to various TDP-glucose 4,6-dehydratases that were well known for deoxygenation at C-6 and oxidation at C-4 of TDP-D-glucose (**17**) to generate TDP-4-keto-6-deoxy-D-glucose (**18**). The start and stop codons of the remainder of the proposed ravidosamine biosynthetic genes (*ravNMT, ravIM* and

52

ravAMT) were translationally coupled. *RavIM* is a unique gene of the RM cluster (see section 2). To test their proposed functions, *N*-terminal His-tag proteins were produced through the
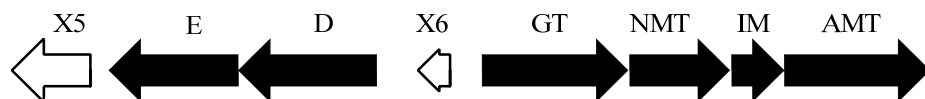
Figure 3.1 TDP-D-ravidosamine biosynthetic gene cluster from *S. ravidus*.

expression of the genes in *E. coli* and were purified using immobilized metal affinity chromatography (IMAC). The expected sizes of the proteins were observed on SDS-PAGE (**Figure 3.2**). Incubation of RavD with G-1-P and thymidine triphosphate (TTP) resulted in the consumption of TTP and concomitant formation of a new peak in HPLC that eluted at the identical retention time (37.7 min) of standard TDP-D-glucose (**Figure 3.3**). The completion of the reaction was noticed in 30 minutes. The molecular ion peak observed in low resolution ESI mass (at 563) was in good agreement with the calculated value (cal. [M-H]⁻ = 563.075). The structure was further confirmed through the comparison of the ^{1}H NMR of the purified compound with that of the reported data (**Table 3.1**).(*131*) Less than 5% turnover was noticed while using UTP, GTP, or CTP indicating TTP to be the preferred substrate for RavD. The results clearly demonstrated that RavD was a TDP-glucose synthase and it likely catalyzes the first step of the TDP-D-ravidosamine biosynthetic pathway.

Incubation of **17** with RavE did not alter the HPLC profile of the reaction mixture, but the peak at 37.7 min was broadened over time (**Figure 3.3**). This indicated a conversion, but the substrate and the products might be eluting at the same retention time.
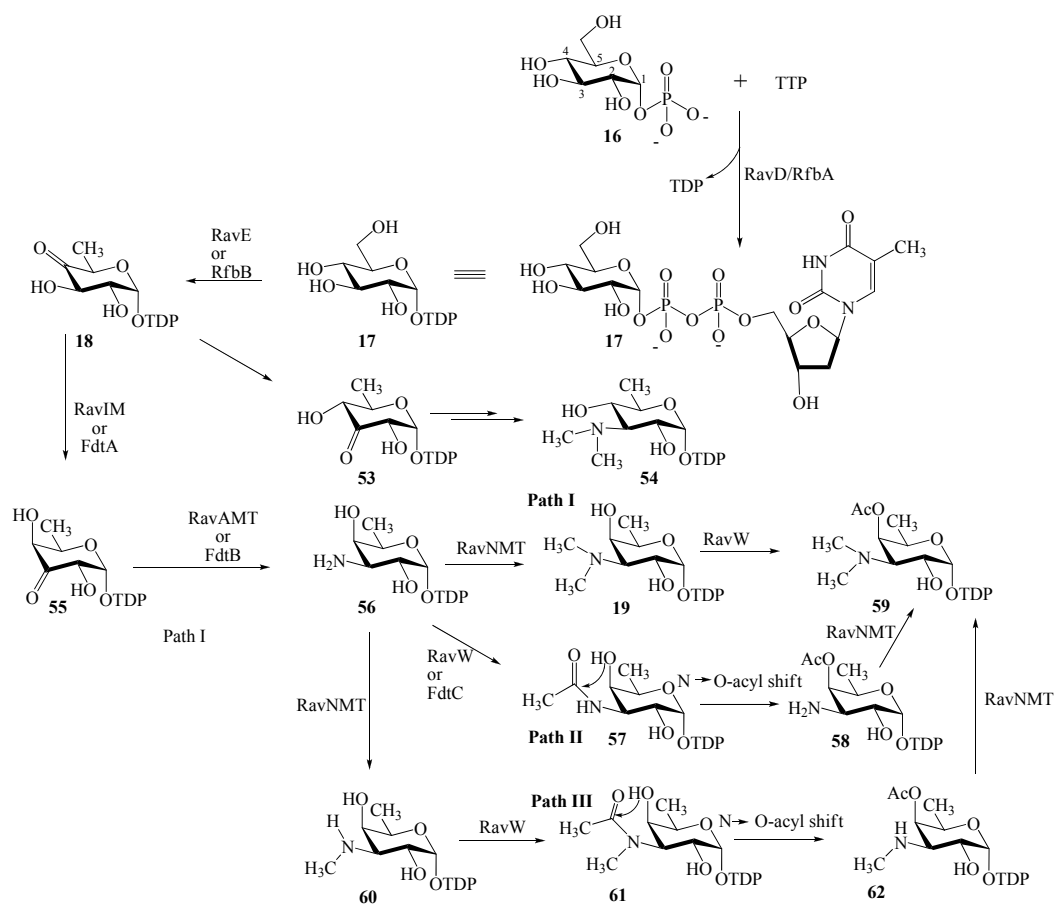
Scheme 3.1 A proposed pathway for TDP-4′-*O*-acetyl-D-ravidosamine biosynthesis.
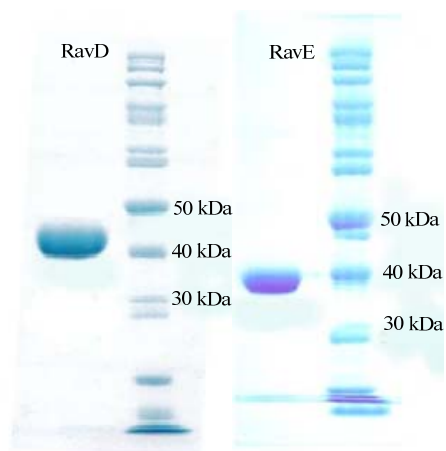
Figure 3.2 SDS-PAGE analyses of recombinant RavD (40.37 kDa) and RavE (37.93 kDa).

The incubation was continued until 1 h to allow the completion of the reaction and the compound was purified through gel filtration. The ^{1}H NMR spectrum of the compound was found to be identical to the previously reported data of TDP-4-keto-6-deoxy-D-glucose (**18**) (**Table 3.1**) (*131*). It is important to note that **18** is the common intermediate for a variety of deoxyhexoses including D-, L-, branched, or amino sugars. The results clearly showed that *ravE* encodes TDP-glucose 4,6-dehydratase which catalyzes conversion of **17** to **18** during the biosynthesis of TDP-D-ravidosamine (**19**).
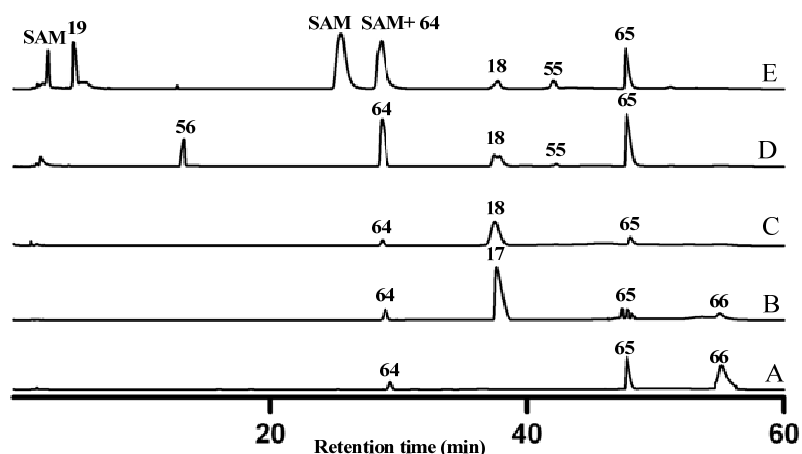
Figure 3.3 HPLC analyses of TDP-sugars generated through the activities of TDP-D-ravidosamine pathway enzymes. Trace A: Thymidine monophosphate (TMP, **64**) incubated with acetate kinase (ACK), thymidine monophosphate kinase (TMK) and adenosine triphosphate (ATP); trace B: Thymidine triphosphate (TTP, **66**) incubated with glucose-1-phosphate (G-1-P) and RavD; trace C: assay mixture B supplemented with RavE; trace D: assay mixture C supplemented with RavIM, RavAMT, L-glutamate and pyridoxol-phosphate (PLP); trace E: assay mixture D supplemented with *S*-adenosyl methionine (SAM) and RavNMT. Reactions were carried out for 1 hour prior to the HPLC analyses. The compounds on the HPLC traces for B, C, D and E remained unchanged when the commercial TTP was replaced by ACK, TMK, acetyl phosphate, TMP and ATP. Three different peaks were observed for commercially available SAM indicating the presence of impurities. TDP-D-glucose (**17**); TDP-4-keto-6-deoxy-D-glucose (**18**); TDP-D-ravidosamine (**19**); TDP-3-keto-3,6-dideoxy-D-galactose (**55**); TDP-3-amino-3,6-dideoxy-D-galactose (**56**); thymidine monophosphate (TMP, **64**); thymidine diphosphate (TDP, **65**); thymidine triphosphate (TTP, **66**)

TDP-D-glucose (17)   TDP-4-keto-6-deoxy-D-glucose (18)
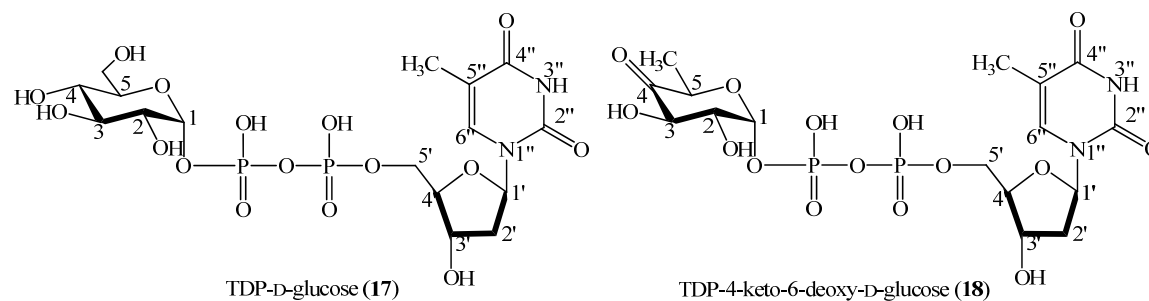
Table 3.1 ^{1}H NMR data of TDP-D-glucose (**17**) and TDP-4-keto-6-deoxy-D-glucose (**18**) (D$_2$O, 500 MHz)

Position number	TDP-D-glucose (17)		TDP-4-keto-6-deoxy-D-glucose (18)	
	[1]H δ [ppm]	Multiplicity (*J*/Hz)	[1]H δ [ppm]	Multiplicity (*J*/Hz)
1	5.58	dd (7.4, 3.5)	5.4	dd (6.8, 3.3)
2	3.51	td (9.8, 3.5)	3.51	td (10.07, 3.3)
3	3.43	t (9.8)	3.63	d (10.07)
4	3.72-3.78	m	-	-
5	3.81-3.90	m	3.93	q (6.5)
5-CH$_3$	3.72-3.78	m	1.06	d (6.5)
1'	6.33	t (6.98)	6.19	t (6.90)
2'	2.31-2.38	m	2.23	dd (6.90, 4.85)
3'	4.62	m	4.43-4.47	m
4'	4.12-4.16	m	3.99-4.06	m
5'	4.12-4.16	m	3.99-4.06	m
5"	1.9	s	1.77	d (0.97)
6"	7.72	s, br	7.57	d (0.97)

Characterization of RavIM and RavAMT

RavIM was the only candidate deoxysugar isomerase identified in the RMV gene cluster. It showed homology to *C*-terminal domains of various WxcM domain-containing proteins. WxcM from *Xanthomonas compestris* was proposed to catalyze two steps: i) isomerization of TDP-4-keto-6-deoxy-D-glucose to TDP-3-keto-6-deoxy-D-galactose and ii) N-acetylation of TDP-3-amino-3,6-dideoxy-D-galactose during the biosynthesis of the TDP-N-acetyl-3-amino-3,6-dideoxy-D-galactose subunit of the surface lipopolysaccharide (LPS). RavIM also showed considerable amino acid sequence identities to recently characterized 3,4-ketoisomerases, such as Tyl1a and FdtA.

```
RavIM    MTDTTAATGTTAGTDVSRVGKVRPCALMKLQTIGDPRGQLAVVEGAKDIGFPVKRLFYLY 60
Srm2     ------------------------------------MIESGVTVDFPVRRVYYMH 19
FdtA     --------------------MEN-KVINFKKIIDSRGSLVAIEENKNIPFSIKRVYYIF 38
WxcM     ------------------MSIERCKIINLPKISDPRGNLTFIESNKHIPFEIKRVYYLY 41
                                            :*     : * ::*::*:.

RavIM    DLPTSS-VRGDHAHRNLEQFVIPINGSFDVAVDDTVDTAVCRLDDPGQGLYIGPMVWNSL 119
Srm2     GQTQSSPHRGLHAHRTLEQLVIAVHGAFSITLDDGFQHATYRLDEPGAGLCIGPMVWRVL 79
FdtA     DTKGEE-HRGFHAHKKLEQVLVCLNGSCRVILDDGNIIQEITLDSPAVGLYVGPAVWHEM 97
WxcM     DVPGGE-TRGGHAHKNLQQLLIAVSGSFDVVVDDGYEKRRYHLNRSYYGLYIPTMIWREM 100
         .     .  ** ***:.*:*.:: : *:  : :**        *:.  ** : . :*. :

RavIM    VNFSEGAIALVLASEHYDEADYYRRYDEFLADAGARPX---- 157
Srm2     KDFDPDTVALVLASQHYEESDYYRDYDTFLHDARSLT----- 116
FdtA     HDFSSDCVMMVLASDYYDETDYIRQYDNFKKYIAKINLEKEG 139
WxcM     DNFSSGSVCLVLASDFYNEDDYIRDYDEFLKEVRK---EK-- 137
          :*. . : :****:.*:* ** * ** *
```

Figure 3.4 Multiple sequence alignment of RavIM and its homologues from the database. Srm2, FdtA, and WxcM are hexose isomerases identified from *Aneurinibacillus thermoaerophillus* L420-91, *Sulfurihydrogenibium yellowstonense* SS-5 and *Streptomyces ambofaciens*, respectively. The sequence highlighted by the box indicates the signature motif of FdtA-type isomerases.

Both enzymes utilize TDP-4-keto-6-deoxy-D-glucose (**18**) but generate two different products TDP-3-keto-6-deoxy-D-glucose (**53** by Tyl1a) and TDP-3-keto-6-deoxy-D-galactose (**55** by FdtA). Holden et al. identified several active site residues of FdtA in this signature region through the combined studies of site directed mutagenesis crystal structure.(*159*) Closer analysis of RavIM revealed the presence of the characteristic signature sequence [RGXHAH(K/R)X(L/I)XQXGS] (**Figure 3.4**). This further supported the aforementioned FdtA-like role of RavIM in 4-*O*-acetyl-TDP-D-ravidosamine (**59**) biosynthesis (see section 2 for more detail discussion). However, a bifunctional (isomerization/acetylation) role of RavIM could not be ruled out. In this context, RavIM might acetylate the amino group of **56** prior to N→O acyl

migration, or it could acetylate 4-OH of **55** before or after the transamination at C-3 position. To assign the exact functional role of RavIM, the enzyme was heterologously produced in *E. coli* and the protein was purified using IMAC. (**Figure 3.5**). The enzyme was then subjected to a series of enzyme assays. No significant change in the substrate concentration was noticed when the assay samples were analyzed after 0 to 2 hours of incubation with **18**. However, the incubation of **18** with RavIM for 24 hours resulted in the complete consumption of the substrate and the appearance of two major compounds TDP (51 min) and a decomposition product, presumably dihydroxy-4-keto-2,3-dihydropyran (**73,** 2.9 min) (data not shown). The UV spectrum of the minor compound observed at 43 minutes was identical to that of the thymidine chromophore, indicating a possibility of being the product (**55**) of RavIM. These results showed that RavIM might be catalyzing the isomerization (**18**→**55**) in a slow rate, and the product might be spontaneously decomposing to TDP in the absence of aminotransferase RavAMT. Our several efforts to isolate and characterize **55** were unsuccessful due to its poor yield and instability.

To carry out the coupled assay with RavIM, RavAMT was overexpressed in *E.coli* and the recombinant protein was purified using IMAC. Incubation of **18** with RavIM and RavAMT in the presence of L-glutamic acid and pyridoxol-5-phosphate (PLP) for 30 minutes resulted in the appearance of a new peak at 12.8 min as a major compound and TDP peak as a minor compound (**Figure 3.3**). The new peak was collected from the HPLC and desalted through gel filtration. Low resolution ESI mass analyses revealed a molecular ion peak at 546 which matched with the molecular weight of **56** (molecular formula $C_{16}H_{27}N_3O_{14}P_2$, cal. $[M-H]^- =$ 546.096). The negative mode high resolution ESI-MS revealed a molecular ion peak at 546.0922 which confirmed the suggested molecular formula. The ^1H NMR of this compound revealed C-5 methyl protons as a doublet at δ 1.16. The proton at C-5 position appeared as a quartet at δ 4.25. The presence of a broad doublet at δ 3.93 was indicative of an equatorial proton at C-4 position. The proton at C-3 position appeared as a double doublet at δ 3.63 with the coupling constants *J* =11.0 Hz and *J* =3.0 Hz. This further confirmed the presence of an equatorial proton at C-4 and axial hydrogen atom at C-2 positions. Likewise in **17** and **18**, the proton at C-2 and C-1 appeared as a double triplet at δ 3.88 and a double doublet at δ 5.54 ppm, respectively, as a result of coupling with the phosphorous atom of the anomeric phosphate group.
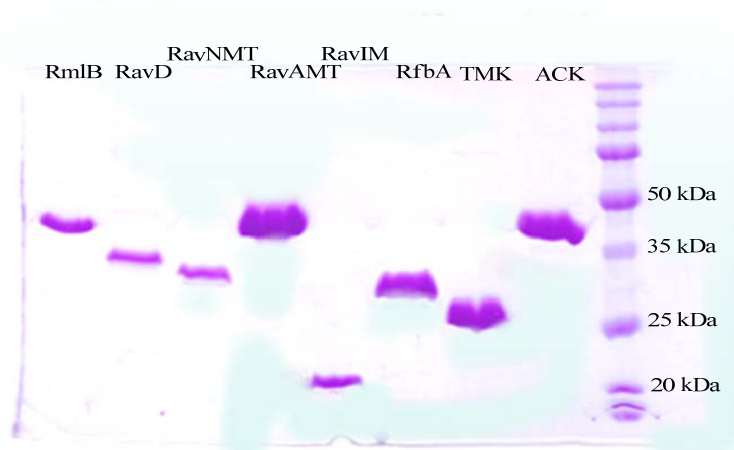
Figure 3.5 SDS-PAGE analyses of TDP-D-ravidosamine pathway enzymes. The sizes of the observed protein bands were in good agreement with the calculated values (RmlB, 40.5 kDa; RavD, 40.37 kDa, RavNMT, 31.06 kDa; RavAMT, 41.71 kDa; RavIM, 19.1 kDa; RfbA, 34.54 kDa; TMK, 25.9 kDa; ACK, 45.41 kDa).

The assignment was made by analyzing the coupling constant of the ^{1}H NMR and ^{1}H, ^{1}H-COSY correlations (**Table 3.2**). As a positive control, FdtA and FdtB were also expressed and purified to compare their activities with RavIM and RavAMT, respectively. Identical HPLC profiles were observed when RavIM and RavAMT were replaced with FdtA and FdtB, respectively. However, the specific activity of FdtA (201 μM mg^{-1} min^{-1}) appeared to be ~6 fold higher than that of RavIM (31.3 μM mg^{-1} min^{-1}). These results clearly demonstrated the identical biochemical actions of these isomerases and aminotransferases. To test whether RavIM is a bifunctional enzyme (i.e. isomerase/acetyltransferase), **18** was incubated with RavIM and acetyl-CoA. The HPLC profile of the assay product was identical to that of the control mixture without acetyl-CoA. Similarly, addition of acetyl-CoA did not alter the products of RavIM and RavAMT. These results clearly established that RavIM is not a bifunctional enzyme but it is, in fact, a 3,4-keto-isomerase that converts **18** to **55**. Our further attempts to isolate **55** in large scale using FdtA were not successful as the compound decomposed during purification. Comparing the active site residues of RavIM with FdtA, we proposed a mechanism of isomerization catalyzed by RavIM. His-71 might abstract a proton from the C-3 carbon of **18** to generate ene-diol **63**. In the next step, His-73 might abstract another proton from the C-3 hydroxyl group of **63** and channels to the rear face of C-4 carbon to yield **55** (**Scheme 3.2**). Alternatively, His-71 could

abstract both protons sequentially and His-73 prior to the transfer to the C-4 position. It is not clear yet whether or not the same C-3 proton of **18** migrates to the C-4 positon.
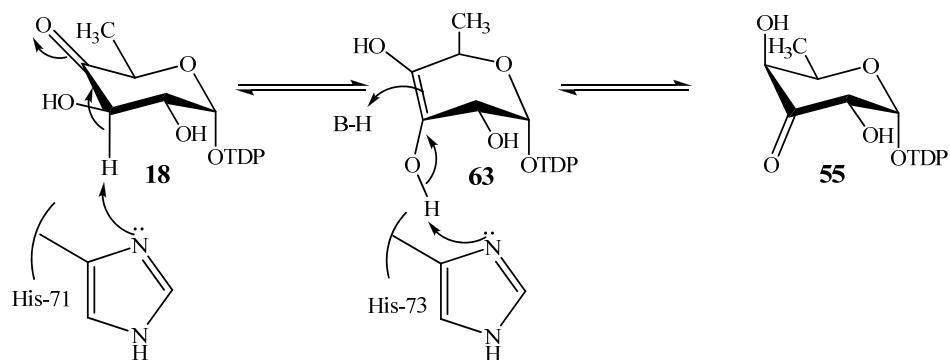
Scheme 3.2 Proposed mechanism for RavIM catalysis
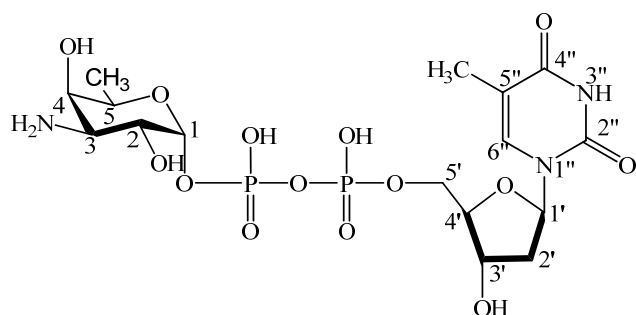
TDP-3-amino-3,6-dideoxy-D-galactose (**55**)

Table 3.2 ^{1}H NMR data of TDP-3-amino-3,6-dideoxy-D-galactose (**55**) (D$_2$O, 500 MHz)

Position number	^{1}H δ [ppm]	Multiplicity (*J*/Hz)	^{1}H, ^{1}H-gCOSY correlations
1	5.54	dd (6.9, 3.5)	2
2	3.88	td (11.1, 3.5)	1, 3
3	3.63	dd (11.1, 3.1)	2, 4
4	3.93	br, d (3.1)	3, 5
5	4.25	q (6.4)	4, 5-CH$_3$
5-CH$_3$	1.16	d (6.4)	5
1'	6.28	t (6.8)	2'
2'	2.36-2.28	m	1', 3'
3'	4.55	br, m	2' 4'
4'	4.10-4.16	m	5', 3'
5'	4.10-4.16	m	4'
5''-CH$_3$	1.87	s	6''
6''	7.66	br, s	5''-CH$_3$

Biochemical studies of FdtA

Although, the functional role of FdtA had been established, the biochemistry of its catalysis including the kinetics and the products were not characterized. Because FdtA was catalytically superior to RavIM in *in vitro* assays, it was chosen for a detailed biochemical characterization. First, we monitored the FdtA catalysis in various time points. Analysis of the

assay mixture from 0 to 110 minutes revealed gradual decrease of **18** with the increasing amount of TDP (**65**) (**Figure 3.6**). The peak corresponding to **55** gradually increased from 0-55 minutes and remained unchanged. However, a new peak at 2.9 minutes increased exponentially throughout the incubation. This indicated that **55** was not stable under the experimental conditions and it decomposed to TDP and **73**. The results were clearly better visualized in the graph plotting the normalized area vs the incubation time (**Figure 3.7**).
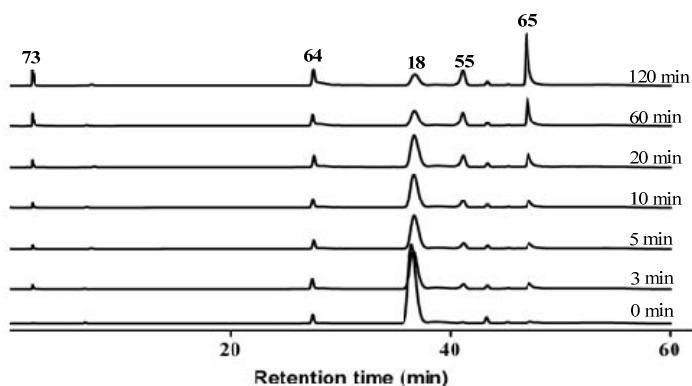
Figure 3.6 HPLC analyses of isomerization reaction catalyzed by FdtA. Assay mixtures were quenched following increasing incubation periods (0-120 min) and were analyzed in HPLC. Thymidine monophosphate (TMP, **64**); TDP-4-keto-6-deoxy-D-glucose (**18**); TDP-3-keto-6-deoxy-D-galactose (**55**); thymidine diphosphate (TDP, **65**); (2*R*,3*S*)-2-methyl-3,5-dihydroxy-4-keto-2,3-dihydropyran (**73**).
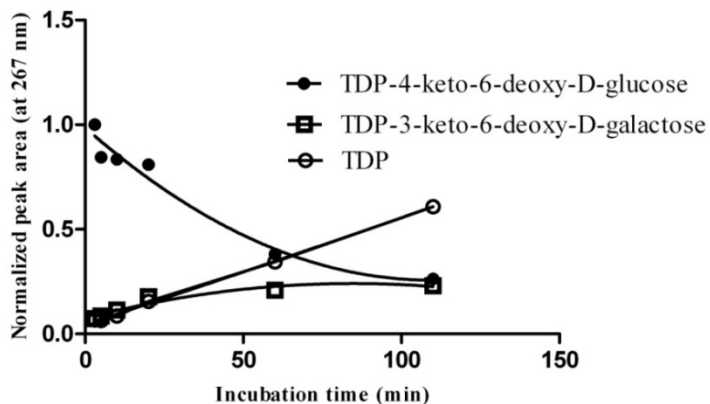
Figure 3.7 Plot of the normalized HPLC peak area of the compounds in the assay mixture using FdtA.

Kinetic studies with FdtA

Because the product of FdtA (**55**) was not stable, the consumption of **18** was monitored to determine the initial velocity of the isomerization reaction. Compound **18** was incubated with FdtA in the presence of increasing substrate concentrations for different time periods (see materials and methods section) and the plot of the initial velocity versus substrate concentration was generated (**Figure 3.8**). The observed values for K_m and k_{cat} were 21.89±6.994 (μM) and 8.95 min^{-1}, respectively. The observed turnover of FdtA was comparable to that of Tyl1a (6.5 min^{-1}) from the TDP-D-mycaminose pathway.
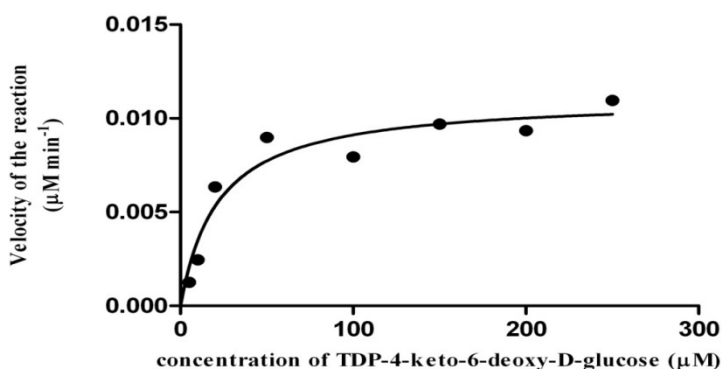
Figure 3.8 Plot of initial velocity (v_0) vs substrate concentration [S].

Functional characterization of RavNMT

Homology searches showed that the encoded amino acid sequence of *ravNMT* is similar to other known deoxysugar *N,N*-methyltransfarases. These homologous enzymes were known to methylate the amino group of the deoxysugar twice sequentially using *S*-adenosyl methionine (SAM) as a co-factor. Well known examples are DesVI and OleM1 from the pikromycin and the oleandomycin pathways (*68, 136*), respectively. In context with the RMV biosynthesis, we assumed that RavNMT might catalyze a similar reaction (**56→19**) utilizing the product of RavAMT (**56**) as the substrate. To validate the proposed function of RavNMT, the enzyme was heterologously produced in *E. coli* and purified through IMAC (**Figure 3.5**). Incubation of TDP-3-amino-3,6 dideoxy-D-galactose (**56**) with RavNMT in the presence of SAM resulted in the complete consumption of **56** with concomitant production of a new peak at 4.8 min in HPLC

(**Figure 3.3**). The compound pertaining to the new peak was isolated and purified through HPLC and gel-filtration.
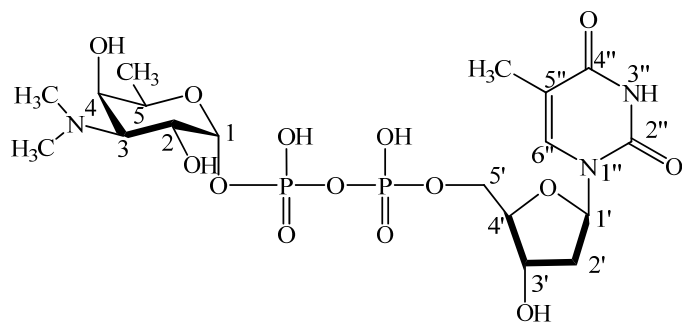
TDP-D-ravidosamine (**19**)

Table 3.3 ^{1}H NMR data of TDP-D-ravidosamine (**19**) (D_2O, 500 MHz)

Position number	[1]H δ [ppm]	Multiplicity (J/Hz)	^{1}H, ^{1}H-COSY
1	5.64	dd (6.9, 3.6)	2
2	4.21	m	1, 3
3	3.60	dd (11.4, 2.7)	2, 4
4	4.21	m	2, 5
5	4.25	q (6.6)	4, 5-CH₃
5-CH₃	1.19	d (6.6)	5
1'	6.33	t (6.8)	2'
2'	2.29-2.38	m	1', 3'
3'	4.6	br, m	2' 4'
4'	4.10-4.16	m	5', 3'
5'	4.10-4.16	m	4'
5"-CH₃	1.91	d (1.1)	6"
6"	7.7	d (1.1)	5"-CH₃

High resolution ESI mass revealed a *quasi* molecular ion peak of 574.1227 dalton which is in agreement with the molecular formula $C_{18}H_{31}N_3O_{14}P_2$ of the proposed product **19** (calcd.

MW, [M-H]$^-$ = 574.1281). In ^{1}H NMR spectrum, the appearance of a new singlet at δ 2.94 was a clear indication of the presence of two methyl groups at the amino nitrogen of **19**. Chemical shifts for the other protons of **19** were assigned through the analyses of ^{1}H, ^{1}H-COSY correlations. The detailed ^{1}H NMR data of **19** are summarized in **table 3.3**. The negative mode high resolution ESI mass spectra of the reaction mixture revealed molecular ion peaks of 574.1227, 560.1073 and 546.0922 daltons which corresponded to the molecular ions of **19**, **60** and **56** at negative mode, respectively. This indicated that RavNMT catalyzes the transfer of methyl groups from SAM to the amino nitrogen of **56** sequentially.

RavW, a candidate enzyme for *O*-acetylation

The product of *ravW* was the only candidate acetyltransferase identified in the gene cluster. Homology searches showed a modest similarity to *N*-acetyltransferases (GCN5 family enzymes). Ravidomycin contains an acetyl group at the C-4′ position, and the amino group at C-3′ position is dimethylated. However, other RMV analogues FE35a (**30**) and FE35B (**31)** which contain *N*-methyl, *N*-acetyl or *N*-methyl, *N*-acetyl and 4-*O*-acetyl functionalities in their sugar moieties have been identified (*85*). In this context, three different routes could be proposed for the biosynthesis of 4′-*O*-acetyl-TDP-D-ravidosamine: i) acetylation of 4′-OH group of **56** at the very last step of the pathway (path I, **Scheme 3.1**); ii) *N*-acetylation of **56**, acyl shift and dimethylation at amino group (path II); iii) *N*-methylation of **56**, *N*-acetylation of **61**, N→O-acyl shift and *N*-methylation of **62** (path III). In order to test these hypotheses, we cloned and expressed RavW in *E. coli*. The protein was expressed only in the form of insoluble inclusion bodies. Unfortunately, our several efforts to optimize expression of RavW in soluble form were not successful. Further experiments directing the prodcution of soluble RavW would be crucial to address the validity of the hypothesized function of RavW.

One-pot enzymatic synthesis of TDP-D-ravidosamine

One of the major goals of this study was to generate a sugar donor substrate for RavGT for *in vitro* glycosylation studies. An identification of sugar donor and acceptor substrates of RavGT would set a framework for assessing its substrate flexibility towards substrate variants and would also provide a means to generate RMV analogues *in vitro* as well as substrate analogoues. This required an easy protocol to generate the substrate for RavGT. Having all of the

biosynthetic enzymes in hand, we attempted to set up a protocol to generate the anticipated sugar donor substrate (TDP-D-ravidosamine) of RavGT in a large scale one-pot setup. However, the fairly expensive starting material TTP would limit the larger scale production of TDP-D-ravidosamine. Therefore, we cloned the enzymes necessary for the biosynthesis of TTP from *E. coli* following a reported protocol (*160*). Thymidine monophosphate kinase (TMK) phosphorylates TMP to generate TDP which is further phophorylated to generate TTP through the activity of acetate kinase (ACK). ATP serves as a phosphate donor and converts into ADP, where acetate kinase regenerates ATP again by transferring the phosphate group of acetyl phosphate to ADP (**Scheme 3.3**).
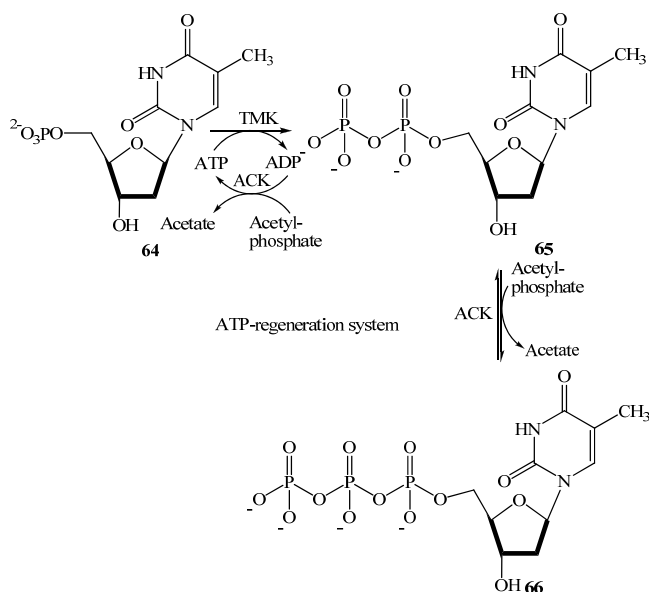
Scheme 3.3. Enzymatic synthesis of TTP (**66**) from TMP (**64**) using acetate kinase (ACK) and thymidine monophosphate kinase (TMK).

Both proteins were over expressed in *E. coli* and purified through IMAC (**Figure 3.5**). Incubation of TMP with acetyl phosphate and catalytic amount of ATP resulted in the production of TDP and TTP in quantitive yield. Since these reactions were reversible, the reaction was driven forward by adding RfbA and G-1-P in the assay mixture (**Figure 3.3**). Considering the fact that TDP can serve as an inhibitor for the downstream pathway enzymes, overall TDP-biosynthesis was divided into two stages in the same pot. In the first stage, TMP (0.48 mM), G-1-P (0.8 mM) and ATP (45 μM) were incubated with ACK, TMK and RfbA. TDP-D-glucose was

produced with ~95% yield in 5 minutes using these enzymes (see detail in the experimental section).
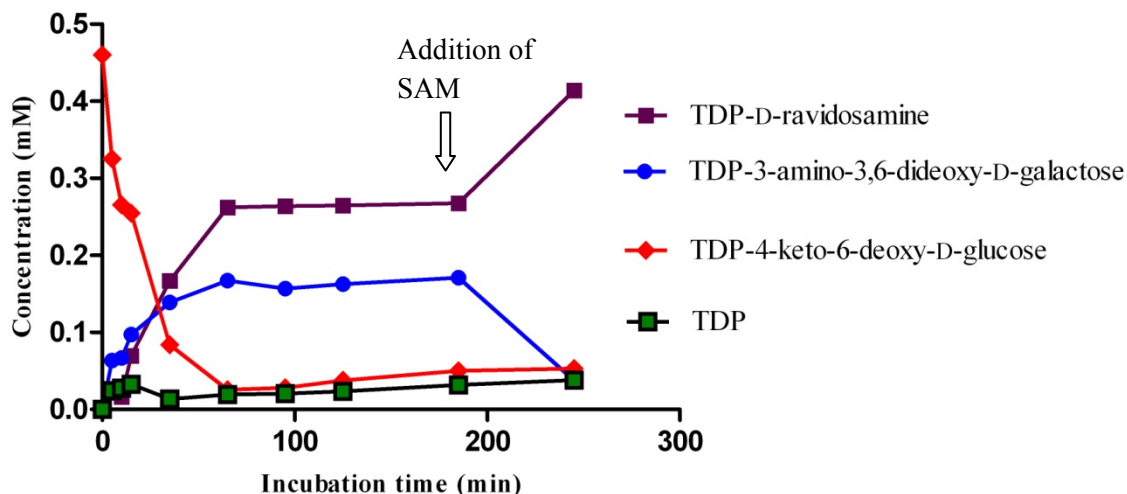
Figure 3.9 The conversion profile of TDP-4-keto-6-deoxy-D-glucose to TDP-D-ravidosamine at the second stage.

The enzymes in the assay mixture were denatured by heating at 95 °C for 4 minutes and then removed by centrifugation. After several batches of protein production, both of the recombinant 4,6-dehydratases (RavE and RfbB) were expressed mostly in insoluble form. Therefore, an additional TDP-glucose-4,6-dehydratase (encoded by *rmlB*) was cloned from *E. coli*, and the protein was produced and purified following the identical protocol described in materials and methods. Addition of RmlB in the assay mixture resulted in the complete conversion of TDP-D-glucose (**17**) to TDP-4-keto-6-deoxy-D-glucose (**18**). RavIM, RavAMT, RavNMT, PLP, L-glutamate and SAM were added in the assay mixture to complete TDP-D-ravidosamine biosynthesis. More than 95% of TDP-4-keto-6-deoxy-D-glucose (**17**) was consumed in 65 minutes (**Figure 3.8**) with the concomitant productions of TDP-3-amino-3,6-dideoxy-D-galactose (**56**) and a mixture of TDP-*N*-demethyl-D-ravidosamine (**60**) and TDP-D-ravidosamine (**19**). Identical retention times for **60** and **19** prevented us to analyze their individual concentrations independently. The productions of **19**, **56** and **60** remained unchanged from 60 to180 min which was a clear indication of the lack of SAM in the assay mixture. Addition of excess SAM resulted in a dramatic increase in production of **19** with the

consumption of **56** as well as **60** in additional 60 minutes, as evident from ^{1}H NMR analyses of the products. The overall yield of TDP-D-ravidosamine was estimated to be 85%.

Enzymatic synthesis of TDP-D-fucofuranose

D-fucofuranose is an uncommon sugar moiety found among glycosylated natural products. It is particularly interesting for biosynthetic studies because of its crucial role in conferring the biological activity of GV, and also due to its unique ring-contraction chemistry. Recently, the TDP-D-fucofuranose (**22**) pathway was characterized from *Escherichia coli* O52 where D-fucofuranose is a subunit of surface *O*-antigen (*161*). The pathway involved four enzymes. The first two enzymes (RmlA and RmlB) were involved in the conversion of **17** to **18** (**Scheme 3.4**). Fcf1 then utilized NADPH to reduce the C-4 carbonyl group of **18** to generate TDP-D-fucose (**67**). Fcf2 was the unique gene of the pathway which converted **67** to **22**. The enzyme did not require any cofactors and the reaction progressed with the contraction of pyranose ring to the furanose ring (**Scheme 3.4,** path A). Since, D-fucofuranose is the component of GV, a similar set of biosynthetic genes could be expected in the GV cluster. While GilD, GilE and GilU were functionally identical to RmlA, RmlB and Fcf1, respectively, no homologue of the *fcf2* gene was found in the GV cluster. This indicated that there might be a structurally very distinct enzyme in the GV pathway to catalyze the identical reaction, or ring contraction could happen after glycosylation (Path B) through the activity of another enzyme whose functional role had not been established. Comparison of the GV and RMV gene clusters outlined GilN or GilL as candidate enzymes which could be contributing to the ring contraction in the D-fucofuranose pathway (discussed in section 2). GilN showed 60-90% homology to GTs from the actinomycetes and belonged to the inverting GT-A category. This could be due to the presence of a common sugar binding pocket in GilN and GTs. Thus, there could be two routes to establish the D-fucofuranose moiety of GV (**Scheme 3.4**): i) through the contraction of the pyranose ring of TDP-D-fucose (**67**) before the *C*-glycosylation (path A); ii) through the ring contraction of the pyranose moiety of **68** after glycosylation (path B). To test these hypotheses, availability of GilN, GilL and GilU was crucial. Therefore, we attempted expression of these enzymes in *E. coli* several times. Unfortunately, all of them were only expressed in insoluble forms. Neither the culture induction conditions nor different expression vectors helped to solubilize these proteins. Future expression of these genes in homologous hosts such as *S. lividans* TK24 or *S. albus* would

possibly produce these proteins in the soluble form and would provide a stage to validate the proposed hypotheses.
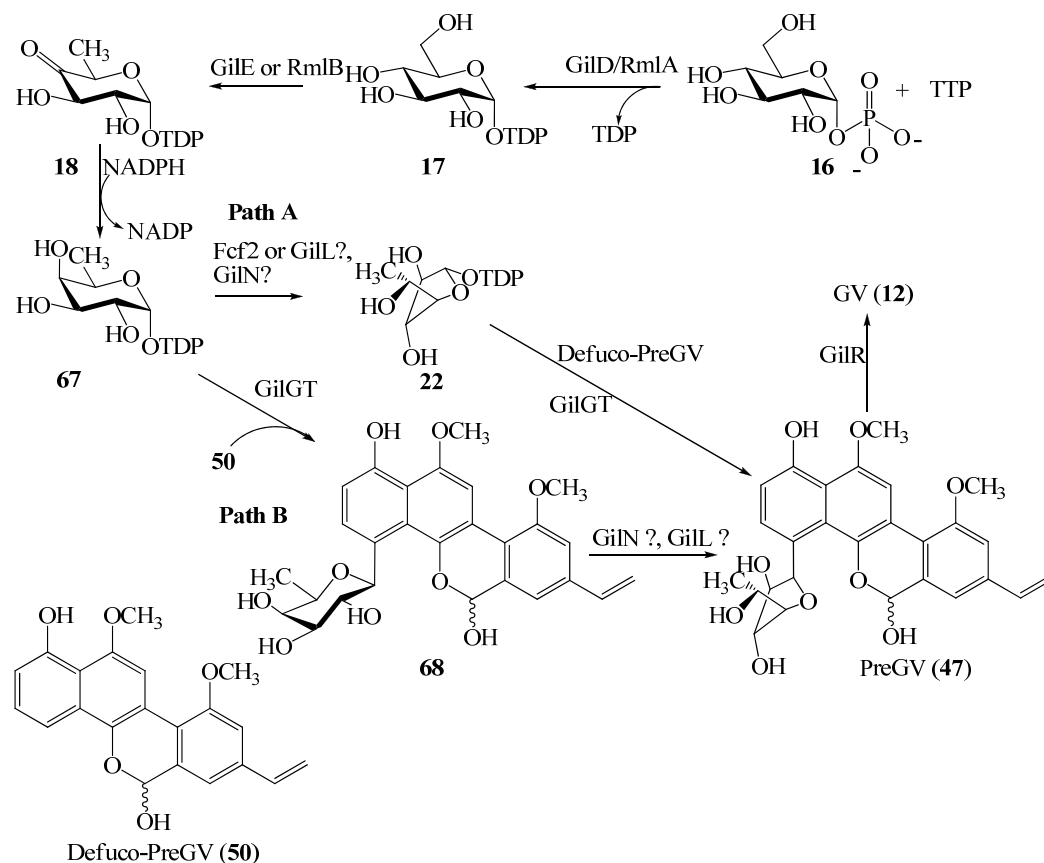
Scheme 3.4 A hypothesized pathway for the biosynthesis of TDP-D-fucofuranose (**22**)

Discussion

In vitro assay results clearly demonstrated that RavD preferentially utilizes TTP (**66**) among NTPs to activate G-1-P (**16**) and the activated sugar converts into TDP-4-keto-6-deoxy-D-glucose (**18**) through the activity of RavE. These two steps are common to a variety of deoxysugar biosynthetic pathway such as TDP-D-olivose (*162*), TDP-L-olivose (*38*), TDP-D-mycarose (*22*), TDP-L-mycarose (*157*), TDP-D-fucose (*163*) and TDP-L-rhamnose (*164*). Isomerization of **18** to **55** catalyzed by RavIM was the unique feature of the TDP-D-ravidosamine biosynthetic pathway. Such a conversion was further confirmed by the formation of an identical product in a parallel assay involving FdtA. Thus far, no such isomerization has

been observed in secondary metabolite biosynthetic pathways. FdtA from *Aneurinibacillus thermoaerophillus* L420-91 was the only enzyme known to catalyze such a transformation during the biosynthesis of 3-acetamido-3,6-dideoxy-D-galactose. However, the lack of suitable separation conditions for the substrate and product of FdtA hindered the earlier efforts to carry out complete biochemical characterization. Taking advantage of the current analytical system, we were able to fully characterize FdtA. Kinetic studies showed that the turnover of FdtA ($k_{cat} = 8.95$ min^{-1}) was comparable to the recently characterized Tyl1a ($k_{cat} = 6.10$ min^{-1}) from the mycaminose pathway. Notably, the TDP-D-mycaminose biosynthetic pathway was identical to the TDP-D-ravidosamine pathway except the isomerization step where Tyl1a converts **18** to **53** with equatorial facing hydroxyl group at C-4 position (**Scheme 3.5**). Incorporation of ravidosamine in place of the mycaminose moiety of a pikromycin derivative through the expression of FdtA in Tyl1a-deletion mutant has further demonstrated the close resemblance of these pathways (*165*). Unfortunately, the small turnover of the isolated RavIM hindered us from the determination of its kinetic parameters. RavIM appeared to have ~6.4 fold less specific activity compared to FdtA. Therefore, FdtA was used to study the 3,4-ketoisomerization. Isolation and characterization of **55** would provide direct evidence for this reaction. We were unable to isolate **55** due to its instability under experimental conditions. While monitoring FdtA's activity, continuous production of TDP (**65**) with the consumption of **18** was observed, but the concentration of **55** remained constant after 55 minutes. This suggested a parallel isomerization and decomposition reaction sequence in the assay mixture. Compound **55** could easily undergo keto-enol tautomerization to generate **72** which could undergo decomposition spontaneously resulting in the formation of (2*R*,3*S*)-2-methyl-3,5-dihydroxy-4-keto-2,3-dihydropyran (**73)** and TDP (**65)** (**Scheme 3.5**). There was precedence for the similar transformation in the mycaminose pathway where the Tyl1a product **53** spontaneously degraded to TDP and (2*R*,3*R*)-2-methyl-3,5-dihydroxy-4-keto-2,3-dihydropyran (**71)**.
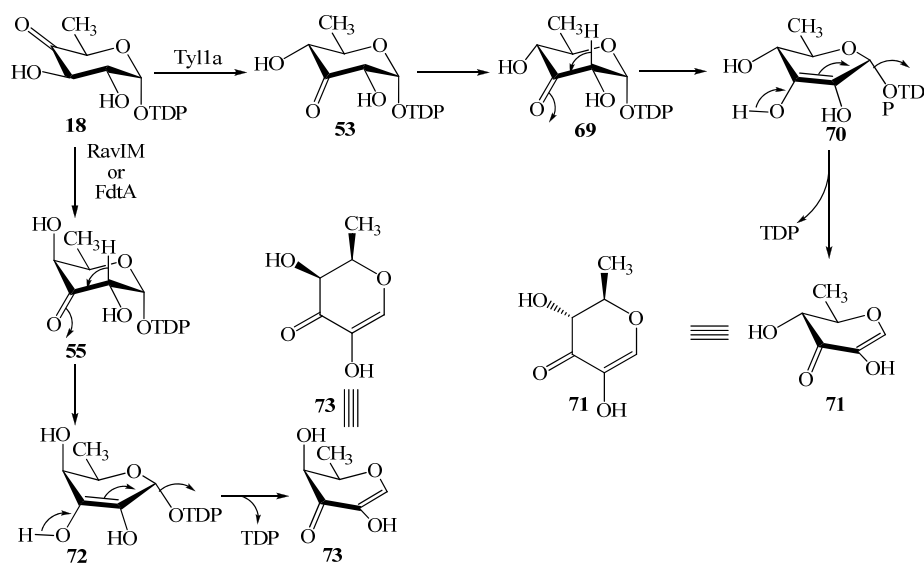
Scheme 3.5 Degradation mechanisms for TDP-3-keto-6-deoxy-D-glucose and TDP-3-keto-6-deoxy-D-galactose.

The coupled enzyme assay of RavAMT and RavIM catalyzed production of **56** from **18** and provided indirect evidence that RavAMT is an aminotransferase which transaminates the product of RavIM. Similarly, the conversion of **56** to **19** through the activity of RavNMT unambiguously proved the enzyme to be an *N,N*-dimethyltransferase. Identification of both **19** and **60** in the assay mixture favored the hypothesis of sequential transfer of methyl groups from SAM to the amino nitrogen. DesVI and TylM1 from desosamine and mycaminose biosynthetic pathways were other examples of such *N,N*-dimethyltransferases which catalyze dimethylation of a sugar amino group sequentially (*68, 166*). Whether RavW is involved in 4′-*O*-acetyl-D-ravidosamine biosynthesis or whether it serves as a part of RavPKS enzyme complex could not be addressed due to the lack of soluble protein. The timing and the sequence of the acetylation (path I -III, **scheme 3.1**) were not clear as well. Clearly, RavIM is not responsible for the acetylation. Upon the production of soluble enzyme, it would be highly interesting to clarify these ambiguities regarding the functional role of RavW.

Reports on one-pot enzymatic syntheses of NDP-sugars are rapidly growing. Synthesis of TDP-L-rhamnose,(*164*) TDP-L-mycarose (*157*), TDP-4-keto-6-deoxy-D-glucose (*167*) are typical representatives. With all of the functional biosynthetic enzymes in hand, we also devised an one-pot enzymatic synthesis protocol for the quick and efficient production of TDP-D-ravidosamine (**19**) in a cost-effective way. Instead of using expensive commercial TTP, the compound was

generated *in situ* from TMP and acetyl phosphate using acetate kinase (ACK) and thymidine monophosphate kinase (TMK) from *E. coli*. An identical strategy was utilized for the enzymatic synthesis of TDP-L-rhamnose (*164*), TDP-4-keto-6-deoxy-D-glucose (*44*), and TDP-4-amino-4,6-dideoxy-D-glucose (*167*). The first stage synthesis resulted in the production of TDP-D-glucose (**17**) in a satisfactory yield (~95%). Further incubation of the 1st stage product with the remaining TDP-D-ravidosamine biosynthetic enzymes resulted in the efficient production of TDP-D-ravidosamine (~85% at the second stage). Suzuki et al. had developed a chemical synthesis protocol for the production of TDP-D-ravidosamine (*168*). Like the synthesis of other TDP-sugars, the reported procedure involved lengthy synthetic steps with overall poor yield. In this context, the enzymatic synthesis approach presented in this work would be a better method for the mass-production of TDP-D-ravidosamine.

In summary, all of the proposed biosynthetic enzymes were characterized *in vitro* to achieve the enzymatic synthesis of TDP-D-ravidosamine. Overall, the biosynthetic pathway closely resembles the biosynthesis of TDP-D-mycaminose (**54**), which is different only with respect to the isomerization step. The RavIM-mediated isomerization step of the TDP-D-ravidosamine biosynthetic pathway was further studied using recombinant FdtA. The one-pot enzymatic synthesis protocol developed in this study would be a valuable tool to produce TDP-D-ravidosamine in large scale for routine use of this compound for *in vitro* activity assay of RavGT and other GTs.

Experimental section

Cloning and preparation of expression constructs

Oligonucleotide primers used in this study are summarized in the **table 3.4**. *E. coli* XL1 Blue was used for the routine amplification of plasmid DNA. *E. coli* BL21 was used for the expression of all of the constructs listed in the **table 3.5**. Genomic DNA of *Salmonella enterica serover* typhimurium LT2*, E. coli* XL1 Blue and *Aneurinibacillus thermoaerophillus* L420-91 were prepared following a standard protocol. The TDP-glucose-4,6-dehydratase gene (*rmlB*), acetate kinase gene (*ack*) and thymidine monophosphate kinase gene (*tmK*) were amplified from the genomic DNA of *E. coli*. Similarly, *rfbA*, *rfbB*, and *fdtA* and *fdtB* were amplified from the genomic DNA of *Salmonella typhimurium* LT2 and *Aneurinibacillus thermoaerophillus* L420-

91, respectively. All of the TDP-D-ravidosamine biosynthetic genes were amplified from the cosmid cosRav32.

Expression and purification of proteins

Pfu polymerase (Stratagene) was used for PCR amplification and the products were cloned into the ZeroBlunt-TOPO vector (Invitrogen). Primers used in this study are listed in **table 3.4**. Sequencing of the TOPO clones were carried out to ascertain the perfect fidelity of the PCR amplification. The acetate kinase (ACK), *ravD, E, IM, AMT, NMT* and *rfbA* were isolated from their TOPO-clones with *Nde*I/*Eco*RI fragments and ligated at the identical sites of pET28a(+) expression vector (Novagen). Similarly, the thymidine monophosphate kinase gene (*tmk*), *rfbB* and *rmlB* were ligated at *Eco*RI/*Hind*III and *Bam*HI/*Hind*III and *Bam*HI/*Xho*I sites of pET28a vector, respectively. All the expression constructs generated in this study are summarized in **table 3.5**. The constructs were then transformed into *E. coli* BL21(DE3) host. A single colony was inoculated into 20 mL of LB liquid supplemented with kanamycin sulfate and was grown for 4 h to prepare a seed culture. A liter of LB (100 mL × 10 flasks) was inoculated with 10 mL of the seed and grown the culture at 37 °C until OD_{600} reached to 0.5. The culture was grown at 23 °C for 12 h following addition of IPTG (0.05 mM final concentration). The cell pellets were collected by centrifugation (4000 × g, 15 min) and were washed twice with 20 mL of lysis buffer (50 mM KH_2PO_4, 300 mM NaCl, 10 mM imidazole, pH 7.6). Disruption of the pellets was carried out using a French Press (Thermo electron corporation), and the crude soluble enzyme fraction was collected through centrifugation (17000 × g). *N*-terminal His-tagged enzymes were purified through immobilized metal affinity chromatography (IMAC) which is further desalted through the Profinia protein purification system (Bio-Rad). The concentration of the enzyme was determined using the Bradford protein assay method (*169*). The purities of the enzymes were estimated by sodium dodecylsulfate-polyacrylamide genel electrophoresis (SDS-PAGE) analyses.

Table 3.4. List of the primers used in the experiments (deoxysugar biosynthesis)

Name of the primers	Oligonucleotide sequence
ACK_ex_for	5′-TCGCATATGTCGAGTAAGTTAGTACTGGTT-3′
ACK_ex_rev	5′-TCGGAATTCTCAGGCAGTCAGGCGGCTCGCGTC-3′
TMK_ex_for	5′-GCTGAATTCATGCGCAGTAAGTATATCGTCATTGAG -3′
TMK_ex_rev	5′-CTGGAAGCTTCATGCGTCCAACTCCTTCACCCA -3′
RmlB_for	5' -TTGGATCCATGAAAATACTTGTTACTGGTGGCGCAGGA -3′
RmlB_rev	5′-ATTCTCGAGTTACTGGCGGCCCTCATAGTTCTGTTC-3′
RavAMT_ex_for	5′-GCGCATATGAAGGTCCCCTATCTGGACCTGAAG-3′
RavAMT_ex_rev	5′-GTGGAATTCTCAGACCGCGGCGCGCACCGCTTC-3′
RavNMT_ex_for	5′-GCGCATATGAGCACCCTCTCGGTGAGCCAGCCC-3′
RavNMT-ex_rev	5′-TCGGAATTCTCATCGCGGTCCCCCGTCGGTCGT-3′
RavE_ex_for	5′-CTCCATATGACCTCGACACACATCCTGGTG-3′
RavE_ex_rev	5′-TAGGAATTCTCACACCATCGCGCGCTCCTTCAA-3′
RavD_ex_for	5′-CCGCATATGAAGGCCCTCGTCCTGTCCGGC-3′
RavD_ex_rev	5′-TGTGAAGCTTGGTCATGAGGAGATCTGCACCTT-3′
RavIM_ex_for	5′-ACCCATATGACCGACACGACCGCCGCGACC-3′
RavIM_ex_rev	5′-AGGGAATTCTCATGGGCGCGCTCCGGCATCCGC-3′
RfbA_ex_for	5′-TGCATATGAAAACGCGTAAGGGCATTATTTTA-3′
RfbA_ex_rev	5′-TATGAATTCTTATAAACCTTTCACCATCTTCAGCAA-3′
RfbB_ex_for	5′-CGGGATCCATGGTGAAGATACTTATTACTGGCGGG-3′
RfbB_ex_rev	5′-TCAAGCTTGTCACTGGCGTCCTTCATAGTTCTGTTC-3′
FdtA_ex_for	5′- AGAGGATAACATATGGAAAATAAAGTTATTAACT-3′
FdtA_ex_rev	5′-TATGAATTCTTATCCTTCTTTTTCTAGATTAATTTTA-3′
FdtB_ex_for	5′- TGGGATCCATGATTCCTTTTTTGGATTTAAGACAA-3′
FdtB_ex_rev	5′-ATCTCGAGCTAGTATCCATACCTGTTAACAGC-3′

Table 3.5. Bacterial strains and plasmids used in this study.

Strain/ plasmid	Characteristics and relevance	References
E. coli XL1-Blue-MRF	Host for routine cloning works and for the construction of genomic libraries	Stratagene
E. coli BL21 (DE3)	Expression host for the various expression constructs	Invitrogen
Salmonella typhimurium LT2	For cloning deoxysugar biosynthetic genes	Liu et al. (*131*)
A.neurinibacillus thermoaerophillus L420-91	For cloning deoxysugar biosynthetic genes	Messner et al.(*132*)
Streptomyces lividans TK24	For expression of gilvocarcin PKS enzymes	Wright et al.(*170*)
pRfbA	*rfbA* gene cloned into pET28a	This work
pRfbB	*rfbB* gene cloned into pET28a	This work
pRavD	*ravD* gene cloned into pET28a	This work
pRavE	*ravE* gene cloned into pET28a	This work
pRavIM	*ravIM* gene cloned into pET28a	This work
pRavAMT	*ravAMT* gene cloned into pET28a	This work
pRavNMT	*ravNMT* gene cloned into pET28a	This work
pFdtA	*fdtA* gene cloned into pET28a	This work
pFdtB	*fdtB* gene cloned into pET28a	This work
pACK	Acetate kinase gene cloned into pET28a(+)	Kim et al.(*134, 164*)
pTMK	*tmk* gene cloned into pET28a(+)	This work
pET28a	Expression vector	Novagen
PCR-Blunt-II-TOPO	To clone PCR products	Invitrogen

Substrates and co-factors

TMP, TDP, TTP, UTP, GTP, CTP, ATP and TDP-glucose were purchased from Sigma-Aldrich for analytical scale enzyme assays. Acetyl phosphate, pyridoxal phosphate (PLP), L-glutamic acid and *S*-adenosyl methionine (SAM) were also purchased from Sigma-Aldrich.

Enzyme assays

Activity assay of RavD and preparation of TDP-D-glucose

Enzyme assays for TDP-glucose synthases (RavD or RfbA) were carried out in 200 μL mixture composed of 50 mM phosphate buffer (pH. 7.5), 20 mM $MgCl_2$, 5 mM G-1-P, 5 mM NTP (ATP, UTP, TTP, GTP and CTP) and 3 μM enzyme. The mixture was incubated for 30 minutes at 37 °C on the water bath. The reaction was stopped by heating the mixture at 90 °C for 3 minutes. The precipitated enzymes were removed by centrifugation and the clear supernatant fraction was passed through amicon filter (mol. Wt. cut off 3000 Da). The filtrate (50 μL) was injected in the HPLC. CarboPack^{TM} PA1 (4×250 mm) column (Dionex co.) was used to separate compounds. A gradient of ammonium acetate (500 mM, solvent B) and water (solvent A) was used to separate the assay products: 5-20% solvent B from 0-15 min, 20-60% solvent B from 15-35 min, 60-100% solvent B from 35-37 min, 100% solvent B from 37-40 min, 100-5% solvent B from 40-45 min and 5% solvent B from 45-60 min. A control sample without enzyme and standard TDP-D-glucose was also injected for the identification of the product peak. For the larger scale production of TDP-D-glucose, TMP (5 mM), G-1-P (5 mM), acetyl phosphate (20 mM) and ATP (5 μM) were mixed in 20 mL solution composed of 50 mM phosphate buffer (pH. 7.5), 20 mM $MgCl_2$, 4 μM ACK, 6 μM TMK and 5 μM RavD. The mixture was incubated for 1 hour and the reaction was quenched by heating the mixture at 90 °C for 5 minutes. The mixture was centrifuged at 10000×g for 10 minutes and the supernatant was lyophilized. The dry powder was dissolved in 5 mL of water and the solution was loaded into the pre-equilibrated Bio-Gel P2 gel filtration column 100 cm × 25 mm. Separation was carried out at 4 °C with water at flow rate of 0.2 mL min^{-1}. Aliquots of 4 mL were collected and the purity of compounds were monitored in HPLC at 267 nM. Pure fractions were combined, lyophilized and analyzed by NMR. About 80 mg of pure TDP-D-glucose was obtained in addition to 125 mg with minor amount TMP and TDP mixture. The purity was assayed through HPLC and ^{1}H NMR.

Activity assay of RavE and preparation of TDP-4-keto-6-deoxy-D-glucose

To carry out TDP-glucose-4,6-dehydratases assays, the enzymes (4 μM RavE or 3 μM RfbB or 3 μM RmlB) were added to 200 μL solution composed of phosphate buffer (50 mM, pH. 7.5), $MgCl_2$ (20 mM), NAD$^+$ (2 μM) and TDP-D-glucose (5 mM). The mixture was incubated for

30 minutes and the reaction was stopped. An enzyme-free solution was analyzed in the HPLC. For the larger scale production, a one-pot enzymatic synthesis was carried out in 50 mL volume composed of phosphate buffer (50 mM, pH 7.5), $MgCl_2$ (20 mM), glucose-1-phosphate (150 mM), TMP (98.3 mM), acetyl phosphate (400 mM), ATP (1mM), ACK (0.8 μM), TMK (2.1 μM), RfbA (2.0 μM). The mixture was incubated for 1 hour at 37 °C and then RavE (final concentration of 2.73 μM) or RfbB (2.3 μM) and NAD^+ (2 μM) were added and then the incubation was continued for additional 2 hours. Surprisingly, after a first few batches of such reactions, both RavE and RfbB appeared exclusively in the form of inclusion bodies. Therefore, the enzymes were replaced by RmlB (6 μM) in the following batches. The assay mixture was heated to denature enzymes. The precipitated fractions were removed by centrifugation. The crude product was lyophilized to powder and dissolved in 10 mL water. Separation of the products was carried out in the Bio-gel P2 following an identical protocol described earlier. The purity of each fraction was analyzed by HPLC. Pure fractions were combined and lyophilized to powder. This afforded 700 mg of TDP-4-keto-6-deoxy-D-glucose with ~95% purity and an additional 630 mg with about >70% purity. The purity and identity of the compound was further confirmed through the comparison of ^{1}H NMR data with the reported literature.

Activity assay of RavIM alone or coupled with RavAMT

RavIM (20 μM final concentration) was incubated at 37 °C with TDP-4-keto-6-deoxy-D-glucose (1 mM) in 500 μL solution composed of phosphate buffer (50 mM, pH 7.5) and $MgCl_2$ (20 mM) for 60 minutes. An aliquot (60 μL) was drawn from the incubation mixture and flash-freezed in dry ice/acetone. The enzyme was removed by further heating the mixture at 90 °C for 3 minutes. The solution was filtered through centricon (MW cut off 3000 Da) and the filtrate (50 μL) was analyzed by HPLC. In the coupled enzyme assays, RavAMT (30 μM), L-glutamate (10 mM) and PLP (60 μM) were added in the RavIM assay mixture and progress of the reaction was monitored until 60 minutes. The new peak appeared at 12.8 minutes in the coupled assays was collected and analyzed by High-resolution ESI-MS. Similarly, FdtA and FdtA-FdtB coupled assays were also carried out in parallel as positive controls. For the preparative scale production of TDP-3-amino-3,6-dideoxy-D-galactose (**56**), the substrate and enzyme concentrations were doubled and the assay volume was increased to 20 mL. After 2 hours of incubation, enzymes were removed and the volume was reduced to 1 mL. Separation of the compounds was carried

out in analytical scale HPLC. The compound eluted at 12.8 minutes was collected and separated from the ammonium acetate buffer using Bio-Gel P2 column. The pure fractions were combined and lyophilized. Through this method, 12 mg powder of TDP-3-amino-3,6-dideoxy- D-galactose (**56**) was purified. The structure of **56** was confirmed through NMR analyses.

Kinetics study of FdtA

Enzyme assay carried out in 100 μL solution was composed of phosphate buffer (50 mM, pH 7.5), magnesium chloride (20 mM), FdtA (8.4 μM) and various concentrations of **18** ranging from 0.05 to 25 μM. Samples with low (0.05 - 2 μM), medium (5 - 15 μM) and high concentration (20 - 25 μM) of substrates were incubated for 3 min, 5 min and 8 minutes, respectively. Reactions were stopped immediately by chilling dry ice/acetone bath (-90 °C). After completion of the assay, the samples were exposed to the boiling water for 5 minutes and centrifuged to remove precipated protein. The supernatant fraction (50 μL) was injected in the HPLC and the concentration of remaining **18** was determined by using a standard calibration curve. This allowed us to calculate decrease in concentration of substrate per unit time. Thus generated initial velocity data were plotted versus the concentration of **18.** The resulting data were fitted into the Michaelis-Menten equation by non-linear regression using PRISM to determine k_{cat} and K_m values.

HPLC analyses of FdtA-mediated isomerization and the decomposition of the TDP-3-keto-6-deoxy-D-galactose

TDP-4-keto-6-deoxy- D-glucose (**18**, 1 mM) was incubated at 37 °C with 24.9 μM enzyme in 400 μL solution containing 50 mM phosphate buffer and 20 mM magnesium chloride. Aliquots (60 μL) were drawn from the assay mixture and quenched the reaction in dry ice/acetone bath. The samples were boiled for 4 minutes and centrifuged to remove protein. The clear supernatant (50 μL) was injected in to the HPLC and the peak areas under the curve were normalized to present data in graph using PRISM.

Activity assay for RavNMT

For an analytical scale assay, 200 μL assay mixture containing 50 mM phosphate buffer, 20 mM $MgCl_2$, 1 mM of **56**, 2.0 mM SAM and 10 μM RavNMT was incubated for two hours.

Enzyme-free solution was injected in HPLC. The new peak, appeared at 4.8 minutes was collected and analyzed by HR-ESI Mass spectroscopy. For the larger scale production, a coupled assay was carried out in 20 mL using **18** (2 mM), L-glutamate (14 mM), PLP (80 μM), SAM (4.0 mM) RavIM (4.8 μM), RavAMT (6.5 μM) and RavNMT (8.0 μM). After two hours of incubation, the enzymes were removed and the solution was lyophilized. The dried powder was dissolved in 1 mL water and separated by HPLC. The compound eluted at 4.8 min was collected and desalted using Bio-gel-P2 column. Lyophilization of the pure fractions afforded 5 mg and 8 mg of TDP-D-ravidosamine (**54**) and a 1:1 mixture of **60** and **54**, respectively. The structures of the compounds were confirmed through the NMR and HR-MS analyses.

Two-stage one-pot enzymatic synthesis of TDP-D-ravidosamine

Overall TDP-D-ravidosamine synthesis was divided into two stages. For the first stage, 20 mL solution containing glucose-1-phosphate (0.8 mM), TMP (0.48 mM), ATP (45 μM), acetyl phosphate (2 mM), ACK (5 μM), TMK (10 μM), RfbA (8 μM), phosphate buffer (50 mM, pH 7.5) and $MgCl_2$ (20 mM) was incubated at 37 °C for 30 minutes. HPLC analyses of the assay mixture revealed 0.47 mM of TDP-D-glucose in the solution. Enzymes were precipitated by heating the solution to 95 °C and removed through the centrifugation. In the second stage, RmlB (5.8 μM) was added to the solution and incubated the mixture for 30 minutes. Broadening of the peak at 37.7 minutes indicated conversion of TDP-D-glucose (**17**) to TDP-4-keto-6-deoxy-D-glucose (**18**). RavIM (2.3 μM), AMT (3.9 μM), RavNMT (4.5 μM), L-glutamate (7 mM), PLP (42 μM) and SAM (0.7 mM) were then added to the assay mixture and continued incubation for 5 hours. Considering the fact that SAM degrades in water, an identical amount of SAM was added after 3 hours. The profile of conversion was monitored using HPLC. A calibration curve was generated varying the concentration of TMP from 62 nM to 15 mM. The concentrations of TDP-D-sugars were determined using this calibration curve. The reaction was stopped by heating at 90 °C for 4 minutes. The denatured protein was removed through centriguration and the liquid fraction was lyophilized. The peak corresponding to the TDP-D-ravidosamine was separated through HPLC. The pure fraction was desalted and then analyzed by NMR spectroscopy.

Section Four: Enzymatic synthesis of defucoGM

Background

Post-PKS tailoring reactions such as oxidation, dehydration, glycosylation and methylation were of our central focus as their in depth understanding was necessary to set a framework for successful combinatorial biosynthesis of GV analogues. Incorporation studies using ^{13}C-labeled acetate showed the loss of two carbon atoms during GV biosynthesis. Furthermore, isotope incorporation patterns were indicative of a C-C bond cleavage and an oxidative rearrangement of the backbone (**Scheme 4.1**) (*97*). To figure out the exact mode of this cleavage and the sequence of the post-PKS tailoring steps the entire GV biosynthetic cluster was cloned and sequenced in our lab (*71*). Most of the proposed post-PKS genes were inactivated individually and the metabolites from the mutant strains were isolated and analyzed (*33, 57, 101, 102*).
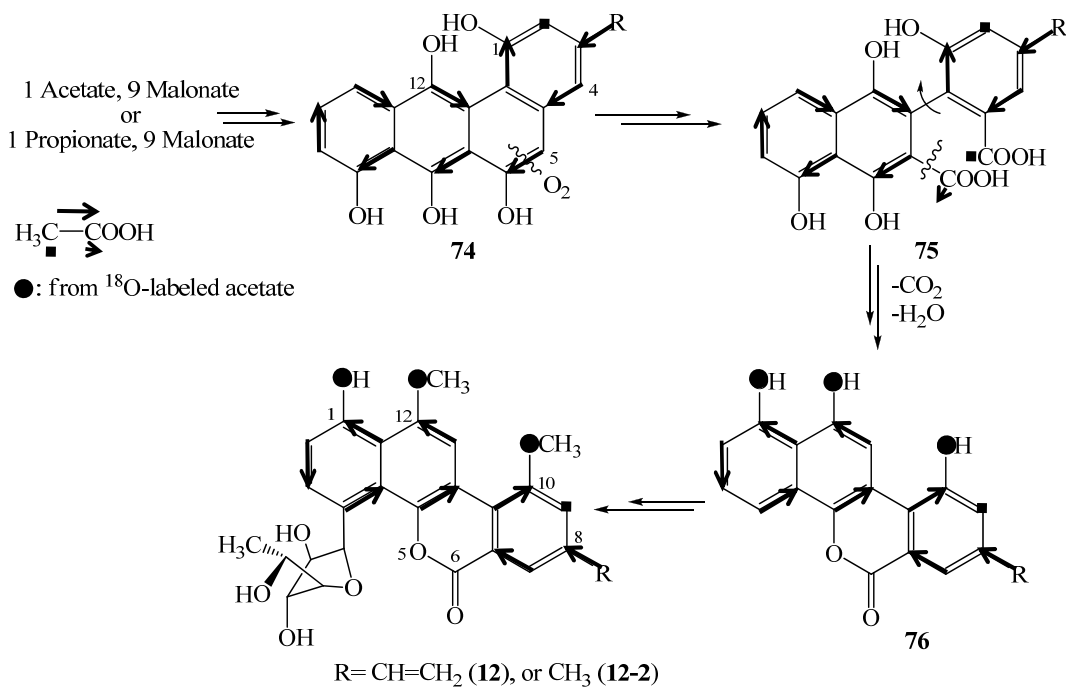
Scheme 4.1 Incorporation of labeled acetate in gilvocarcin (figure was adapted from Rohr et al.) (*57*).

GilOI, GilOII and GilOIV were the only candidate monooxygenases for the oxidative C-C bond cleavage. ΔGilOI and ΔGilOI-ΔGilOII strains accumulated 2,3-

dehydro-homo-UWM6 (**41**) as the major metabolite which gradually oxidized to pre-GE-o-quinone (**42-c**). Homo-rabelomycin (**48**) and dehydro-homo-rabelomycin (**49**) appeared to be the major metabolites of ΔGilOIV and ΔGilOIV-ΔGilOII, and ΔGilOII mutants, respectively (*56*). All of these accumulated metabolites had an intact angucyclinone backbone which confirmed the hypothesis that a C5-C6 bond cleavage of **42** is indeed necessary and might require a complex of the all three enzymes. The results also demonstrated the bifunctional (dehydratase/monooxygenases) role of GilOI and GilOIV. GilOIV appeared to catalyze 2,3-dehydration of **40** to **41** in addition to its oxygenation activities namely the hydroxylation at the C-5 position of **41**. Similarly, GilOI was able to catalyze a 4a,12b dehydration of 2,3-dehydrohomo-UWM6 (**41**) in addition to its oxidation activity, which might be the Baeyer-Villiger oxidation of the carbonyl of **42** or 12-hydroxylation. Similar set of metabolites was also identified in the closely related jadomycin pathway through the inactivation of oxygenase homologues (*61*). These inactivation results were further verified by complementation experiments where pWHM1238, a jadomycin PKS expression construct, which is responsible for UWM6 generation, was complemented with various combinations of gilvocarcin/jadomycin-oxygenases. Jadomycin A was only detected when pWHM1238 was co-expressed with JadF, G, H or GilOI, OII, OIV in the presence of L-isoleucine. This was a clear indication for the requirement of all three-oxygenases to form a complex for the oxidative C-C bond cleavage in both pathways (*56*). These results also showed that the biosynthesis of both antibiotics proceed through common intermediates up to **44/44-a**. Based on these results, a pathway for the oxidative cleavage was proposed. GilPKS and other associated enzymes synthesize the presumed intermediate homo-UWM6 (**40**), where GilOIV catalyzes its 2,3-dehydration and then subsequent hydroxylation at C-5 position. Further oxidation of **42** through the activity of the yet unknown enzyme results in the formation of pre-GE-o-quinone (**42-c**). Most likely, GilOI might serve as a Baeyer-villiger monooxygenase which could introduce an oxygen atom between C-5 and C-6 of **42**. Spontaneous hydrolysis of the lactone moiety of **43** could provide the acid-aldeyde intermediate **44,** which could be further oxidized by anthrone oxygenase homologue GilOII.
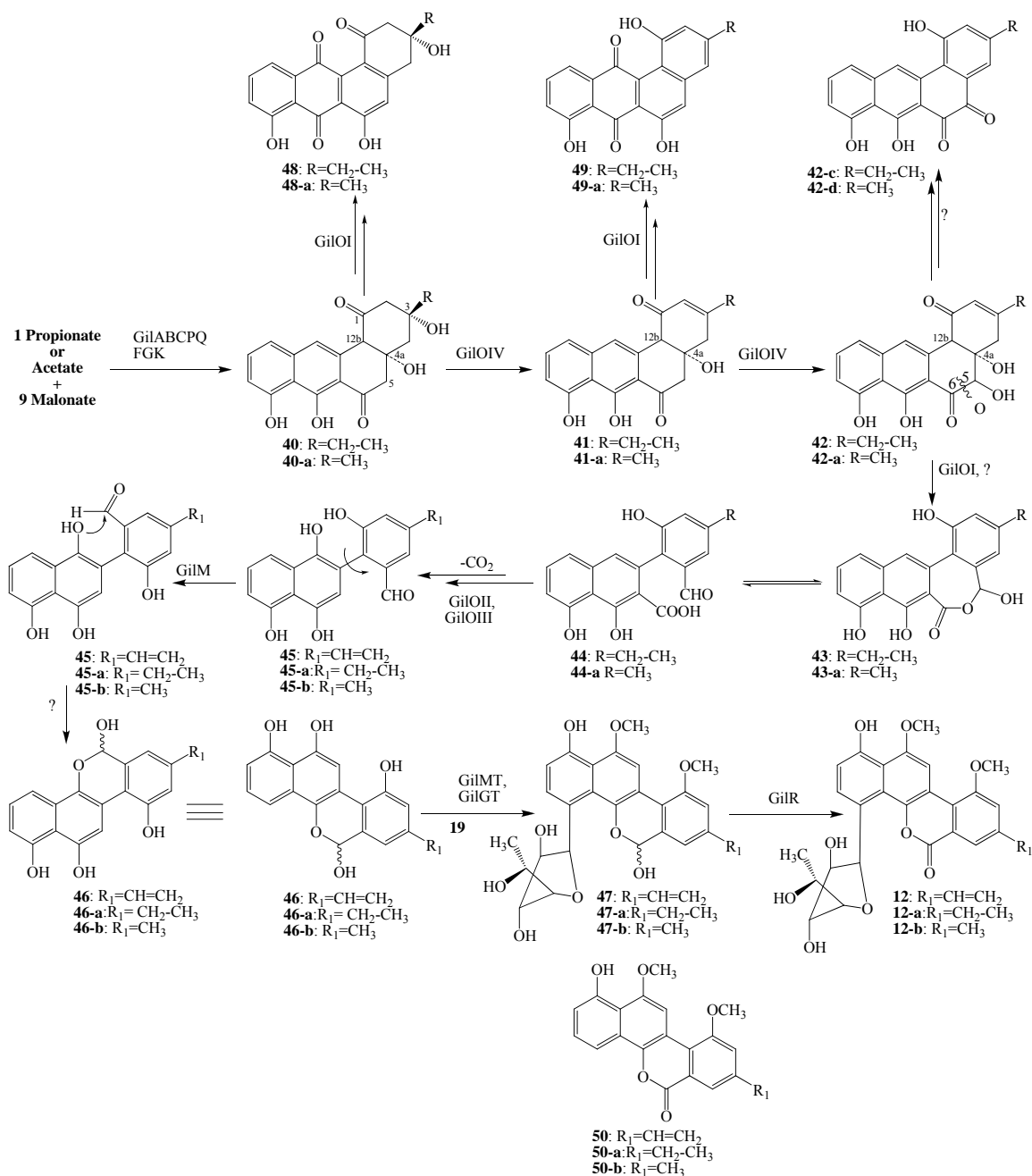
Scheme 4.2 Proposed pathway for the biosynthesis of gilvocarcins

Inactivation of GilOIII and GilGT clearly showed their direct roles in the formation of vinyl group and the C-glycosylation, respectively. Various attempts to produce metabolites from a GilMT deleted-mutant cosmid (cosG9B3-Δ*gilMT*) were not successful. However, the enzyme was proposed to catalyze the methylation of phenolic hydroxyl groups. The ΔGilM mutant abolished the production of GV and accumulated

the biosynthetic shunt product homorabelomycin (**48**) in addition to a yet uncharacterized, highly unstable compound. Inactivation of GilR resulted in the accumulation of the pathway intermediate pre-GV (**47**) indicating its dehydrogenase activity. The functional roles of *gilN*, *gilV* and *gilL* were not assigned as the corresponding gene deleted mutant cosmids failed to express in *S. lividans* TK24.

Several previous *in vivo* studies had indicated that angucycline post-PKS tailoring reactions were often accompanied by follow up bioconversion events in *Streptomyces*. As a result, accumulated products deviated from the pathway toward shunt products with unexpected alterations in the structures (*30, 61, 125, 171-174*). In addition, the involvement of non-enzymatic reactions was proposed as well as unusual bifunctionality combined with an atypical enzymology associated with these biosynthetic steps. These posed challenges to the separation of the real associated biosynthetic events from nonspecific spontaneous or host-contributed endogenous reactions, and also complicated the functional assignment of individual enzymes. Despite having many results from gene inactivation experiments in hand, the GV biosynthetic pathway remained largely unclear. The exact functional roles of GilOI, GilOII and GilOIV appeared to be the central question. Although the results indicated the necessity of these three enzymes for the C5-C6 bond cleavage, these studies failed to address the possible contributions from host enzymes. The results clearly showed the contribution of GilGT, GilM, GilR and GilOIII to GV biosynthesis. However, their substrates and mechanisms of actions were not identified. Similarly, the sequence of the entire post-PKS biosynthetic events remained unclear. Expression of the entire post-PKS enzymes and reconstitution of the pathway *in vitro* would not only provide new insights into the sequence of biosynthetic events and the involved intermediates, but would also provide a means to test the substrate flexibility of the individual enzymes. In addition, the assay system would be free of the the host enzymes, and thus could provide cleaner results which could be applied to engineer the pathway to biosynthesize novel GV derivatives. Since defucoGM (**50-b**) possesses all structural elements of GV except the vinyl group and the sugar moiety, its biosynthesis could be used as a model for *in vitro* studies.

Prejadomycin (**41-a**), the dehydration product of UWM6 (**40-a**), was first isolated from the mutant lacking JadH, an oxygenase of the jadomycin pathway (*61*). Through feeding experiments, it was proven that compound **41-a** was also a jadomycin biosynthetic pathway intermediate (*30*). The compound was later isolated as a minor compound from the ΔGilOI mutant (*57*). Further inactivation and complementation experiments had concluded that **41-a** was likely an intermediate of the GM biosynthetic pathway. However, the hypothesis could not be tested through feeding experiments due to the lack of a GilPKS-deleted mutant strain. Therefore, we have chosen this compound as the starting point to biosynthesize defucoGM through enzymatic synthesis *in vitro*. A detail of the experimental approach and the results are discussed below.

Results

Activity assay of the post-PKS enzymes using 2,3-dehydro-UWM6 (41-a, prejadomycin) as a substrate

As discussed earlier, the complex of GilOI, GilOII and GilOIV might be necessary to catalyze conversion of **41/41-a** to the hypothetical acid-aldehyde intermediate **44/44-a**. To test this hypothesis, all of the three enzymes were expressed in *E. coli*. Unfortunately, GilOIV appeared in the form of inclusion bodies. Various attempts to express the protein in soluble form were unsuccessful. Given the fact that the jadomycin pathway oxygenase (JadF) could complement the GV biosynthesis, which was halted by the deletion of GilOIV (*56*), identical activity of both enzymes was expected. Therefore, JadF was overexpressed in *E. coli* and purified using IMAC. JadF and GilOI were eluted as intense yellow proteins indicating them being flavo-proteins. The molecular weights of the expressed proteins observed in SDS-PAGE were in good agreement with their calculated weights (**Figure 4.1**).
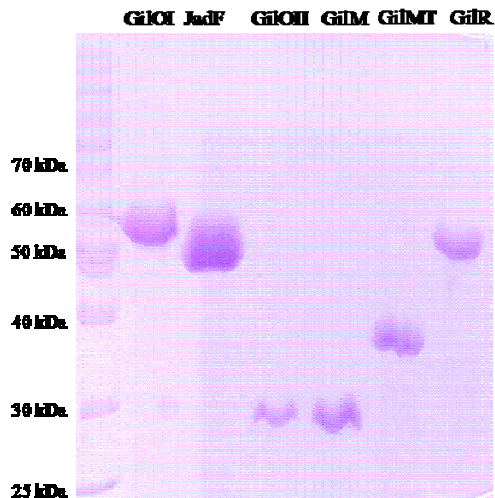
Figure 4.1 Expression and purification of post-PKS tailoring enzymes. Theoretical molecular weights of these enzymes are as follows: GilOI, 55.23 kDa; JadF, 53.72 kDa; GilOII, 27.78 kDa; GilM, 29.35 kDa; GilMT, 39.99 kDa; GilR, 56.15 kDa.

HPLC analyses of cofactor released from GilOI and JadF clearly showed the presence of non-covalently bound FAD. Incubation of prejadomycin (**41-a**), FAD and NADPH with JadF or JadF and GilOII at various conditions did not reveal any change in the HPLC profile of the assay mixture compared to that of the control sample (**Figure 4.2**). To provide excess of FADH *in situ*, NADH, FAD and NADH-FAD reductase were added in the assay mixture, and the assay products were analyzed in parallel. Most of the prejadomycin (**41-a**) was converted into dehydrorabelomycin (**49-a**) when GilOI was added in place of JadF. The reaction involved a C4a-C12b dehydration and an oxygenation at the C-12 position of **41-a**. Dehydrorabelomycin (**49-a**) was the only new compound observed when **41-a** was incubated with GilOI and JadF, GilOI and GilOII, or GilOI, JadF and GilOII (data not shown). The results raised two questions: i) was prejadomycin not a true intermediate? ii) were other post-PKS enzymes (GilM, GilR, GilMT) necessary to catalyze the C5-C6 bond cleavage of **41-a** and then further convert the cleavage product to defucoGM (**50-b**)? To address these issues, the remainder of the defucoGM post-PKS biosynthetic enzymes were cloned, expressed and purified (**Figure 4.1**). Incubation of **41-a** with three oxygenases (GilOI, GilOII and JadF), GilM, GilMT and GilR in the presence of all anticipated cofactors did not reveal any new product, except again the formation of **49-a**. Notably, co-incubation of the assay mixture with

NADH-FAD reductase resulted in poor conversion of prejadomycin (**41-a**) to dehydrorabelomycin (**49-a**) (**Figure 4.2**) indicating the inhibition of the GilOI activity.
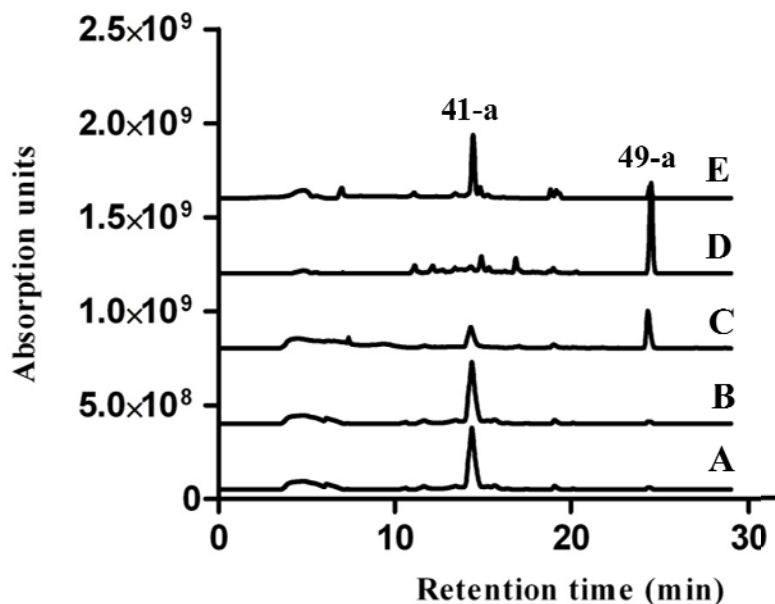
Figure 4.2 HPLC analyses of enzyme assay products using various oxygenases and prejadomycin (2,3-dehydro-UWM6). Trace A: control sample without enzyme; trace B: assay products of prejadomycin and JadF; trace C: assay products of prejadomycin, GilOI and NADH:FAD reductase; trace D: assay products of prejadomycin and GilOI, OII, R, M, MT, JadF; trace E: assay products of prejadomycin and GilOI, OII, R, M, MT, JadF and NADH:FAD reductase.

In contrast to the previous hypothesis, these assay results clearly showed that prejadomycin (**41-a**) was not a defucoGM/GM biosynthetic intermediate. Neither the complex of the three oxygenases alone nor combined with GilR, GilM, GilMT could convert **41-a** to defucoGM (**50-b**) or to other pathway intermediates. The results further indicated that PKS enzymes might be necessary for the optimal activities of the oxygenase complex to proceed the oxidative ring opening, the crucial step of the defucoGM biosynthesis. To envisage this possibility, PKS enzymes were expressed in *Streptomyces lividans* TK64 or in *E. coli* to reconstitute the entire defucoGM pathway *in vitro*.

Expression of PKS enzymes *in vitro*

Gilvocarcin PKS genes identified in the cluster represented typical type II PKS genes. The encoded GilPKS enzymes [GilA (KS_α); GilB (KS_β); GilC (ACP); GilP (AT); GilF (KR); GilK (cyclase); GilG (cyclase)] had been proposed to biosynthesize the anticipated earliest intermediate, UWM6 (**40-a**) of the GM/defucoGM pathway (**Scheme 4.3**). The malonyl-CoA-ACP acyltransferase (GilP) picks up the starting building block (acetyl-CoA) or the extending blocks (malonyl-CoA) and loads then on to the ACP (GilC). Thus, this enzyme along with GilABC constitutes a minimal PKS whose activities would condense nine molecules of malonyl-CoA to an acetyl-CoA to generate a 20-carbon poly-*β*-keto-acyl intermediate, where the associated keto-reductase (GilF) would reduce the keto-group at the C-8 carbon of **77** to a secondary alcohol. The associated cyclase/aromatase (GilK) and cyclase (GilG) could catalyze intramolecular aldol condensations of the poly-*β*-keto-ester which could set up the first, second, the third ring, respectively. The GilG-catalyzed cyclization/dehydration would immediately follow the second cyclization to produce UWM6 (**40-a**).

To reconstitute the biosynthetic pathway of **40-a**, these genes were cloned into pET28a (+) expression vector and thus generated constructs were expressed in *E. coli*. GilA and GilB were expressed together to allow their hetero-dimer formation. Unfortunately, none of the constructs except pGilP and pGilF could produce enzymes in soluble form. GilP and GilF were produced in 8-10 mg/L and 6-9 mg/L culture, respectively. To improve their solubilities, *gilC*, *gilK*, *gilG* and *gilAB* were cloned into pXY200 (*175*) and expressed under the control of a thiostrepton-inducible promoter (*ptipA*) for *Streptomyces* enzymes in a genetically more friendly host *S. lividans* TK64. Surprisingly, 29 mg/L of soluble GilAB could be produced through the expression of the construct pGilAB. Purity and size of proteins were estimated on SDS-PAGE (**figure 4.3**). However, the productions of GilK and GilG were still poor (0.4-0.7 mg/L culture). Several attempts to produce soluble GilC were unsuccessful. Therefore, homologues of these enzymes from closely related pathways were cloned and expressed. Thus, pJadD (contains the homologue of *gilK*) and pRavG (contains the homologue of *gilG*) were generated using jadomycin and ravidomycin biosynthetic pathway genes, respectively.
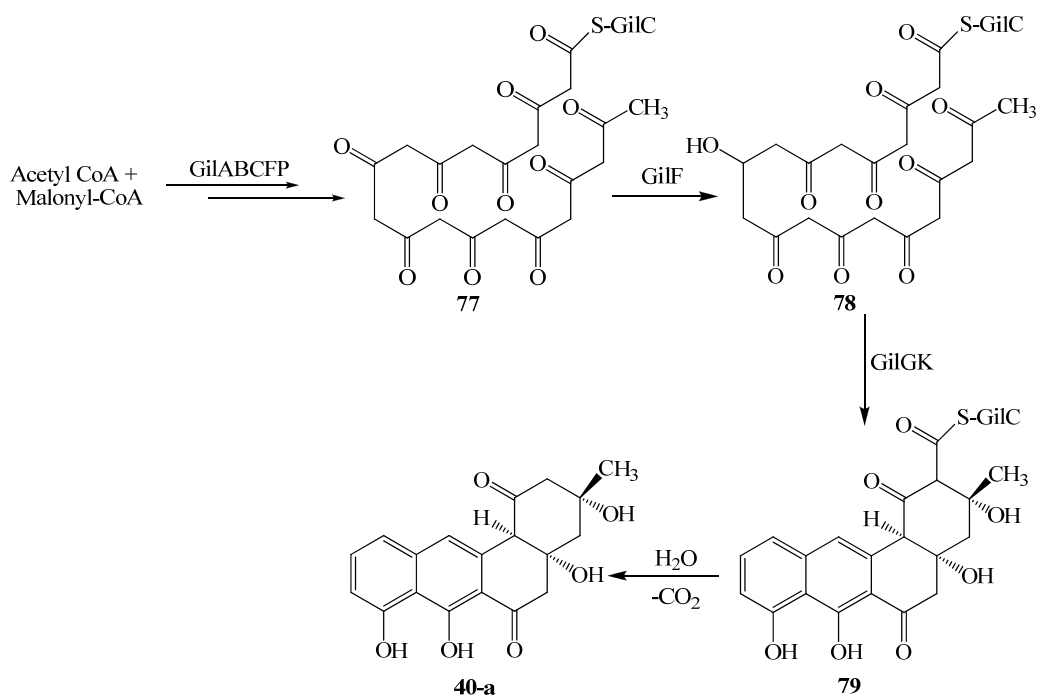
Scheme 4.3 Biosynthesis of UWM6 (**40-a**) through the activity of GilPKS enzymes.
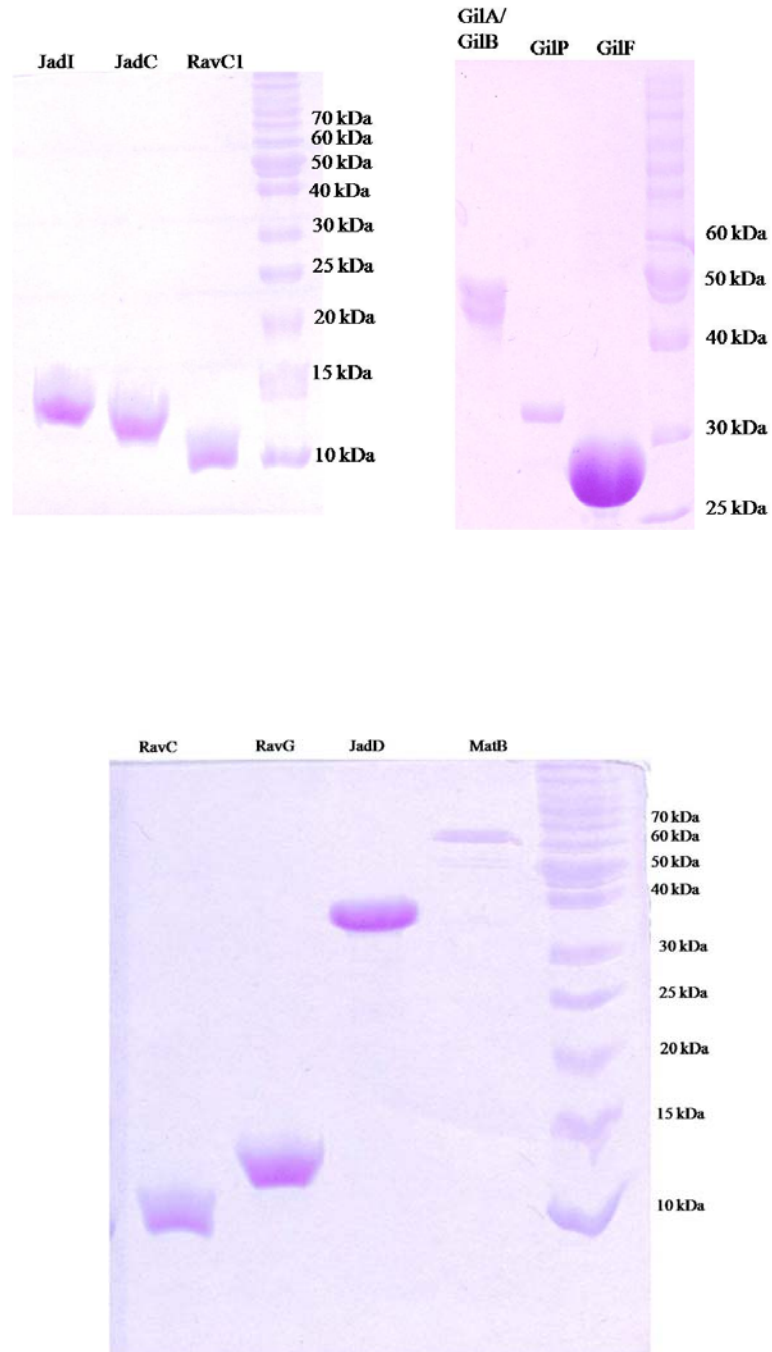
Figure 4.3 SDS-PAGE analyses of PKS enzymes and malonyl-CoA synthetase (MatB). Theoritical molecular weights of the enzymes are as follows: GilA, 43.85 kDa; GilB, 40.33 kDa; GilP, 33.75 kDa; GilF, 26.9 kDa; JadI, 12.4 kDa; JadC, 11.7 kDa; RavC1, 10.8 kDa; RavC; 11.3 kDa; RavG, 12.48 kDa; JadD, 34.74 kDa.

In addition, *jadI* (another homologue of GilG) from jadomycin gene cluster was also cloned and expressed. As discussed earlier in Section 2, there were two putative ACPs (RavC and RavC1) in the RMV cluster, unlike GilC and JadC of the GV and jadomycin pathways, correspondingly. Considering the potential role of either or both of these RavACPs in defucoGM biosynthesis, both were cloned into pET28a vector to generate the expression constructs pRavC and pRavC1. In addition, JadC expression construct (pJadC) was also generated and expressed in *E. coli*. While RavG (10 mg L^{-1} culture) and JadD (9 mg L^{-1} culture) were produced from *E. coli* BL21 (DE3), 5-9 mg L^{-1} of each of the ACPs were also produced from the same host. The sizes of proteins observed in SDS-PAGE were relevant to the calculated molecular weight of the proteins (**Figure 4.3**).

It was well established that post-translational modification of apo-ACP was crucial for the activities. Such modification involved the phosphopantetheinylation of the active site serine residue of the apo-ACP. The reaction was catalyzed by phosphopantetheinyl transfearase (PPTase) which transfers phosphopantetheine moiety of co-enzyme A (CoA) to the hydroxyl group of a serine residue of the *apo*-ACP to generate the functionally active *holo*-ACP (**Scheme 4.4**). The phosphopantetheine arm of the ACPs provides an anchor site to tether starter/extender units. Since *E. coli* harbored a functional PPTase to generate *holo*-ACPs for its fatty acid biosynthesis, it was possible that RavC, RavC1 and JadC were expressed in their *holo* forms, but that was not sure. Therefore, each enzyme was purified, desalted, and then subjected to the MALDI-TOF analysis.
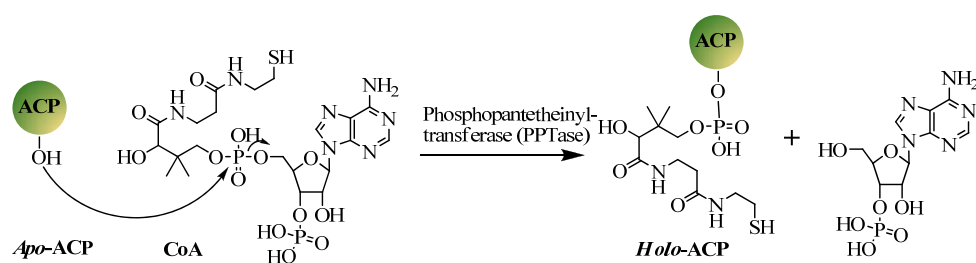
Scheme 4.4. Conversion of *apo*-ACP to *holo*-ACP.

Table 4.1 MALDI-TOF analyses of ACPs

Enzymes	*Apo*-ACP (Mol Wt.)		*Holo*-ACP (Mol. Wt.)		*Holo*-ACP-Acetyl-CoA (Mol. Wt.)	
	Calculated	**Observed**	**Calculated**	**Observed**	**Calculated**	**Observed**
RavC	11197	11203	11537	11550	11585	11593
RavC1	10732	10735	11072	11085	11115	11128
JadC	11613	11619	11953	11964	11996	12007

Only molecular weights of the *apo*-ACPs were observed in the spectra with reasonable accuracy (± 10 amu) ruling out the phosphopantetheinylation of RavC, RavC1 and JadC through the activities of the host enzymes (**Table 4.1**). Thus, to produce these enzymes in their *holo*-forms, the genes were expressed in the *E. coli* BAP1 host (*176*), which was generated through the integration of the promiscuous phosphopantetheinyl transferase (*sfp*) gene from *Bacillus subtilis* into the genome of *E. coli* BL21(DE3), thereby generating a mutant strain capable of efficiently expressing ACPs in their *holo* form. The amount of production of ACPs from *E. coli* BAP1 host was similar to that from *E. coli* BL21 host. MALDI-TOF analyses of the purified proteins revealed the presence of both *apo*- and *holo*-forms in about equal proportions. (**Table 4.1**).

Activity assay of PKS enzymes: an enzymatic synthesis of angucyclines

With all of the anticipated PKS enzymes in hand, enzyme assays were conducted in analytical scale and the assay products were analyzed by HPLC-MS. Malonyl-CoA was synthesized *in situ* using sodium malonate, CoA and malonyl-CoA synthetase (MatB) (see details in materials and methods). Preliminary screening for the efficient production of polyketides was conducted using a "mix and match" approach. The experimental design using various combinations of ACPs (RavC, RavC1 and JadC), CYCs (RavG, JadD) and other PKS enzymes (GilF, GilP, GilAB) are summarized in **Table 4.2**.

Table 4.2 An experimental design for the preliminary PKS enzyme assay

Enzymes	Experiment no.							
	1	2	3	4	5	6	7	8
GilA/GilB	+	+	+	+	+	+	+	-
RavC	+	+	-	-	+	+	+	+
RavC1	+	-	+	-	+	+	+	+
JadC	-	-	-	+	-	-	-	-
GilF	+	+	+	+	-	+	+	+
GilP	+	+	+	+	+	+	+	+
RavG	+	+	+	+	+	-	+	+
JadD	+	+	+	+	+	+	-	+

+ and – indicates presence or absence of enzymes, respectively.

A cocktail of acetyl-CoA, malonyl-CoA and NADPH in phosphate buffer was incubated with eight different combinations of PKS enzymes in 200 μL solution. Rabelomycin (**48-a**) was produced as the major compound from assay no. 1 through 4 (**Figure 4.4**). The authenticity of the compound was confirmed through the comparison of retention time, UV spectrum and mass data with an authentic sample compound. This clearly demonstrated that GilABFP, JadD, RavG and either of the ACPs (JadC, RavC or RavC1) were sufficient to biosynthesize rabelomycin from acetate and malonate building blocks. Surprisingly, UWM6 (**40-a**) was not observed in any of these experiments while **48-a** was produced as the major product. This could be due to the instability of **40-a** under the experimental conditions, where it could convert to **48-a** spontaneously. Such a non-enzymatic conversion was reported previously while producing UWM6 in *S. lividans* TK24 using jadomycin PKS biosynthetic enzymes (*177*). Since the reaction was carried out in a small scale, UWM6 could be present in the assay mixture in undetectable small amounts. The new peaks **a**, **b** and **c** which were detected in assay no. 5 and 6 were not identified. However, their UV spectra were not typical to those of angucyclines. This indicated that **a**, **b** and **c** could be aroused from the partial cyclization of the newly formed poly-*β*-keto-ester. No new peak was detected from experiment 7 that lacked

cyclase JadD. This could be due to the spontaneous cyclization of the newly formed polyketide to several products not sufficient enough to appear in the HPLC chromatogram. As expected, no new compounds were detected in experiment 8 (the control experiment that lacked GilAB). In summary, assay no. 2 and 4 appeared to be the best enzyme combinations for the production of rabelomycin (**48-a**). Therefore, GilAB, RavC, GilF, GilP, RavG and JadD (assay no. 2) were chosen for the production of angucyclines in larger scale which potentially could synthesize UWM6 besides rabelomycin, and for the coupled assay with the post-PKS enzymes to biosynthesize defuco-GM.

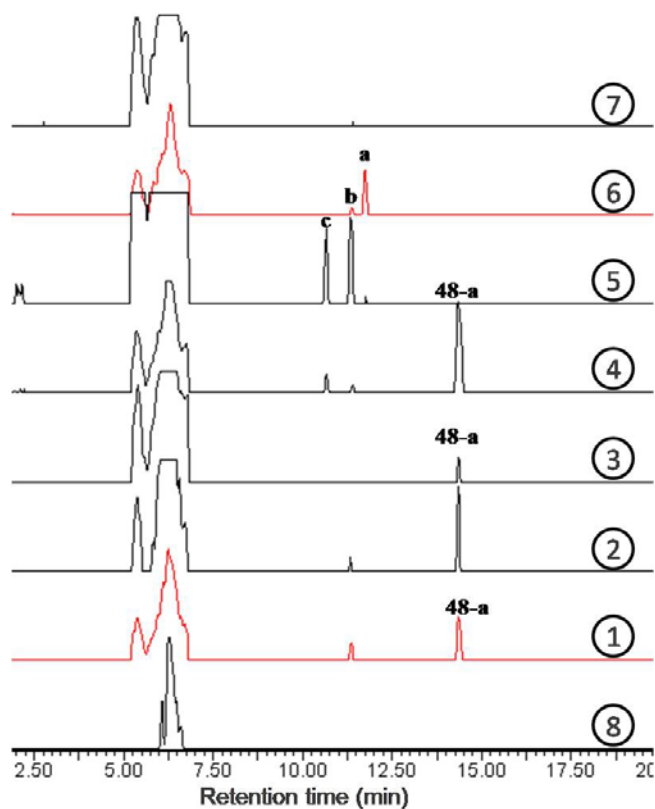**Figure 4.4** Analyses of the PKS enzyme assay products. Traces 1 through 8 correspond to the experiment 1 through 8 described in **table 4.2**. Compounds were detected at 380 nm. The peaks **a-c** were not identified. Compound **48-a** represents rabelomycin.

The assay volume was increased to 10 mL and the experiment was conducted under identical conditions. HPLC-MS analyses of the assay products revealed **48-a** as the

major product (**Figure 4.5**). Compound **40-a** was not observed in any of the experimens. However, a new peak (peak **X**, **Figure 4.5**) with the UV spectrum typical of **40-a** was detected. The compound was purified through HPLC. Low resolution ESI-MS revealed a molecular ion [M-H]$^-$ peak of 383 daltons at negative mode suggesting a MW of 384. A molecular ion peak of 383.1113 daltons corresponding to $C_{20}H_{16}O_8$ observed in high resolution ESI-MS (at negative mode) further confirmed the MW of the compound to be 384 daltons. However, we were unable to further characterize this compound through NMR due to its poor yield.
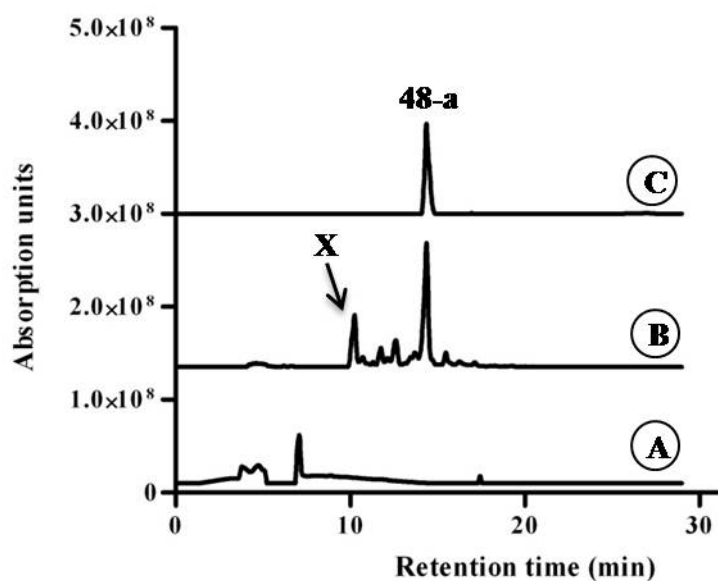
Figure 4.5 Enzymatic synthesis of rabelomycin and other angucyclinones. Trace A: control sample without enzymes; trace B: GilABFP, RavCG and JadD incubated with acetyl-CoA, malonyl-CoA and NADPH; trace C: standard rabelomycin (**48-a**). Compound X was the unknown compound with MW=384 g Mol^{-1}.

Enzymatic synthesis of defuco-GM

Production of the biosynthetic shunt product rabelomycin as the major product through the activity of the PKS enzymes suggested that the PKS and post-PKS oxygenase complex might function intimately to catalyze the oxidative ring opening of a *in situ* formed angucycline intermediate, e.g. **40-a**. The product of the oxygenase complex could

be further processed by other enzymes, presumably GilM, GilMT and GilR, to yield defuco-GM. Successful reconstitution of the entire defuco-GM pathway would be highly interesting as it would outline the minimum number of enzymes required to biosynthesize this compound. Identification of the compounds generated from various combinations of PKS and post-PKS enzymes would also provide a platform to identify the pathway intermediates as well as the sequence of biosynthetic events. Therefore, we attempted the co-incubation of PKS and post-PKS enzymes and analyzed the metabolites in HPLC-MS.
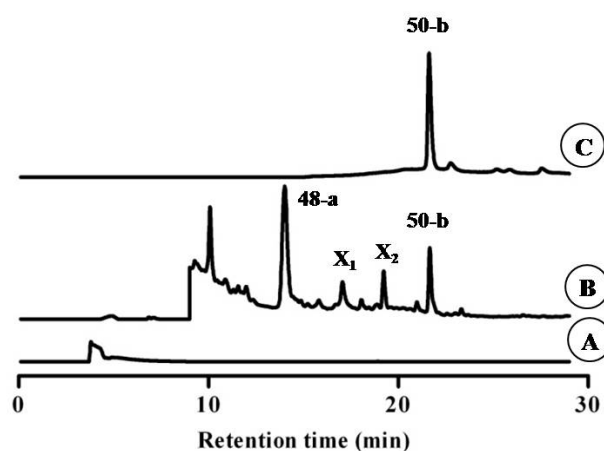

Figure 4.6 HPLC analyses of the enzyme assay products: trace A, control sample without GilAB; trace B, assay sample with GilA, B, F, P, OI, OII, R, M, MT, RavC, G, JadD and JadF; trace C, standard defucoGM. X_1 and X_2 represent yet uncharacterized compounds. They were also observed in the culture broth of the ΔGilGT mutant.

Acetyl-CoA and malonyl-CoA were incubated with GilA, GilB, GilF, GilP, GilOI, GilOII, GilM, GilMT, RavC, RavG, JadD, JadF, NADPH, FAD and SAM. Here, NADPH could serve as a co-factor for the PKS-associated ketoreductase GilF, and for the FAD-dependent oxygenases GilOI and JadF. GilF could utilize NADPH directly whereas the oxygenases would utilize it to generate chemically reactive FADH *in situ*. Similarly, SAM served as a methyl donor for the MTs (GilMT, GilM). Overnight incubation of the assay mixtures at 37 °C resulted in the formation of several new compounds besides rabelomycin. Formation of defuco-GM in the assay mixture was confirmed through the comparison of retention time, UV spectrum and ESI-MS with the standard compound (**Figure 4.6**). UV spectra and retention times of the other two unknown minor products

were also observed in culture extract of the defuco-GM producer ΔGilGT mutant (*33*). This further indicated that defuco-GM production *in vitro* indeed had an identical biosynthetic fingerprint to that of the *in vivo* production. As expected, no defucoGM or rabelomycin were observed when GilAB were removed from the assay mixture.

Discussion

Polyketide-derived metabolites show tremendous structural diversity. A majority of their structural and functional variations are aroused from post-PKS tailoring reactions. Biosynthetic pathway of most of the bio-active compounds such as GV (*71*), jadomycins (*61*), urdamycins (*106*), landomycins (*178*), kinamycin D (*179*) and marmycins (*178*) utilize identical building blocks and functionally identical PKS enzymes to generate an angucycline backbone, presumably UWM6, which is further processed through the activity of specialized post-PKS tailoring enzymes to furnish their final structures. Oxygenases and dehydratases have appeared to be the earliest post-PKS modifying enzymes in many biosynthetic pathways of angucycline-derived natural products (*30, 59, 171*). Oxygenases have drawn special interest in biosynthetic studies due to their tremendous capacities to alter structures through hydroxylation, C-C bond cleavage, rearrangement of the backbone and desaturation that ultimately lead to the generation of highly complex biologically active compounds, whose biosyntheses are often difficult to decipher without the study of the involved enzymes. For instance, the oxygenation cascade of jadomycin biosynthetic pathway has been labeled as a "biosynthetic black box" as *in vivo* and *in vitro* experiments attempted so far failed to address the exact functional roles of the oxygenases JadFGH (*61*). The biosyntheses of GV and GM also involve such modifications, for instance, C-C bond cleavage, a series of oxygenations, a glycosylation and methylations. Previous experiments concluded that GilOI, GilOII and GilOIV were the earliest post-PKS tailoring enzymes that could convert the anticipated intermediate UWM6 (**40-a**) and homo-UWM6 (**40**) to GM (**12-b**) and GV (**12**), respectively, through the presumed intermediates prejadomycin (**41-a**) and 2,3-dehydro-homo-UWM6 (**41**), respectively. However, the exact catalytic actions of these enzymes, the sequence of their actions and the fate of the involved intermediates were not clear. Therefore, the first task of this study was to demonstrate **41/41-a** to be the intermediates

of the GV/GM pathway. There were three advantages in focusing on GM instead of on GV. First, the pathway did not require a desaturation step to install the vinyl side chain, which makes the study simpler. Second, prejadomycin could be obtained easily in large amounts from a landomycin oxygenase (LndE)-deletion mutant. Lastly, the jadomycin pathway oxygenases JadF and JadH could be used in place of GilOIV and GilOI, respectively, whenever necessary. Here, it was important to note that JadF and JadH completed GV biosynthesis that was disrupted by the deletion of GilOIV and GilOI, respectively, indicating their functional identities.

To unravel the molecular details of the GV oxygenation cascade, prejadomycin (**41-a**) was incubated with various combinations of oxygenases. Surprisingly, the substrate remained unchanged in all cases except those involving GilOI, where the compound was converted into the shunt product 2,3-dehydrorabelomycin (**49-a**). The results clearly outlined two possibilities: the first, **41-a** might not be an appropriate substrate of GilOIV or GilOI; the second, the downstream post-PKS enzymes might be necessary for the oxygenases to carry out their optimal functions. The product **49-a** was originally produced *in vivo* from the ΔJadG mutant from the jadomycin pathway, and later from the ΔGilOII mutant as well (*30, 56, 61*). Identical transformation was observed *in vitro* when the GilOI homologue JadH was incubated with prejadomycin (*166*). JadH and GilOI had been proposed to be bifunctional dehydratase/oxygenase enzymes where they could catalyze dehydration to establish a double bond between carbons 4a and 12b (*101*). It was not clear whether C-12 oxygenation could happen spontaneously after 4a,12b dehydration, or whether JadH might contribute to or is responsible for this oxygenation. The observed activity of GilOI was identical to that of JadH and it did not reveal any new insights into the exact functional role of this enzyme in the context of the GV/GM biosynthesis. The FAD-dependent monooxygenases could be one component or two component enzymes. One component enzymes can produce $FADH_2$ themselves utilizing NADPH whereas two component oxygenases require the second enzyme that utilizes NADH or NADPH to generate $FADH_2$ to carry out their activities. In this context, GilOI and JadF could be two component enzymes and $FADH_2$ might not be available to carry out their oxygenation activities. To provide access of $FADH_2$, NADH, FAD and NADH:FAD reductase were supplied to the assay mixtures. Such extra supply

of FADH$_2$ *in situ* did not alter the product profiles except the inhibition of conversion of **41-a** to **49-a**. It is possible that the robust production of FADH$_2$ could induce conformational change in GilOI which could prevent the docking of **41-a** to the active site, thereby preventing the 4a,12b dehydration as well as C-12 oxygenation or NADH:FAD reductase might block the substrate from docking into the active site of GilOI. The results again indicated that prejadomycin was not an appropriate substrate for GilOI and JadF or that the oxygenases might need to form a complex with the other post-PKS enzymes to convert **41-a** to defucoGM (**50-b**). To test the second possibility, **41-a** was incubated with GilOI, GilOII, JadF, GilM, GilMT and GilR in the presence of cofactors SAM, FAD, NADH/NADPH and NADH:FAD reductase. As a result, the substrate remained unchanged. This result revealed that **41-a** was not a defuco-GM pathway intermediate, and that PKS and post-PKS enzymes might form a giant multi-enzyme complex, where the nascent yet unknown PKS product could be utilized immediately by the post-PKS enzymes to convert to defucoGM. Alternatively, these enzymes might not be forming a complex but need to be physically close enough to transfer the nascent PKS product to the post-PKS enzymes, as hypothesized earlier for the gaudimycin biosynthetic pathway (*59*).

To reconstitute the entire defuco-GM pathway *in vitro*, we had further cloned and expressed PKS enzymes from GV, RMV and jadomycin pathways. Preliminary screening for the production of UWM6 was conducted using a "mix and match" approach. Interestingly, rabelomycin (**48-a)** was produced as the major product when GilA, B, F, P, JadD and RavG were incubated with any of the three ACPs (JadC or RavC or RavC1). This clearly demonstrated that the three ACPs are functionally identical. It was an interesting result for several reasons: i) it was the first report on the successful reconstitution of an angucycline pathway *in vitro*; ii) rabelomycin was produced through a one-pot enzymatic reaction; iii) the assay demonstrated that PKS enzymes (GilABFP, RavC, RavG and JadD) were sufficient to generate rabelomycin, presumably through the spontaneous oxidation of UWM6 after cleaving off from the ACP (RavC); iv) GilA and GilB could accept ACPs from other angucycline pathways without discrimination; v) it indicated that PKS enzymes might require post-PKS enzymes to channel the PKS-product into the defuco-GM biosynthetic pathway thus preventing it from converting to

shunt product rabelomycin. Interestingly, incubation of the PKS and post-PKS (GilOI, GilOII, JadF, GilM, GilMT, GilR) enzymes with acetyl-CoA, malonyl-CoA (generated *in situ*), FAD, NADPH, SAM, NADH and NADH:FAD reductase resulted in the formation of defucoGM (**50-b**) and rabelomycin (**48-a**) in addition to other minor yet uncharacterized compounds. Several conclusions could be drawn from this results: i) the enzymes used in this assay were sufficient for the production of defuco-GM; ii) post-PKS tailoring enzymes must be coupled with PKS enzymes to convert the nascent PKS product to defuco-GM; iii) several other unknown gene products (GilN, GilL and GilV) of the GV pathway did not contribute to the biosynthesis of defuco-GM (**50-b**); iv) the imperfect stochiometry of the PKS and post-PKS enzymes might result in the production of rabelomycin (**48-a**) where the post-PKS enzymes might not be able to sufficiently utilize the products of the PKS enzymes leading partly to a spontaneous decomposition of the yet uncharacterized intermediate.

Overall *in vitro* results were in agreement with the current hypothesis that post-PKS tailoring reactions often require a close synchronization between steps to prevent intermediate degradation or shunt product formation. The post-PKS oxygenation cascade catalyzed by the pathway oxygenase (PgaE) and the bifunctional oxygenase/dehydratase (PgaM) in the gaudimycin C biosynthetic pathway represents another typical example (*59*). Through *in vitro* assays, it was demonstrated that both enzymes are required to be in close vicinity, not in the form of a non-covalent complex, to convert UWM6 to gaudimycin C through a oxygenation/dehydration cascade. However, such a phenomenon can not be responsible in the GV post-PKS oxygenation cascade, as both the *in vitro* reconstitution of the post-PKS oxygenases and the entire post-PKS enzymes failed to convert the hypothesized intermediate prejadomycin (**41-a**) to any compound except to the shunt product 2,3-dehydrarabelomycin (**49-a**). In light of an emerging new class of oxygenases that work on ACP-tethered substrates, combined with the results of this work, which showed the formation of defuco-GM only after reconstitution of the entire PKS and most of the post-PKS enzymes, we suggest to revise the GV pathway as described in **Scheme 4.4.** According to this, an ACP-tethered angucycline, presumably **79-a/79**, might serve as the true substrate of JadF/GilOIV. Similar oxygenation activitiy has been discovered recently in the biosynthetic pathway of the antitumor antibiotic C-1027, where

the two-component monooxygenases SgcC catalyzes hydroxylation of the *β*-tyrosine moiety covalently attached to the peptidyl carrier protein (SgcC2) (*180*). GilOIV/JadF might also catalyze C2-C3 dehydration as well as C-5 hydroxylation of **79/79-a** to generate the **80/80-a** which could undergo GilOI-catalyzed Baeyer-Villiger oxidation to generate a lactone-type intermediate **81/81-a**. Opening of the 7-membered lactone ring might take place before or after hydrolysis from the ACP and subsequent decarboxylation. GilOII might catalyze the hydroxylation at the napthol moiety of **44/44-a**. GilM or GilMT might methylate one of the two p-hydroxyl groups which would protect the compound from spontaneously converting to the quinone, as assumed for jadomycin biosynthesis, in whose gene cluster no GilM homologue is found. Desaturation of C-8 ethyl side chain could be catalyzed by GilOIII, along the way, although the exact substrate for this enzyme remains elusive. The hemiacetal moiety of **82**, **82-a** and **82-b** might form spontaneously through the rotation of the C-C bond and the high reactivity of the aldehyde function. Methylation of the other hydroxyl group followed by dehydrogenation of the hemiacetal would generate the defucogilvocarcins (**50**, **50-a**, **50-b**), however, glycosylations might precede these events to yield gilvocarcins. Further *in vitro* activity assays would be necessary to validate this hypothetical pathway. In this context, co-incubation of individual or combinations of post-PKS enzymes with PKS enzymes, and the identification of the products of these assays would be the most crucial experiments.
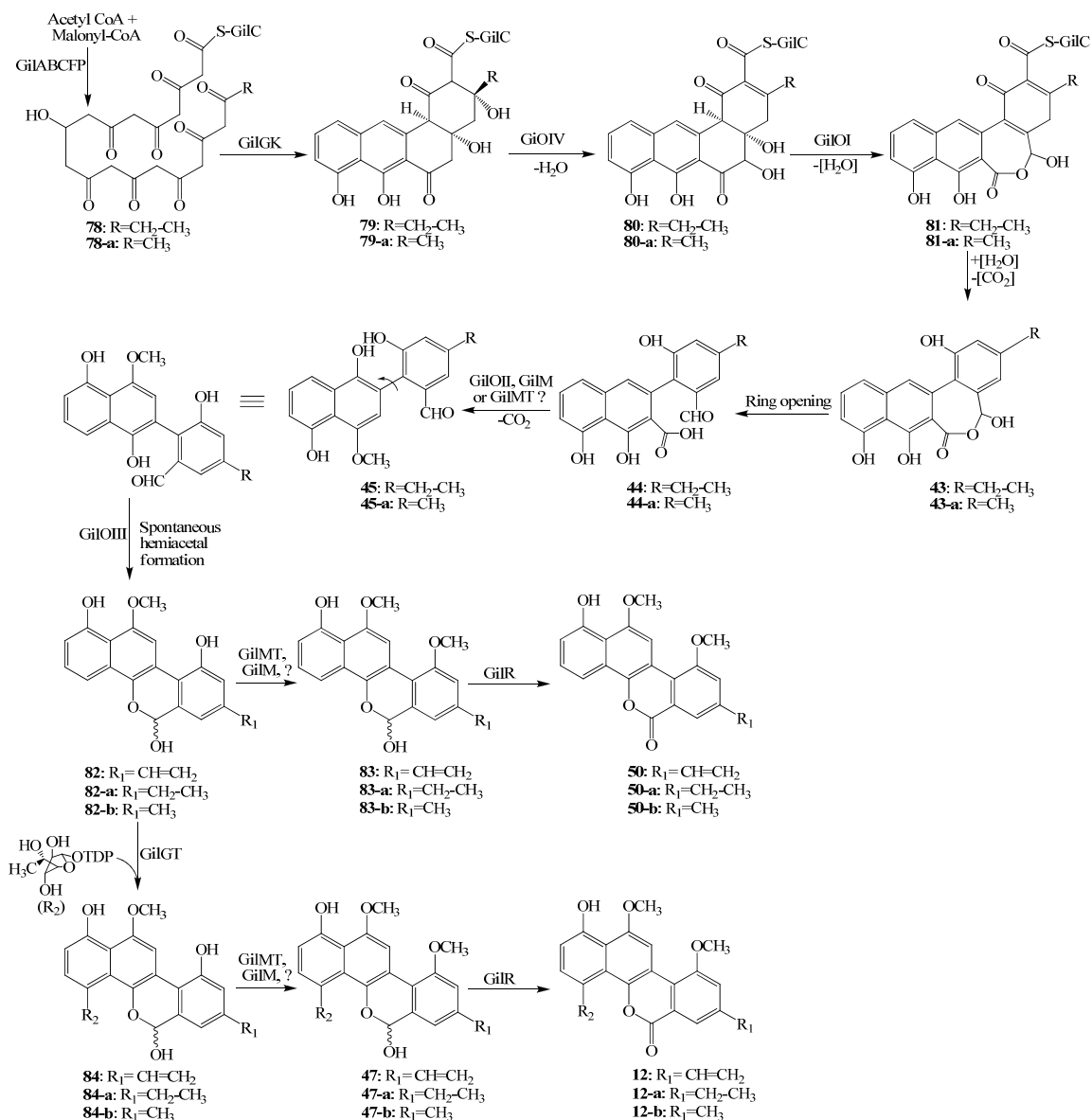
Scheme 4.5 A revised hypothetical pathway for GV biosynthesis. Note that the key-oxidative events occur, while the oxygenase substrates are still tethered to the ACP GilC

Rabelomycin was first isolated from *Streptomyces olivaceous* ATCC21549 (*181*). The compound was later isolated from oxygenase deficient mutants from the jadomycin pathway (ΔJadF), urdamycin (ΔUrdM) and gilvocarcin (ΔGilOIV) biosynthetic pathways (*56, 57, 61, 180, 182*). In addition, the compound was also produced *in vivo* through the engineering of PKS enzymes or through the expression of a cryptic pathway (*125, 174, 177*). Rabelomycin was recognized as an antibiotic, it displayed an activity comparable to tetrangomycin against gram-positive bacteria. Several synthetic routes to produce

101

rabelomycin and related compounds had been developed over the past decades (*183-187*). Similarly, several chemical synthesis routes for the preparation of defucoGM (*187-190*) have been developed. However, all of these methods involve multistep syntheses and suffer from poor overall yields. Enzymatic total syntheses of defucoGM and rabelomycin presented in this work provide alternative approaches, which are environmentally friendlier, more efficient and cost-effective than any of the previously reported synthetic approaches. In addition, the enzymatic synthesis of defucoGM, to our knowledge, represents the first successful one-pot enzymatic synthesis of a natural product that involves consecutive activities of more than 12 functionally distinct enzymes.

Lactone generation by GilR as the last step of the GV biosynthetic pathway

GilR was one of the unknown genes of the GV cluster. Homology searches showed that the enzyme GilR (498 aa) belongs to an oxidoreductase family enzyme, whose amino acid sequence displays its identity/similarity to various putative dehydrogenases from various species: 43%/57% (identity/similarity) to StfE (CAJ42334) from *Streptomyces steffisburgensis*, 41%/55% to a dehydrogenase (ZP_00377889) from *Brevibacterium linens* BL2 and 41%/58% to BusJ and SpnJ from butenyl spinosyn and spinosyn producer *Saccharopolyspora pogona* and *Saccharopolyspora spinosa*, respectively. The conserved domain analysis revealed two hits: the N-terminal FAD-binding domain (pfam01565) and the *C*-terminal berberine-like domain (pfam08031). Pfam01565 family proteins belong to plant enzymes that utilize FAD as a co-factor and catalyze oxidation reactions, whereas pfam08031 family enzymes oxidize the *N*-methyl group of *S*-reticuline into the C-8 methylene bridge carbon of (S)-scoulerine in presence of FAD.
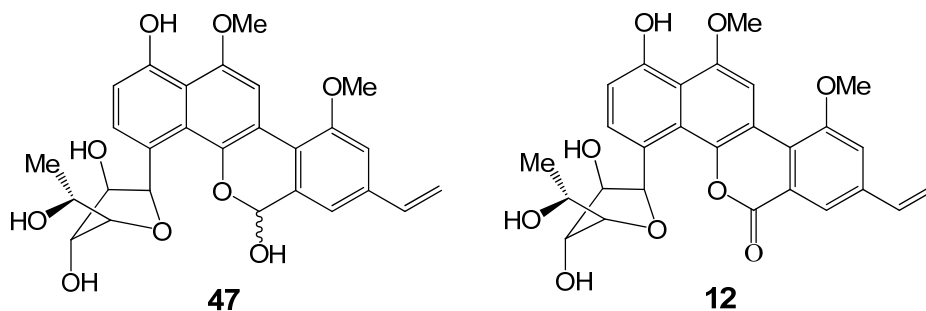
To deduce its functional role in the GV biosynthesis, the gene was inactivated and the ΔGilR mutant accumulated preGV (**47**) which contains a hemiacetal moiety instead of the lactone moiety of GV (*68*). PreGV was completely converted into the GV when fed in the culture of early pathway enzyme-blocked mutant thus, proving the former compound an intermediate of the GV pathway. Chemically, the conversion requires elimination of two protons along with two electrons from preGV. Although many dehydrogenases that catalyze such reactions were reported in primary as well as secondary metabolism, none of them were obseverd in the context of hemiacetal oxidation. To deduce the exact function of the *gilR*-product, the gene was cloned into the pET28a vector, expressed in *E. coli*, and the generated *N*-terminal His-tagged protein was purified using IMAC (**figure 4.7**). The intense yellow color of the purified enzyme and its typical FAD-like UV spectrum clearly indicatied the presence of bound FAD (**Figure 4.8**).
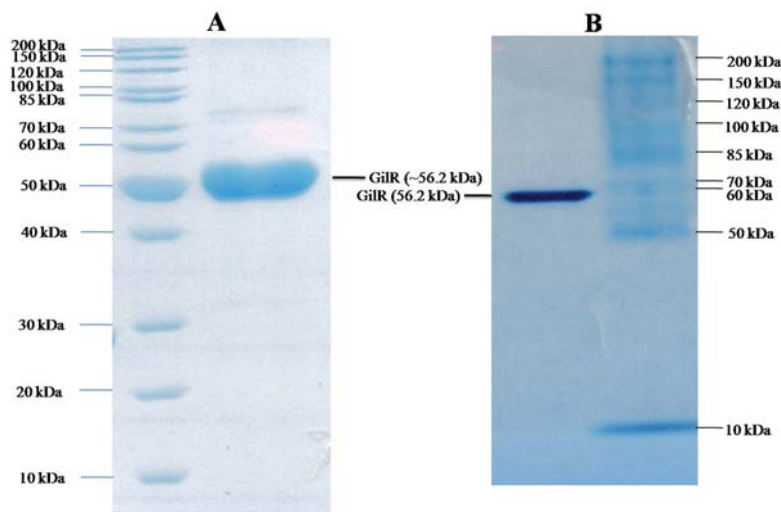

Figure 4.7 Separation of purified GilR at denaturating conditions (SDS-PAGE, A) and at native conditions (native gradient PAGE, B)

To figure out whether FAD is covalently bound or not, the enzyme was denatured using different techniques and the protein-free supernatant liquid was analyzed (see materials and methods for the detail). UV analyses of the liquid did not reveal the presence of FAD indicating a covalently bound FAD. The results were further supported by the visualization of intense golden yellow protein pellets. To determine the quaternary structure of GilR, the enzyme was separated in native 15% polyacrylamide gel (Bio-Rad).

A single band was observed in between 50 and 60 kDa. A similar size was observed when the protein was separated in 12% SDS-PAGE (**Figure 4.7**). This was a clear indication for the monomeric existence of GilR in its natural form.
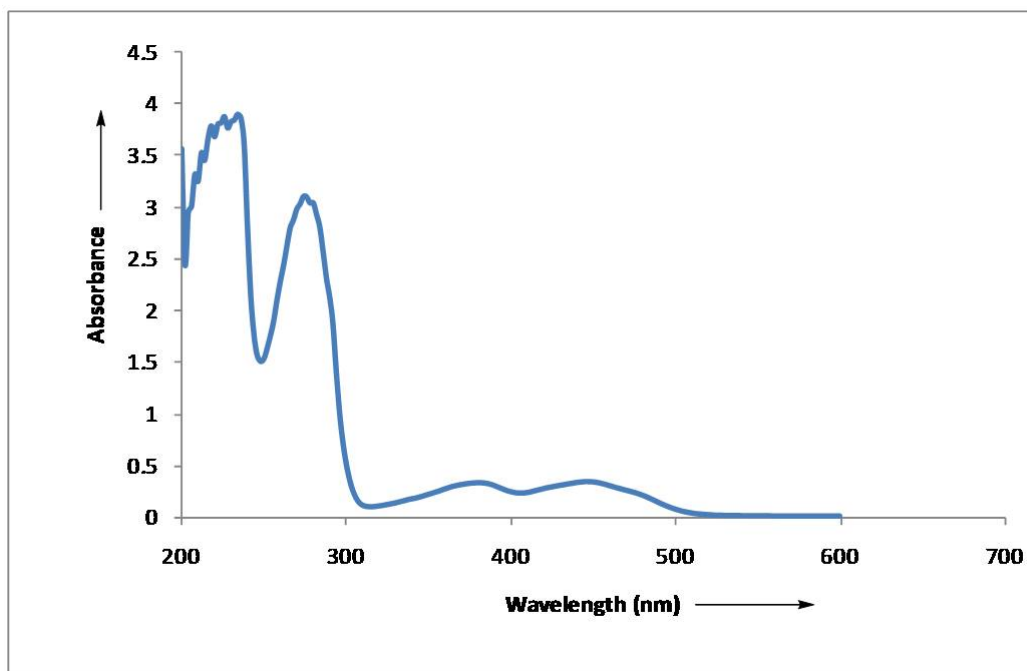

Figure 4.8 UV spectrum of the purified GilR (200 μL, 4.0 mg mL^{-1} concentration). Desalting buffer was used as a reference.

Activity assay of GilR

The pathway intermediate **47** was incubated with GilR both in presence or absence of FAD. The HPLC/MS analysis of the assay mixture revealed a dramatic decrease of the substrate with concomitant formation of **12** (**Figure 4.9**, traces A and B) in the both cases. The addition of FAD did not alter the GilR catalysis indicating the saturation of GilR with FAD. The results confirmed that GilR catalyzes the oxidation of the hemiacetal moiety of **47** to the lactone found in **12**. The accumulation of significant amount of defuco-GV (**50**) by both the GilGT (*C*-glycosyltransferase)- and GilU (deoxysugar ketoreductase)-deletion mutants indicated that GilR might be capable of oxidizing defuco-PreGV (**83**) to **50**, and the tethered defucofuranose residue might not be required for its activity (*102*). The accumulation of **47** by the GilR-deletion mutant

suggested the occurrence of the *C*-glycosylation prior to the lactone formation. These results raised a question, which one (**47** or **83**) is the true substrate of GilR? To address this ambiguity, and to address the sequence of biosynthetic events, **83** was prepared semi-synthetically through one-step diisobutyl aluminium hydride (DIBALH) mediated reduction of **50**, and the kinetics of GilR catalysis were compared for both substrates. Interestingly, GilR was capable to oxidize **83** to **50**. However, the reaction rate was much slower as compared to the reaction with substrate **47** (**Figure 4.9**, trace C, D and E). The K_m values of GilR were determined to be 117.0 μM and 93.4 μM for **47** and **83**, respectively (**Table 4.3**, **Figure 4.10**, **4.11**). However, the k_{cat} for **47** was found to be more than four fold higher than for **83.** The ratio of k_{cat} and K_m for **47** was 3.5 fold higher than for **83** clarifying the former compound to be the true substrate of GilR.
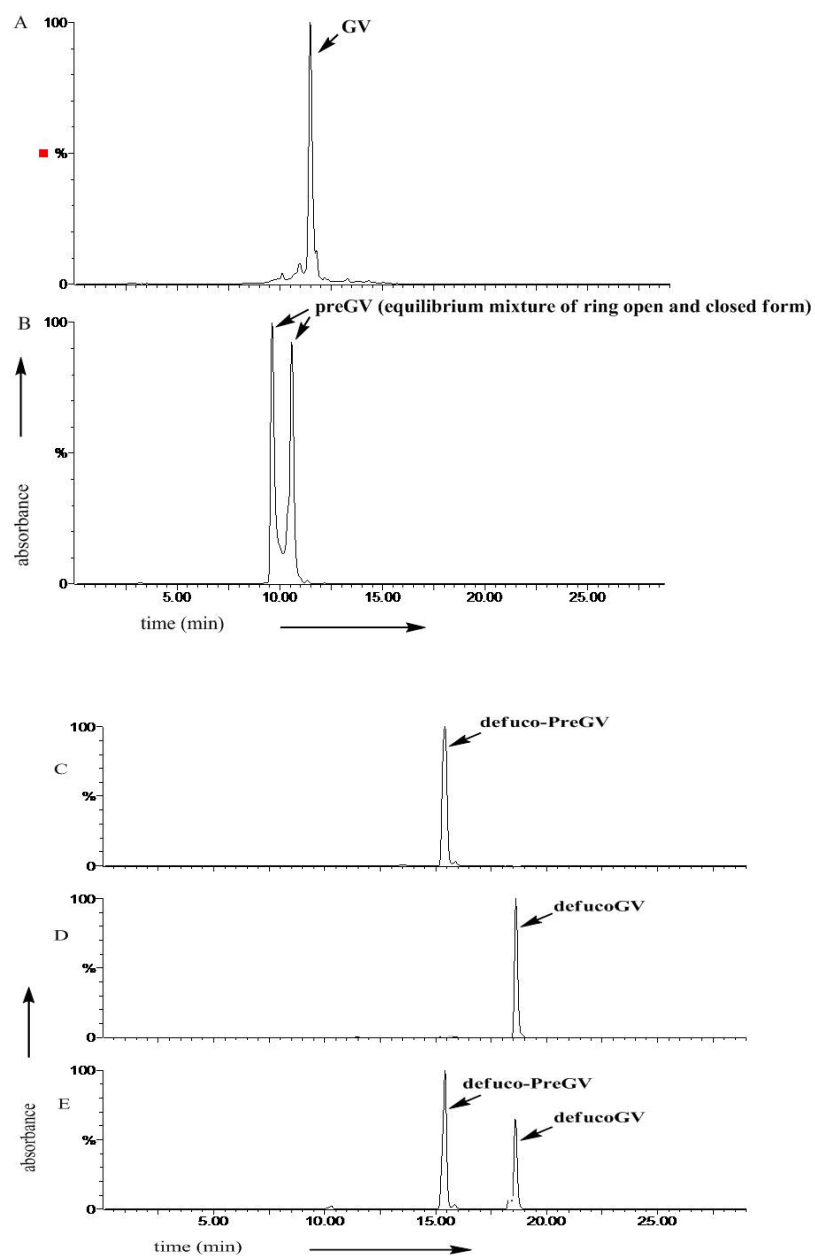
Figure 4.9 HPLC profile of enzyme assay mixtures: trace A; preGV (0.1 mM) incubated with GilR (1 μM) for 10 minutes, trace B; a control sample (without GilR) at the identical conditions, trace C; standard defuco-preGV, trace D; defucoGV, trace E; defuco-preGV (0.1 mM) incubated with GilR (0.5 μM) for 8 minutes.
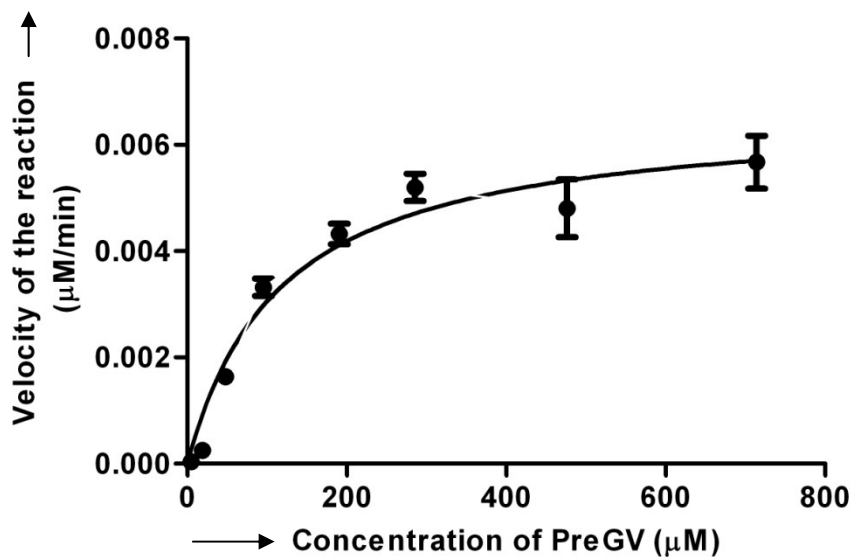
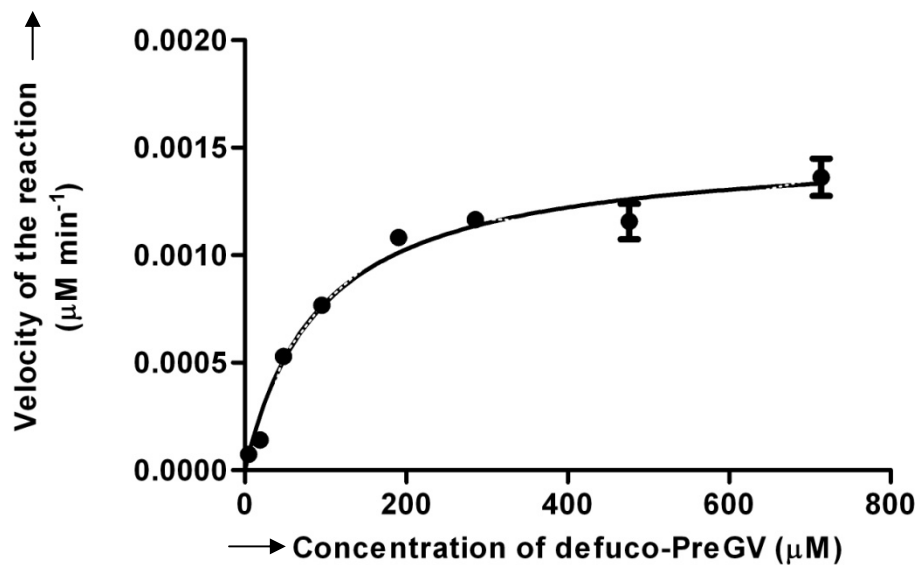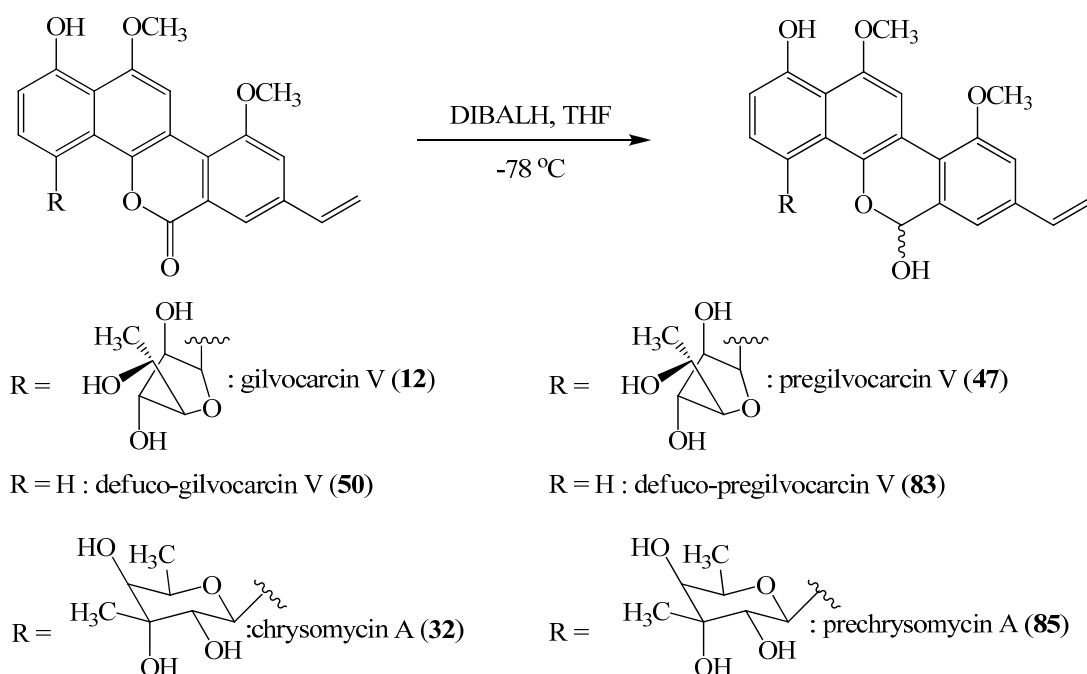**Figure 4.10** Kinetic profile of GilR with preGV.

Figure 4.11 Kinetic profile of GilR with defuco-preGV.

Scheme 4.6 Chemical preparation of preGV and analogues through the reduction of the corresponding natural lactones.

Table 4.3 Kinetic parameters for GilR

Substrate	k_m (μM)	$^a k_{cat}$ (min^{-1})	k_{cat}/k_m	Sp.activity (U mg^{-1}min^{-1})
PreGV (**47**)	117.0 ±24.8	268.31±40.77	2.29±0.03	3.62±0.88
Defuco-PreGV (**83**)	93.4 ±12.73	61.24±6.28	0.65±0.086	0.84 ±0.10
PreChryV (**85**)	nd	nd	nd	1.4×10^{-3}

nd; not determined, ^{a}one active site was assumed per enzyme

Substrate specificity of GilR

One of the major objectives of biosynthetic pathway studies is to generate unnatural derivatives of natural products through the exploitation of the substrate flexibility of pathway enzymes. In this context, alteration of the deoxysugar moiety of GV provides a good means to generate GV analogues by combinatorial biosynthesis. This requires relaxed substrate specificity of both GilGT and GilR. To test the flexibility of GilR regarding the *C*-glycosidically linked deoxysugar moiety, prechrysomycin A (**85**) was prepared through chemical reduction of chrysomycin A (**32**). Surprisingly, the

incubation of **85** with GilR under identical assay condition barely generated **32,** limiting the complete kinetic study of the reaction. However, the specific activity of the enzyme with **85** was determined to be 2500 and 600 fold lower than those with **47** and **83**, respectively, thus indicating its poor specificity for a substrate analogue with a branched pyranose moiety. We were previously able to generate a GV analogue with an extra hydroxyl group at 4' position of the D-fucofuranose moiety through the targeted inactivation of the dedicated ketoreductase gene *gilU* (*102*). This analogue maintained a five-membered sugar residue, and the poor conversion of **85** could be due to the presence of the branched pyranose (D-virenose) which could impose some sterical hindrance in the binding pocket of GilR that appeared optimal for substrates with a 5-membered sugar residue.

Significance

Lactones are often found in varieties of bioactive natural products such as PKS-derived gilvocarcin (GV, **12**), lovastatin (**3**), erythromycin A (**5**), FK-506 (**4**), amphotericin B (**7**), as well as in terpene-derived natural products, such as pentalenolactone and artemisinin. In many cases lactones modulate the biological activity of the molecules directly by interacting with target sites (as in camptothecin)(*191, 192*) or indirectly by providing structural rigidity to the molecule (as in erythromycin). Lactones were also found or proposed as intermediates of various biosynthetic pathways such as jadomycin, gilvocarcin, mithramycin and urdamycin, where the formation of a lactone intermediate is crucial for the installation of polar residues of the final structures (*54, 171, 193*). Interestingly, the ways through which nature installs lactones in secondary metabolites are found to be limited to two general routes (**Scheme 4.6**): (i) through the intramolecular condensation between a hydroxyl group and a carboxylic acid or an ester (route A) and (ii) through Baeyer-Villiger oxidatation of a cyclic ketone (route B). The vast majority of the macrolide lactones are generated through thioesterase (TE)-domain mediated intramolecular condensation of thioester and alcohol (route A),(*194-196*) whereas the lovastatin lactone formation is truly pH-dependent (route A) (*197*). Literature report on Baeyer-Villiger monooxygenases (BVMO)-catalyzed lactone formation in natural products is rapidly growing (*54, 58, 198*). In addition, lactone generation through

the oxidation of a hemiacetal moiety is well established, but was limited so far to primary metabolism, and plays a role especially in the pentose phosphate pathway (route C). We observed such unique chemistry for the first time in the secondary metabolism, in the GV biosynthetic pathway.

Previous gene inactivation results clearly demonstrated that none of the four candidate oxygenases of the GV pathway (GilOI, GilOII, GilOIII, GilOIV) contribute to the formation of the lactone moiety of GV. Complementing the previous gene inactivation results, the current *in vitro* assay results have clearly established that GilR is an FAD-dependent oxidoreductase and is responsible for oxidizing hemiacetal **47** to lactone **12** (**Scheme 4.7,** route I). In the pentose phosphate pathway, glucose-6-phoshate (G-6-P) dehydrogenase catalyzes a similar reaction converting G-6-P to 6-phospho-D-gluconolactone (*199*). However, unlike GilR, this enzyme utilizes free standing $NADP^+$ thereby placing GilR into a new class of oxidoreductases. Since FAD is bound covalently to GilR, and no other reducing equivalent was used in the *in vitro* assay, we anticipate that $FADH_2$ reacts with dissolved oxygen to regenerate FAD. The results also clarify that the lactone generation is the last step of the post-PKS tailoring steps of GV biosynthesis, and the enzyme is flexible enough to handle both sugar-tethered and sugar-free substrates, while its flexibility is limited by the nature of the linked deoxysugar. Understanding the substrate binding mode of GilR and widening its substrate specificity through site-directed mutagenesis might be necessary to generate further GV analogues by combinatorial biosynthesis. The results of this study also show indirectly that defuco-preGV, and not defuco-GV as previously assumed, is the substrate of GilGT, the *C*-GT of the GV pathway.
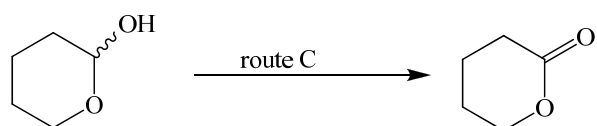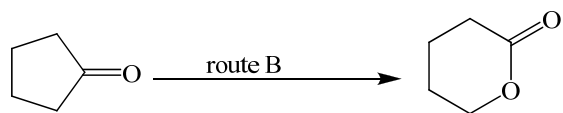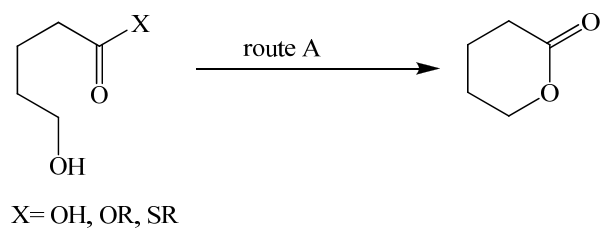
Scheme 4.7 General routes for lactone formation in natural products.
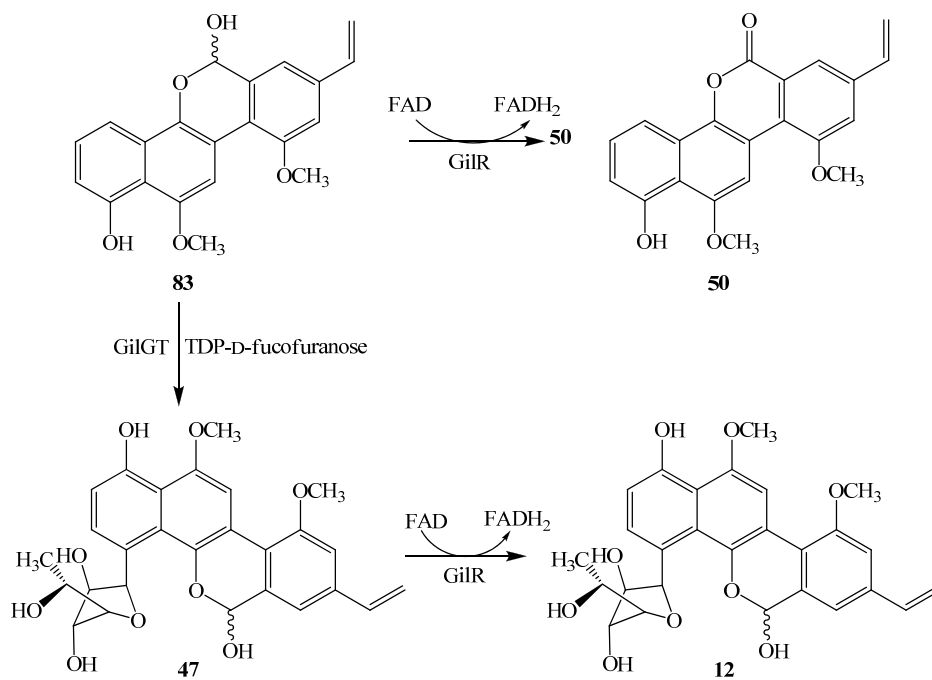
Scheme 4.8 Proposed pathway for GV biosynthesis.

Experimental section

Bacterial strains, culture conditions

Streptomyces strains *S. griseoflavus* Gö 3592, GilGT-minus mutant *S. lividans* TK24 (cosG9B3-GT‾), GilR-minus mutant *S. lividans* TK24 (cosG9B3-R‾) and *S. albaduncus* AD819 were cultured on M2 agar (*56*) or in SG liquid medium (*56*) at 28 °C and 200 rpm. *E. coli* strains were grown in Luria Broth (LB) or on agar at 37 °C. Kanamycin sulfate (50 μg mL^{-1}) and apramycin sulfate (50 μg mL^{-1}) were supplemented to the media, whenever necessary. Standard pregilvocarcin V was isolated from the *S. lividans* TK24 (cosG9B3-R‾) mutant following the previously reported procedure (*56*). The compound was generated in larger scale by chemical reduction of GV (described later in this section). Chrysomycin A was isolated from *S. albaduncus* AD819. Spores of each *Streptomyces* strain were inoculated into 100 mL of SG medium in 500 mL-Erlenmeyer flasks, and the culture was grown for 3 days to prepare a seed. The seed culture (3% by vol) was then taken for inoculation of production scale SG medium (5 L; 100 mL × 50 flasks) and grown for 7 days.

Preparation of expression constructs

All of the anticipated defucoGV biosynthetic genes (*gilA, B, C, F, K, P, OI, OII, OIV, OIII, R, M, MT*) were amplified from cosG9B3 using specific primers (**Table 4.4**) and *pfu* polymerase (Stratagene). Similarly, some of the ravidomycin (*ravG, C, C1*) and jadomycin PKS (*jadC, D, I*) genes were amplified using cosRav32 and pJV779, respectively. PCR products were cloned into Zero-Blunt TOPO vector (Invitrogen) and sequenced. Jadomycin oxygenase (encoded by *jadF*) was also cloned following an identical protocol and *gilA* and *gilB* were amplified together. Topo clones were digested with *Nde*I *and Eco*RI to excise the gene fragments and ligated at the identical sites of pET28a vector for all genes except *jadI* which was cloned at *Nde*I and *Hind*III site. *GilAB* fragments were cloned at *Nde*I and *Eco*RI sites of pXY200. Expression constructs generated in this study are summarized in the **Table 4.5**.

Table 4.4 List of primers used in this study

Name of the primers	Oligonucleotide sequence
GilOIV_ex_for	5′-GGGCATATGACGGAGCCCGAGACCTCGGAC-3′
GilOIV_ex_rev	5′-GAGAATTCTCACCGCGTCACCTCGGGTTCCA-3′
GilOII_for	5′-GGCCATATGCCGATCATCGAGCCCGAGAAC-3′
GilOII_rev	5′-CGGGAATTCTCACGACGCGTACCCCTCGACCGA-3′
GilR_ex_for	5′-CGCCATATGACCGCTTCCGTACCGCCGTTCACGGTG-3′
GilR_ex_rev	5′-CCAGAATTCTCAGAGTCCTATGGACATGCTGTG-3′
GilM_ex_for	5′- CGCCATATGCCAACGGGCAGCACGGAGAAGATC-3′
GilM_ex_rev	5′-CGGGAATTCTCATGCCGGGCTCTCCGATCGGGT -3′
GilMT_ex_for	5′-GACCATATGACCATTACTGCATCGGGATCG-3′
GilMT_ex_rev	5′-CGGGAATTCTCACCGGCTGCGGGGGAGAGCCCGG-3′
JadF_ex_for	5′-ATTCATATGACGGAGCCCCGCCGGGCAGGA-3′
JadF_ex_rev	5′-TTAGAATTCTCAGAGGTGGGTGGGGGCGCC -3′
GilA_exp_for	5′- ATTCATATGAGCCGCAGGGTCTTCATCACC-3′
GilB_exp_rev	5′-ATTGAATTCTCATTTCCCGACGATCAGGGCCGA-3′
GilC_exp_for	5′- ATCCATATGTCCGCACGCGTCACCATGGAC-3′
GilC_exp_rev	5′-ATGAATTCTCACGATGCCGCCGCCTGACGC-3′
GilF_exp_for	5′- ATTCATATGACCTCTCCCCGTCATGCCCTG-3′
GilF_exp_rev	5′-ATGAATTCCTAGAAGTTGCCGAGGCCGCC-3′
GilK_exp_for	5′-ATCATATGGCCGATCCGGCTCGCACAGAC-3′
GilK_exp_rev	5′-ATGAATTCTCACCTCTGCGACACGGCGGTG-3′
GilP_exp_for	5′-ATCATATGAGAGCGTTCCTGTTCCCCGGT -3′
GilP_exp_rev	5′- ATGAATTCTCACCAGCCCGTCACGCC -3′
RavC1_for	5′- ATTCATATGACCACCGGCACGTTCACC-3′
RavC1_rev	5′- ATTGAATTCTCACGCCGCGTTGACCAGCTC-3′
RavG-for	5′-ATTCATATGCACAGCACCCTGATCGTCGCC-3′
RavG-rev	5′-ATTGAATTCTCAGTCCTGCTGCCAGTGGTA-3′
RavC_for	5′-ATTCATATGAGCAGCTTCTCGATCGACGACCTC-3′
RavC_rev	5′-ATTGAATTCTCAGCCGCGCGGCGCGGG -3′
JadC_for	5′-ATTCATATGAGCAGCAAGACCTTCACCCTC-3′
JadC_rev	5′-ATTGAATTCTCAGGCGGCGGCGACGGC-3′
JadI_for	5′-ATTCATATG CACAGCACTCTGATCGTGGCC-3′
JadI_rev	5′-ATTAAGCTTTCACGCGTTCGCCTCCCAGTT-3′
JadD_for	5′-ATTCATATGACGACCCGTGAGGTCGAGCAC-3′
JadD_rev	5′-ATTGAATTCTCAGCGCTTGCCCTCGGCGTA -3′

Table 4.5 List of plasmids constructed/used in this study.

Strain/ plasmid	Characteristics and relevance	References
E. coli XL1-Blue-MRF	Host for routine cloning works and for the construction of genomic libraries	Stratagene
E. coli BL21 (DE3)	Expression host for the various expression constructs	Novagen
E. coli (BAP1)	Expression of holo ACPs	Pelzer et al.(*176*)
S. lividans TK24	For expression of gilvocarcin PKS enzymes	Hopwood et al.(*170*)
*S. lividans*TK24/cosG9B3-GilGT⁻	For the production of defucoGV	Rohr et al.(*33*)
S. albandancus	For the production of chrysomycin A	Asheshov et al.(*200*)
pGilL	*gilL* gene cloned into pET28a	This work
pGilN	*gilN* gene cloned into pET28a	This work
pGilU	*gilU* gene cloned into pET28a	This work
pGilOI-1	*gilOI* gene cloned into pET28a	This work
pGilOII-1	*gilOII* gene cloned into pET28a	This work
pGilOIV-1	*gilOIV* gene cloned into pET28a	This work
pGilR-1	*gilR* gene cloned into pET28a	This work
pGilM-1	*gilM* gene cloned into pET28a	This work
pGilMT-1	*gilMT* gene cloned into pET28a	This work
pGilGT-1	*gilGT* cloned into pET28a(+)	This work
pGilF	*gilF* cloned into pET28a(+)	This work
pGilK	*gilK* cloned into pET28a(+)	This work
pGilP	*gilP* cloned into pET28a(+)	This work
pGilAB	*gilA* and *GilB* cloned into pXY200	This work
pGilC	*gilC* cloned into pET28a(+)	This work
pRavC	*ravC* cloned into pET28a(+)	This work
pRavC1	*ravC1* cloned into pET28a(+)	This work
pRavG	*ravG* cloned into pET28a(+)	This work
pJadC	*jadC* cloned into pET28a(+)	This work
pJadD	*jadD* cloned into pET28a(+)	This work
pJadI	*jadI* cloned into pET28a(+)	This work
pWJ292	Malonyl CoA synthetase expression construct	Tang et al(*201*).
pET28a	Expression vector	Novagen
pXY200	Thiostrepton inducible expression vector	Zabriskie et al.(*175*)
PCR-Blunt-II-TOPO	To clone PCR products	Invitrogen

Expression and purification of enzymes

All of the genes were expressed with *N*-terminal polyhistidine tag. pET-28a-expression constructs were expressed in *E. coli* BL21 (DE3) following the protocol discussed in section **3**. Purification of protein was achieved using IMAC through the Profinia protein purification system (BioRad). Enzymes were purified occasionally using metal talon affinity resin (BD Biosciences). Concentrations of proteins were determined by the Bradford method using a calibration curve of the known concentrations of BSA

(*169*). Purified enzymes were stored in 50 mM phosphate buffer (pH 7.5) supplemented with glycerol (25% final volume). Protoplasts of *S. lividans* TK64 were transformed with pGilAB following the standard protocol (*143*). The transformants were selected for apramycin resistance (50 μg mL^{-1}) phenotypes. A single transformant colony was inoculated into a 500 mL Erlenmeyer flask containing 100 mL TSB medium. The culture was grown for 3 days at 28 °C with constant shaking at 250 rpm. A fully grown seed culture (2 mL) was inoculated into 100 mL (× 40 flasks) of R2YE medium. After 24 hours of incubation, thiostrepton (20 μg mL^{-1}) was added in the culture to induce the expression of GilA-GilB complex. Incubation was continued for another 48 hours and the culture was harvested through centrifugation (4000 × g, 15 min). Cell pellets were washed twice with lysis buffer prior to their lyses. Disruption of cells was carried out in a French Press (Thermo Electron Corporation) and the crude enzyme was separated from the cell debris through centrifugation (15000 × g, 60 min). Thus obtained crude enzymes were filtered through 0.5 μM syringe filter (Millipore) and the filtrate was subjected to the purification using the Profinia (Bio-Rad).

Preparation of *holo*-ACPs

Holo form of ACPs (RavC, RavC1 and JadC) were achieved directly by expressing the enzymes in the genetically engineered *E. coli* BAP hosts. Purified enzymes (5-8 mg) were subjected to the PD10 desalting column (GE Healthcare) pre-equilibrated with water. Water was used to elute the enzymes and aliquots of 200 μL each were collected and analyzed in a UV spectrophotometer. The first fraction containing enzyme was analyzed by MALDI-TOF mass at University of Kentucky Mass Spectrometry facility.

Co-factor analyses of GilOI and JadF

GilOI and JadF were eluted as intense yellow enzymes. To identify the enzyme bound cofactors, 1 mL enzyme solution (4 mg mL^{-1}) of each enzyme was boiled for 5 minutes to release the cofactor from the enzyme. The precipitated enzyme was removed by centrifugation (12000 ×g, 5 min) and the yellow solution was analyzed by HPLC. Standard solution (1 μM) of FAD and FMN were used in parallel for comparison.

Carbopac PA1 (Dionex Inc.) analytical column (4 × 250 mm) was injected with 50 μL sample. A gradient program of ammonium acetate (500 mM) and water (ammonium acetate: 5 to 20% over 15 min, 20 to 60% over 20 min, 60 to 100 % over 2 min, 100% over 3 min and 100% to 5% over 5 min and 5% over 15 min) was used to elute compounds. The flow rate and detector were set to 1 mL min^{-1} and 267 nm, respectively. FMN and FAD were eluted at 35 min and 50 min, respectively.

Preparation of malonyl-CoA *in situ*

Malonyl-CoA was prepared enzymatically from sodium malonate, ATP and coenzyme A (CoA). Assay mixtures of variable volumes (200 μL- 20 mL) composed of sodium malonate (5 mM), magnesium chloride (5 mM), ATP (25 mM), CoA (1.5 mM) and MatB (6 μM) were incubated for 30 minutes. The formation of malonyl-CoA was monitored by HPLC. CarboPack^{TM} PA1 (4×250 mm) column (Dionex co.) was used to separate compounds. A gradient of ammonium acetate (500 mM, solvent B) and water (solvent A) was used to separate the assay products: 5-20% solvent B from 0-15 min, 20-60% solvent B from 15-35 min, 60-100% solvent B from 35-37 min, 100% solvent B from 37-40 min, 100-5% solvent B from 40-45 min and 5% solvent B from 45-60 min. The flow rate was maintained to 1 mL min^{-1} throughout the separation. ATP, CoA and malonyl-CoA were eluted at 52.5, 33.0 and 31.5 minutes, respectively.

PKS Enzyme assay

One-pot PKS enzyme assays were carried out in two scales. For the analytical scale, assay mixtures composed of phosphate buffer (pH 7.5, 50 mM), sodium malonate (5 mM), magnesium chloride (5 mM), ATP (25 mM), CoA (1.5 mM), MatB (6 μM), acetyl CoA (100 μM) were added with various combinations of PKS enzymes [GilAB, ~0.2 μM; RavC, 2.1 μM; RavC1, 2.0 μM; JadC, 2.5 μM; GilP, 3.1 μM; GilF, 0.1 μM; RavG, 0.15 μM; JadD, 0.06 μM, JadI, 0.2 μM] and the final volume of the assay mixture was maintained at 200 μL. Incubation of the assay mixture was carried out at 28 °C for 4 hours and the reaction was stopped by thorough mixing with 400 μL of ethyl acetate. The organic layer was collected and dried in a speed vac. The extract was dissolved in methanol (150 μL) and then filtered through a syringe filter (pore size 0.25 μM) prior to

the HPLC/MS analyses. Separation and MS analyses were carried out following the procedure described in Section 2. For larger scale assays, the volume was increased to 20 mL and the mixtures were incubated for 12 hours.

Enzymatic synthesis of defucoGM

One pot enzymatic synthesis of defuco-GM was carried out using PKS, post-PKS enzymes and the anticipated cofactors. The assay mixture (10 mL) containing phosphate buffer (pH 7.5, 50 mM), sodium malonate (5 mM), magnesium chloride (5 mM), ATP (25 mM), CoA (1.5 mM), MatB (6 μM), acetyl CoA (100 μM)**,** GilAB (~ 0.2 μM), RavC (2.1 μM), GilP (3.1 μM), GilF (0.1 μM), RavG (0.15 μM), JadD (0.06 μM) GilOI (1.58 μM), GilOII (0.9 μM), JadF (0.95 μM), GilM (0.15 μM), GilR (0.10 μM), JadI (1.35 μM), FAD (12.5 μM), NADH (0.62 mM), NADPH (0.5 mM), SAM (25 μM), NADH:FAD-reductase (10 μM) was incubated at 28 °C for 8 hours. The reaction was stopped by extracting the products with 40 mL ethyl acetate. The organic layer was dried under reduced pressure and dissolved in methanol. Then the sample was passed through a 0.25 μM syringe filter for the HPLC-MS analysis. In the control experiment, buffer was added in place of GilAB.

Cofactor analysis of GilR

The UV spectrum of the purified enzyme was typical to that of FAD (**Figure 4.8**). To analyze the bound cofactor, the enzyme (2.37 mg) was boiled for 5 minutes and centrifuged (12000 × g, 15 minutes). Interestingly, the supernatant was a clear solution, and a yellow solid residue was noticed at the bottom of the tube. No UV-absorption was found for the supernatant but a typical of FAD spectrum found when the solid residue was dissolved in 100 μL of 10% SDS solution and analyzed in a spectrophotometer. The protein was precipitated again with the addition of trichloroacetic acid (TCA, 400 μL) and then centrifuged to obtain the supernatant. Again, the supernatant did not absorb in the UV, but the pellet fraction remained yellow. The results clearly indicated that FAD is covalently bound to the enzyme.

**Kinetic study of GilR catalysis**

A typical assay mixture (105 μL) composed of 0.1 mM substrate, 9.5 mM MgCl$_2$, 50 mM phosphate buffer and 0.1-1 μM enzyme and 15% DMSO (final concentration) were incubated at 20-45 °C for 4-10 minutes to estimate the optimal assay conditions for the kinetic analysis. The reaction was terminated through flash freezing at -80 °C. The product was extracted twice with 300 μL of ethyl acetate, and the extract was combined and dried in vacuom. The extract was then dissolved in 100 μL of acetonitrile and filtered through 0.2 μM membrane. 70 μL of the filtrate was injected in the HPLC/MS. The amount of product formation was estimated through plotting the product peak area in the standard calibration curve. For the kinetic analysis, 8 different substrate concentrations (4.76 to 714 μM final concentration) were incubated with the enzyme (0.23 μM final concentration) for 5 minutes at 37 °C. The resulting data were fitted into the Michaelis-Menten equation by nonlinear regression using GraphPad Prism 5.0 to determine the k_{cat} and K_m values. Three replicates of each sample were analyzed and the average was taken as a result.

Isolation and purification of chrysomycin A, defuco-gilvocarcin V and gilvocarcin V

The fully grown culture of *S. albaduncus* was harvested by centrifugation (4000 × g, 15 min). The cell pellets were resuspended in 400 mL of acetone, sonicated for 20 min and centrifuged to obtain the acetone extraction fraction. The process was repeated twice, the acetone fractions were combined and the solvent was removed using a rotary evaporator. Finally, the extract was resuspended in 200 mL of water and mixed with the culture broth. Reverse phase silica (Rp$_{18}$, 200 g) was loaded into a column (5 × 65 cm) and equlibrated with deionized water. The culture broth was passed through the column and the effluent was discarded. The column was washed with 1 L of 10% acetonitrile (in water). Elution was carried out using an increasing gradient of acetonitrile (20, 40, 60, 80 and 100%, 1L each). HPLC/MS analyses revealed that chrysomycin A was the major compound in the 40% to 60% fractions. Acetonitrile was removed from these fractions and the aqueous fractions were combined and dried in a lyophilizer. The dried extract was dissolved in 30 mL dimethylsulfoxide (DMSO), filtered through a 0.25 μM filter, and the filtrate was subjected to preparative HPLC separation. Fractions containing pure

chrysomycin A were collected and dried. The purity and authenticity of the compound was confirmed through ^{1}H and ^{13}C NMR analyses. A linear gradient of acetonitrile and acidified water (solvent A = 0.1% formic acid in H_2O; solvent B = acetonitrile; 0-15 min 25% B to 100% B; 16-24 min 100% B; 25-26 min 100% to 25% B; 27-29 min 25% B) was used to separate compounds. SunFire™ prepC$_{18}$ (19 × 150 mm, 5 μm) and Symmetry C$_{18}$ (4.6 × 250 mm, 5 μm) columns were used for preparative and analytical scale separations, respectively. The flow rate was maintained at 10 mL min^{-1} and 1.0 mL min^{-1} for preparative and analytical scale separations, respectively. Micromass ZQ 2000 (Waters corporation) equipped with HPLC (Waters alliance 2695 model) and photodiode array detector (Waters 2996) were used to analyze the compounds. Atmospheric pressure chemical ionization (APCI) probe was used to detect molecular ions. A similar procedure was followed to isolate defucogilvocarcin V and gilvocarcin V from *S. lividans*TK24/cosG9B3-GilGT¯ and *S. griseoflavus*, respectively.

Preparation of pregilvocarcinV (preGV, 47), defuco-preGV (83) and prechrysomycin A (85)

47
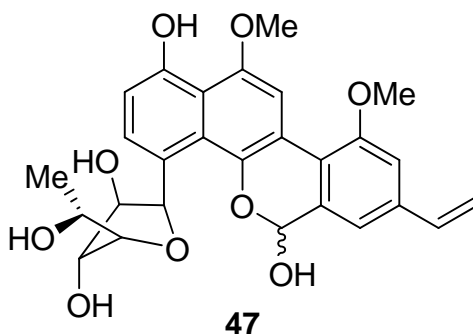
Pregilvocarcin V (**47**). A solution of DIBALH (1M in hexane, 0.4 mL, 0.4 mM) was added dropwise to a stirred solution of gilvocarcin V (**1**) (20 mg, 0.04 mM) in THF (7 mL) at - 78 °C. After 1 h ethyl acetate (10 mL) was added to quench the reaction. The solvent was removed in vacuom and the mixture was purified through reverse phase HPLC (Symmetry Prep C$_{18}$7μm column, 19 × 150 mm) using a linear gradient of acetonitrile and water to afford **11** (14 mg, 70%) as colorless solid. The structure was confirmed through the comparison of NMR and MS data with the standard compound isolated from the GilR-minus mutant (*56*).

83
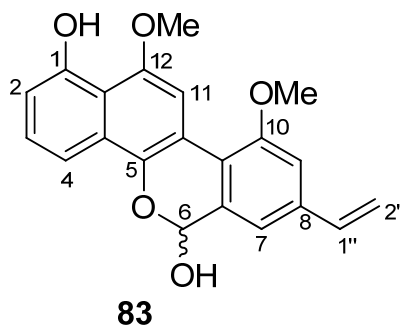
Defuco-pregilvocarcin (**83**). A solution of DIBALH (1M in hexane, 0.4 mL, 0.4 mM) was added drop wise to a stirred solution of defuco-gilvocarcin V (**9**) (30 mg, 0.09 mM) in THF (10 mL) at -78 °C. After 1 h ethyl acetate (10 mL) was added to quench the reaction. Solvent was removed in vacuom and the mixture was purified through reverse phase HPLC (Symmetry Prep C_{18} 7μm column, 19 × 150 mm) using a linear gradient of acetonitrile and water to afford **12** (22 mg, 73%) as light yellow solid. ^{1}H NMR (DMSO-d_6, 500 MHz) δ 9.48 (s, 1H, 1-OH), 8.03 (s, 1H, H11), 7.72 (d, J = 8.0 Hz, 1H, H4), 7.37 (t, J = 8.0 Hz, 1H, H3), 7.32 (d, J = 5.5 Hz, 1H, 6 -OH), 7.28 (s, 1H, H9), 7.19 (s, 1H, H7), 6.82 (d, J = 8.0 Hz, 1H, H2), 6.80 (dd, J = 17.0, 11.0 Hz, 1H, H1″), 6.37 (d, J = 4.0 Hz, 1H, H6), 5.98 (d, J = 17.0 Hz, 1H, H2″$_E$), 5.35 (d, J = 11.0 Hz, 1H, H2″$_Z$), 4.06 (s, 3H, -OMe), 4.02 (s, 3H, OMe); ^{13}C NMR (DMSO-d_6, 125 MHz) δ 156.2, 153.8, 149.3, 140.6, 137.8, 136.2, 134.8, 128.1, 127.2, 116.4, 116.3, 115.0, 114.3, 113.1, 111.0, 109.9, 103.6, 92.3, 56.2, 56.1.

85
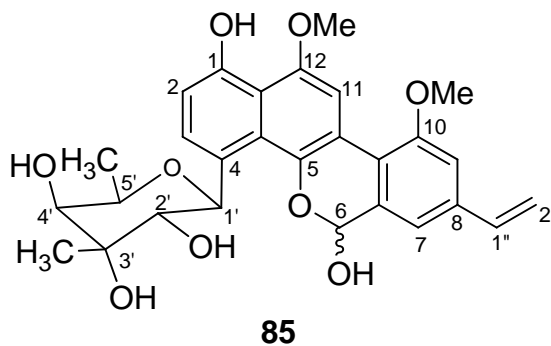
Prechrysomycin A (85). A solution of DIBALH (1M in hexane, 0.4 mL, 0.4 mM) was added dropwise to a stirred solution of chrysomycin A (**10**) (20 mg, 0.04 mM) in THF (7 mL) at - 78 °C. After 1 h ethyl acetate (10 mL) was added to quench the reaction. Solvent was removed in vacuom and the mixture was purified through reverse phase HPLC (Symmetry Prep C_{18}7μm column, 19 × 150 mm) using a linear gradient of acetonitrile and water to afford **13** (11 mg, 55%) as colorless solid. It exists as 1:1 mixture of two diastereomers. ^{1}H NMR (DMSO-d_6, 500 MHz) δ 9.81 and 9.78 (2 × s, 1H, -OH, D_2O exchangeable), 8.00 and 7.94 (2 × s, 1H, H11), 7.75 and 7.73 (2 × d, J = 8.5 Hz, 1H, H3), 7.30 and 7.08 (2 × d, J = 6.5 Hz, 1H, -OH, D_2O exchangeable), 7.27 (2 × s, 1H, H9), 7.18 (2 × s, 1H, H7), 6.83 (2 × d, J = 8.5 Hz, 1H, H2), 6.80 (2 × dd, J = 17, 11 Hz, 1H, H1″), 6.46 and 6.33 (2 × d, J = 9.0 Hz, 1H, H1′), 6.28 (2 × d, J = 5.0 Hz, 1H, H6), 5.99 (2 × d, J = 17.0 Hz, 1H, H2″$_E$), 5.36 (2 × d, J = 11.0 Hz, 1H, H2″$_Z$), 4.67 (d, J = 8.0 Hz, 1H, -OH, D_2O exchangeable), 4.67 and 4.51 (2 × s, 1H, -OH, D_2O exchangeable), 4.25 and 4.20 (2 × q, J = 7.0 Hz, 1H, H5′), 4.06 (2 × s, 3H, -OMe), 4.01 and 4.00 (2 × s, 3H, -OMe), 3.97 and 3.61 (2 × d, J = 8.0 Hz, 1H, -OH, D_2O exchangeable), 3.38 and 3.28 (2 × d, J = 9.5 Hz, 1H, H2′), 3.13 (2 × d, J = 8.0 Hz, 1H, H4′), 1.25 and 1.21 (2 × s, 3H, 3′-OMe), 1.13 and 1.08 (2 × d, J = 7.0 Hz, 3H, 5′-OMe); ^{13}C NMR (DMSO-d_6, 125 MHz) δ (156.2, 156.1), (153.2, 153.0), (149.8, 149.7), (142.5, 142.2), (137.9, 137.8), (136.3, 136.2), (135.6, 134.5), (130.2, 129.9), (127.9, 127.3), (127.2, 127.0), 117.0, (116.8, 116.6), (116.1, 115.8), (115.3, 115.2), (114.7, 114.6), (111.0, 110.9), (110.1, 109.9), (103.7, 103.6), (92.5, 91.8), (76.5, 76.4), (76.1, 75.8), 74.2, (73.5, 73.2), (71.1, 71.0), 56.4, 56.1, (23.5, 23.4), 17.2.

Summary and future studies

In summary, the studies reported in this dissertation cover the cloning and characterization of the ravidomycin biosynthetic gene cluster. Functional roles of many of the PKS, the post-PKS and the deoxysugar biosynthetic enzymes involved in the biosynthetic pathway of this antibiotic were established. The insights gained into the activities of these enzymes provide a framework for the generation of gilvocarcin V/ravidomycin V analogues through combinatorial biosynthesis.

In **Specific aim 1,** we cloned and sequenced the entire gene cluster responsible for ravidomycin biosynthesis. The implication of this cluster in ravidomycin biosynthesis was demonstrated through the production of 4′-*O*-deacetylravidomycin E in a heterologous host. The functional role of RavOIII that generates the vinyl side chain of ravidomycin was established. An unprecedented sugar donor substrate flexibility of the RavGT that is capable of utilizing 6-membered amino sugar as well as 5-membered neutral sugar was also demonstrated.

In **Specific aim 2**, *in vitro* enzyme assays of the recombinant enzymes established the functional roles of all of the proposed TDP-D-ravodosamine biosynthetic genes *ravD*, *ravE*, *ravIM*, *ravAMT* and *ravNMT*. Kinetics of keto-isomerizations catalyzed by RavIM and its homologue FdtA were studied in detail. An efficient protocol for one-pot *in vitro* enzymatic synthesis of TDP-D-ravidosamine was developed through a combination of suitable biosynthetic enzymes from various pathways.

In **specific aim 3**, the angucyclinone rabelomycin was enzymatically synthesized *in vitro* using enzymes from GV, RMV and jadomycin pathways. The requirement of synchronous actions of PKS and post-PKS enzymes to biosynthesize GV was demonstrated *in vitro* through the one-pot enzymatic synthesis of defucoGM, and thus outlined the minimal set of enzymes necessary to biosynthesize the benzo[*d*]naptha[1,2-*b*]pyran-6-one moiety of GV. Through the enzyme assay and kinetic studies, GilR was determined to be a dehydrogenase and was shown to be responsible for dehydrogenation of the hemiacetal moiety of preGV at the last step of the GV biosynthetic pathway.

Future studies will focus on the characterization of the remaining GV/RMV post-PKS tailoring enzymes. Analyses of the products generated through co-assay of the aforementioned PKS enzymes with the individual post-PKS enzymes in all possible combinations will potentially generate all of the intermediates of the defucoGM pathway. In this context, co-assay of PKS enzymes, the oxygenases (GilOI, GilOII or GilOIV) and GilM or GilMT will be the first experiment to test. The isolation and characterization of the products (pathway intermediates or shunt products) will help to assign the exact functional roles of GilM and GilMT as well as their sequence of action. In addition, these intermediates can be tested as sugar acceptor substrates of RavGT. The sugar donor substrate of RavG, presumably TDP-D-ravidosamine, can be routinely produced following the protocol developed in this study. Having all of the biosynthetic enzymes in hand, a one-pot enzymatic synthesis of ravidomycin M (RM) will be an interesting experiment to try as the successful reconstitution of the pathway. This will not only set up a framework for the *in vitro* biosynthesis of RMV/GV analogues but also demonstrate the tremendous power of biosynthesis to generate highly complex molecules, which are otherwise impossible to generate in such a concerted way using currently available technology.

Given the fact that the C-8 vinyl side chains of GV/RMV serve as a warhead of these antibiotics, future experiments will involve the characterization of the dedicated enzymes GilOIII/RavOIII. In this context, acetyl-CoA will be replaced by propionyl-CoA and a potential propionyl-CoA:ACP acyltransferase (GilQ/RavQ) will be included in the assays. As discussed earlier, co-assay of the PKS and the post-PKS enzymes will result GV/RMV pathway intermediates with ethyl side chain instead of methyl group. These compounds will potentially serve as substrates for GilOIII/RavOIII. Successful characterization of GilOIII/RavOIII together with the other enzymes (Gil/RavM, MT and GT) will clarify all of the current ambiguities regarding the structure of the intermediates and the sequence of the biosynthetic events. In addition, the results will provide an opportunity to assess the gatekeeping vs promiscuity of the individual enzymes with regard to their unnatural substrates. Promiscuity of these enzymes can be exploited to generate new GV analogues through mutasynthesis as demonstrated earlier in generating enterocin,(*202*) rapamycin,(*203*) or lacticin 481 analogues.(*204*) The unusual sugar donor

substrate flexibility of RavGT observed in this study is particularly promising for its application in generating GV/RMV analogues with diverse sugars. In this context, experiments involving the co-expression of *ravGT* with various deoxysugar biosynthetic plasmids in the GilGT-deletion mutant (ΔGilGT) seem promising to generate novel GV/RMV analogues.

Appendices

Figure A1. Multiple sequence alignment of ketoacyl synthase (α) from several angucycline biosynthetic pathways (CmmP from chromomycins, GilA from GV, PgaA from gaudimycins, UrdA from uradamycins and Stfp from steffimycin biosynthetic pathways). Consensus amino acid sequence (highlighted by arrow) were used to design conserve primers for the amplification of internal gene sequence of ketoacyl synthase (α) from *S. ravidus*. The expected amplification (~0.65 Kb) is shown in the DNA agarose gel picture.

Figure A2. Multiple sequence alignment of 4,6-dehydratases from several deoxysugar biosynthetic pathways (CosH from cosmomycin, RhoH from rhodomycin, Gra-ORF1 from granaticin, LanH4 from landomycin and SchS5 from Sch47554 and Sch4755 biosynthetic pathways). Consensus amino acid sequences (highlighted by arrow) were used to design conserve primers for the amplification of the internal gene sequence of 4,6-dehydratase from *S. ravidus*. The expected amplification (~0.60 Kb) is shown in the DNA agarose gel picture.
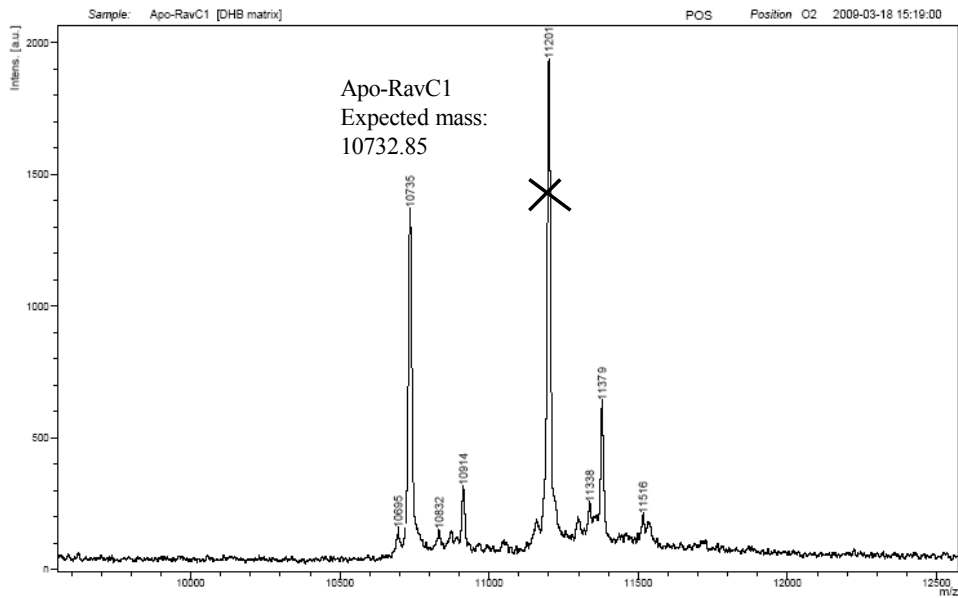
Figure A3. MALDI-TOF analysis of *apo*-RavC1
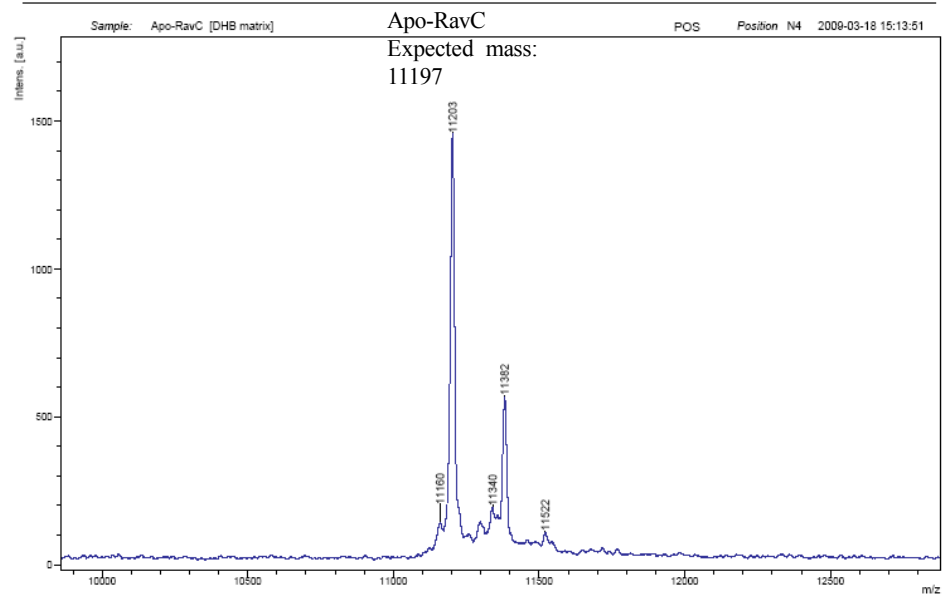
Figure A4. MALDI-TOF analysis of *apo*-RavC
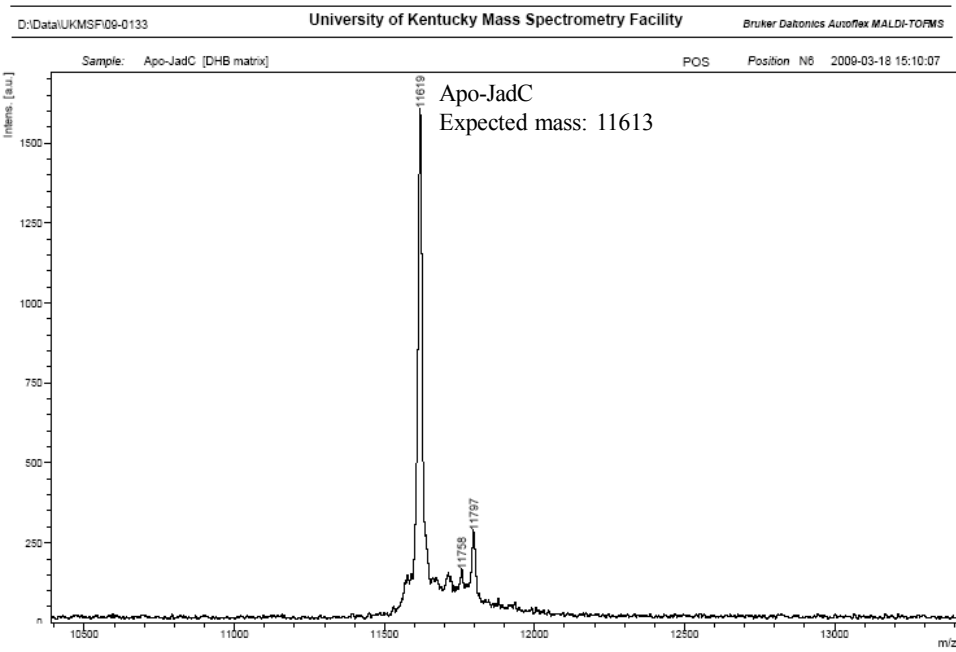
Figure A5. MALDI-TOF analysis of *apo*-JadC
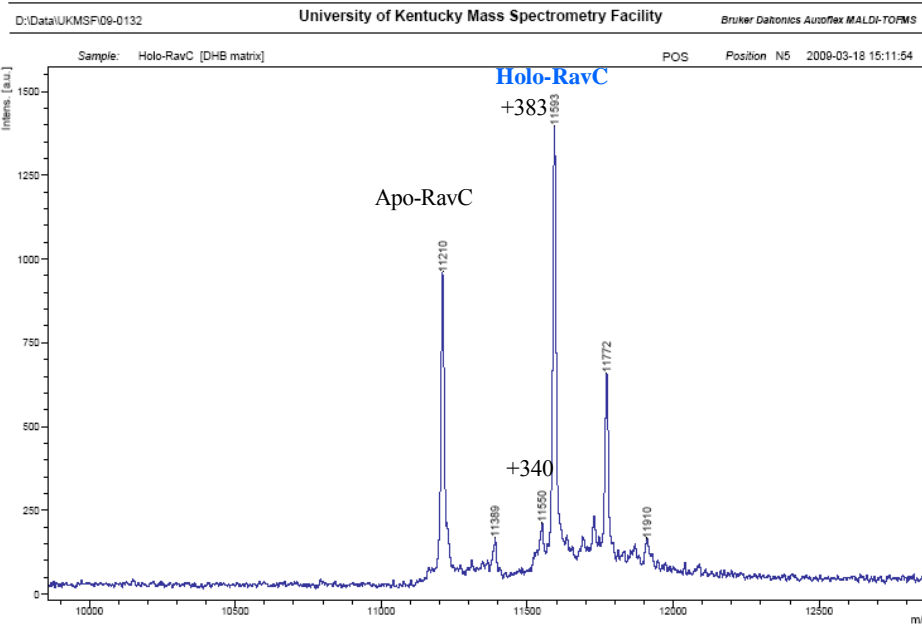
Figure A6. MALDI-TOF analysis of *holo*-RavC
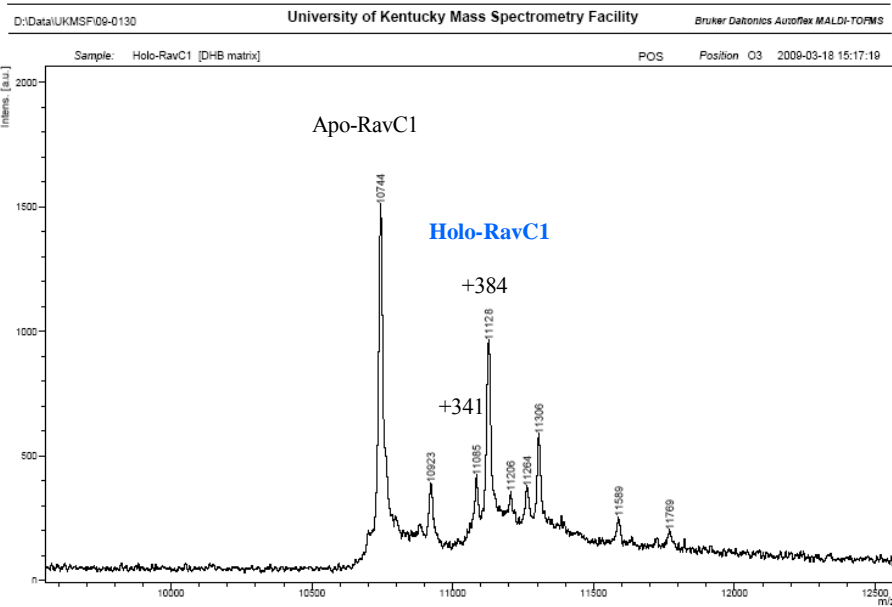
Figure A7. MALDI-TOF analysis of *holo*-RavC1
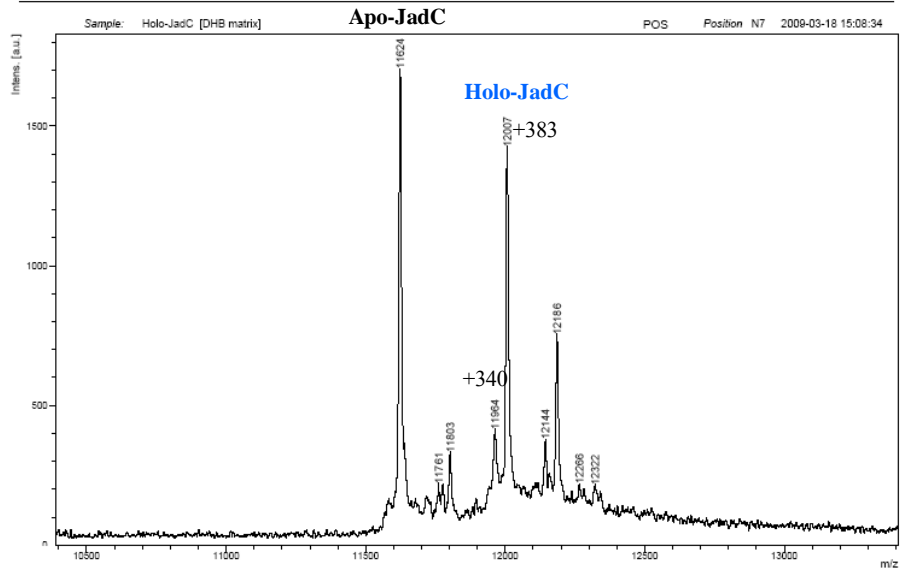
Figure A8. MALDI-TOF analysis of *holo*-JadC
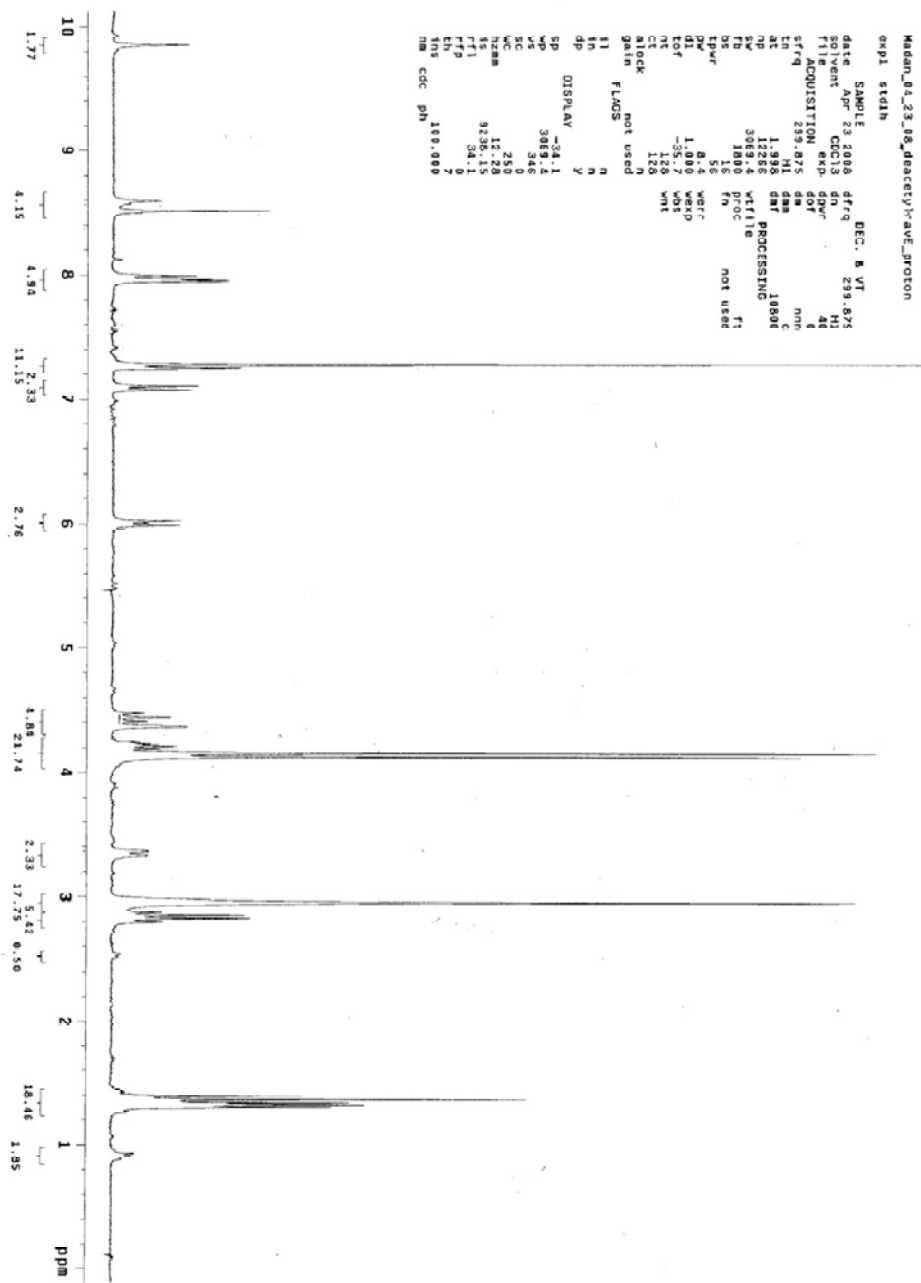
Figure A8. ^{1}H NMR spectrum of deacetylravidomycin E in CDCl$_3$(300 MHz, Varian)
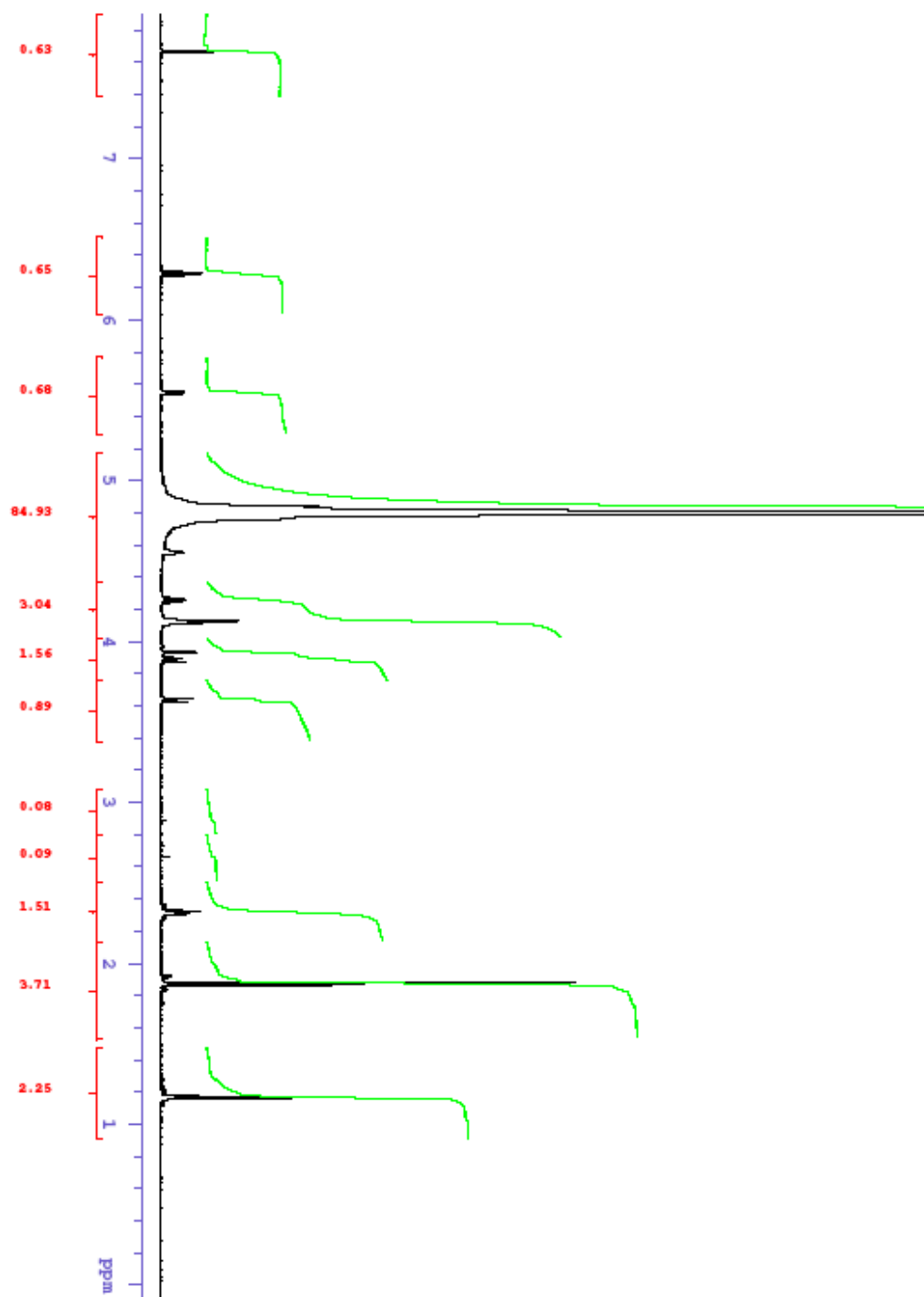
Figure A9. ^{1}H NMR spectrum of TDP-3-amino-3,6-dideoxy-D-galactose (**55**) in D$_2$O (500 MHz, Varian)
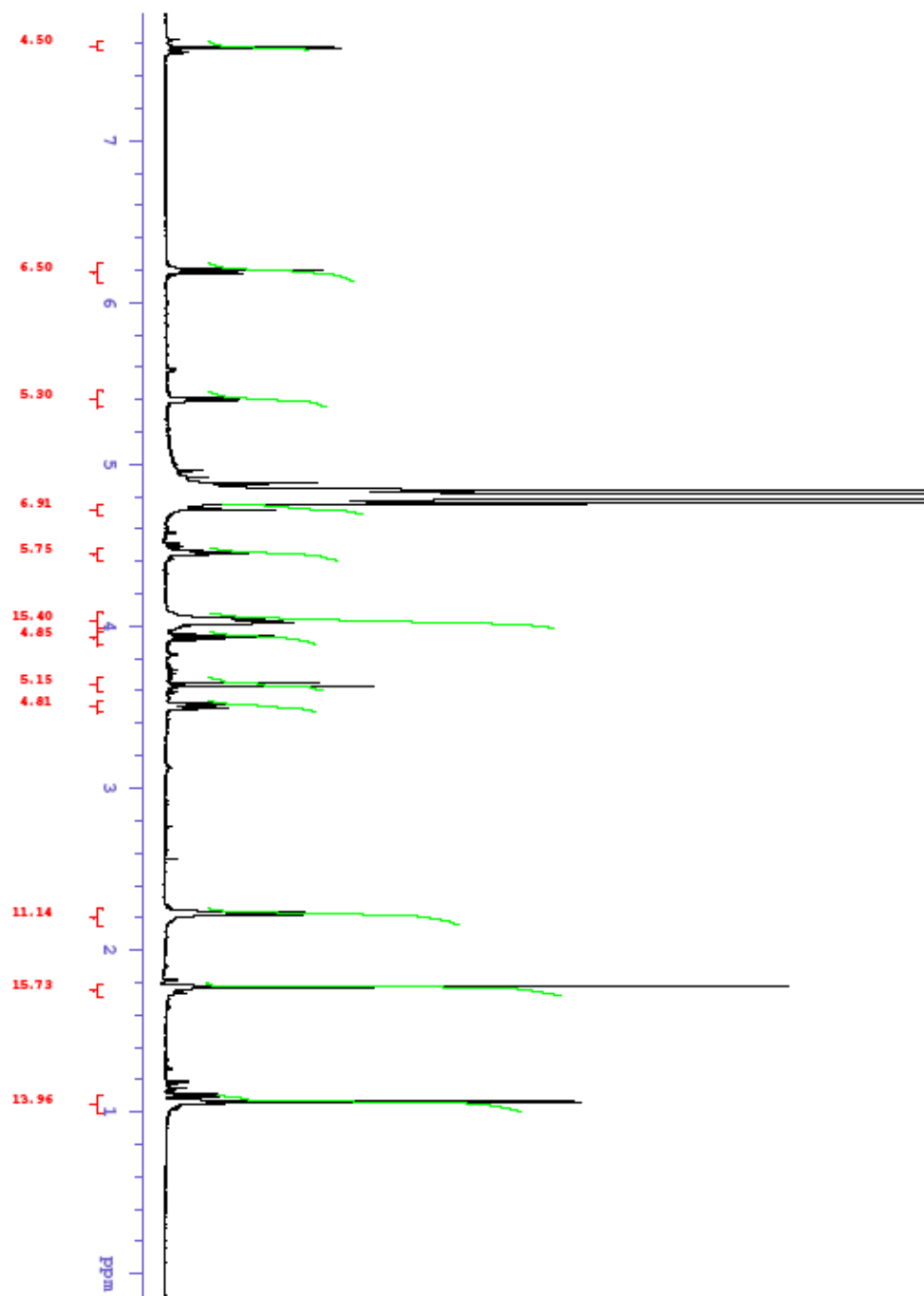
Figure A10. ^{1}H NMR spectrum of TDP-4-keto-6-deoxy-D-glucose (**18**) in D$_2$O (500 MHz, Varian)
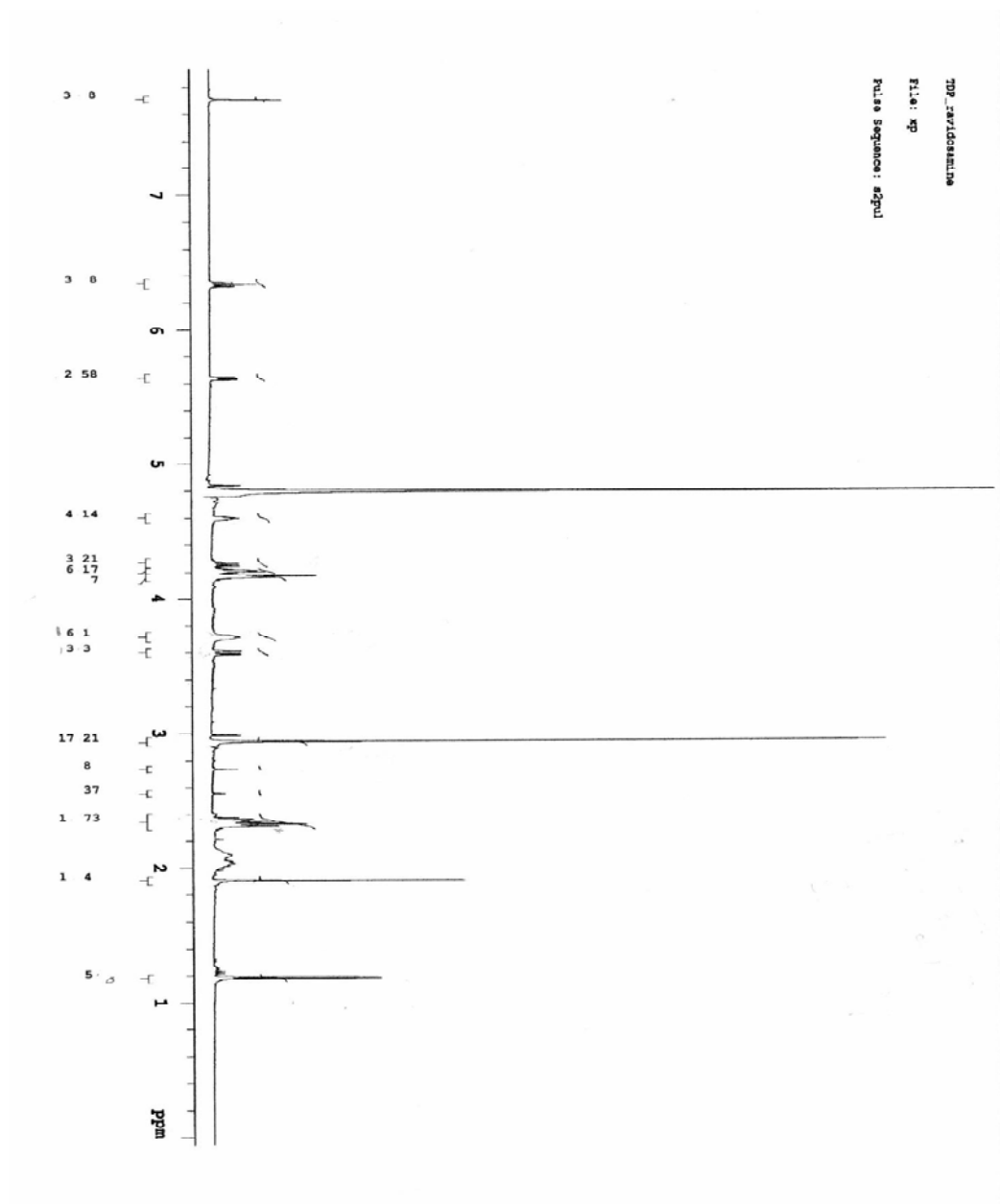
Figure A11. ^{1}H NMR spectrum of TDP-D-ravidosamine (**19**) in D_2O (500 MHz, Varian)

References

1.   Hertweck, C. (2009) The Biosynthetic Logic of Polyketide Diversity, *Angew. Chem. Int. Ed. 48*, 4688-4716.
2.   O'Hagan, D. (1991) *The Polyketides Metabolites, Ellis Horwood, Chichester.*
3.   Walsh, C. T. (2003) Antibiotics, *Antibiotics, ASM press, washington DC, USA*, 175-191.
4.   Smith, S., and Tsai, S. C. (2007) The type I fatty acid and polyketide synthases: a tale of two megasynthases, *Nat. Prod. Rep. 24*, 1041-1072.
5.   Hopwood, D. A. (1997) Genetic contributions to understanding polyketide synthases, *Chem. Rev. 97*, 2465-2497.
6.   Sattely, E. S., Fischbach, M. A., and Walsh, C. T. (2008) Total biosynthesis: in vitro reconstitution of polyketide and nonribosomal peptide pathways, *Nat. Prod. Rep. 25*, 757-793.
7.   Rawlings, B. J. (2001a) Type I polyketide biosynthesis in bacteria (part A-erythromycin biosynthesis), *Nat. Pord. Rep. 18*, 190-227.
8.   Donadio, S., Staver, M. J., McAlpine, J. B., Swanson, S. J., and Katz, L. (1991) Modular organization of genes required for complex polyketide biosynthesis, *Science 252*, 675-679.
9.   Cheng, Y. Q., Tang, G. L., and Shen, B. (2003) Type I polyketide synthase requiring a discrete acyltransferase for polyketide biosynthesis, *Proc. Natl. Acad. Sci. USA. 100*, 3149-3154.
10.   Piel, J. (2002) A polyketide synthase-peptide synthetase gene cluster from an uncultured bacterial symbiont of Paederus beetles, *Proc. Natl. Acad. Sci. USA. 99*, 14002-14007.
11.   Schumann, J. H., C. (2006) Advances in cloning, functional analysis and heterologous expression of fungal polyketide synthase genes, *J. Biotechnol. 124*, 690-703.
12.   Cox, R. J. (2007) Polyketides, proteins and genes in fungi: programmed nano-machines begin to reveal their secrets, *Org. Biomol. Chem. 5*, 2010-2026.
13.   McDaniel, R., Ebertkhosla, S., Hopwood, D. A., and Khosla, C. (1995) Rational design of aromatic polyketide natural products by recombinant assembly of enzymatic subunits, *Nature 375*, 549-554.
14.   Malpartida, F., and Hopwood, D. A. (1984) Molecular cloning of the whole biosynthetic pathway of a *Streptomyces* antibiotic and its expression in a heterologous host, *Nature 309*, 462-464.
15.   Brachmann, A. O., Joyce, S. A., Jenke-Kodarna, H., Schwar, G., Clarke, D. J., and Bode, H. B. (2007) A type II polyketide synthase is responsible for anthraquinone biosynthesis in *Photorhabdus luminescens*, *Chembiochem 8*, 1721-1728.
16.   Joyce, S. A., Brachmann, A. O., Glazer, I., Lango, L., Schwar, G., Clarke, D. J., and Bode, H. B. (2008) Bacterial biosynthesis of a multipotent stilbene, *Angew. Chem. Int. Ed. 47*, 1942-1945.
17.   Moore, B. S., and Hopke, J. N. (2001) Discovery of a new bacterial polyketide biosynthetic pathway, *Chembiochem 2*, 35-38.
18.   Pfeifer, V., Nicholson, G. J., Ries, J., Recktenwald, J., Schefer, A. B., Shawky, R. M., Schroder, J., Wohlleben, W., and Pelzer, S. (2001) A polyketide synthase in glycopeptide biosynthesis - The biosynthesis of the non-proteinogenic amino acid (S)-3,5-dihydroxyphenylglycine, *J. Biol. Chem. 276*, 38370-38377.
19.   Seshime, Y., Juvvadi, P. R., Fujii, I., and Kitamoto, K. (2005) Discovery of a novel superfamily of type III polyketide synthases in *Aspergillus oryzae*, *Biochem. Biophys. Res. Commun. 331*, 253-260.

20. Jacobsen, J. R., Hutchinson, C. R., Cane, D. E., and Khosla, C. (1997) Precursor-directed biosynthesis of erythromycin analogs by an engineered polyketide synthase, *Science 277*, 367-369.
21. Rix, U., Fischer, C., Remsing, L. L., and Rohr, J. (2002) Modification of post-PKS tailoring steps through combinatorial biosynthesis, *Nat. Prod. Rep. 19*, 542-580.
22. Perez, M., Baig, I., Brana, A. F., Salas, J. A., Rohr, J., and Mendez, C. (2008) Generation of new derivatives of the antitumor antibiotic mithramycin by altering the glycosylation pattern through combinatorial biosynthesis, *Chembiochem 9*, 2295-2304.
23. Remsing, L. L., Garcia-Bernardo, J., Gonzalez, A., Kunzel, E., Rix, U., Brana, A. F., Bearden, D. W., Mendez, C., Salas, J. A., and Rohr, J. (2002) Ketopremithramycins and ketomithramycins, four new aureolic acid-type compounds obtained upon inactivation of two genes involved in the biosynthesis of the deoxysugar moieties of the antitumor drug mithramycin by *Streptomyces argillaceus*, reveal novel insights into post-PKS tailoring steps of the mithramycin biosynthetic pathway, *J. Am. Chem. Soc. 124*, 1606-1614.
24. Trefzer, A., Blanco, G., Remsing, L., Kunzel, E., Rix, U., Lipata, F., Brana, A. F., Mendez, C., Rohr, J., Bechthold, A., and Salas, J. A. (2002) Rationally designed glycosylated premithramycins: hybrid aromatic polyketides using genes from three different biosynthetic pathways, *J. Am. Chem. Soc. 124*, 6056-6062.
25. Baig, I., Perez, M., Brana, A. F., Gomathinayagam, R., Damodaran, C., Salas, J. A., Mendez, C., and Rohr, J. (2008) Mithramycin analogues generated by combinatorial biosynthesis show improved bioactivity, *J. Nat. Prod. 71*, 199-207.
26. Remsing, L. L., Gonzalez, A. M., Nur-e-Alam, M., Fernandez-Lozano, M. J., Brana, A. F., Rix, U., Oliveira, M. A., Mendez, C., Salas, J. A., and Rohr, J. (2003) Mithramycin SK, a novel antitumor drug with improved therapeutic index, mithramycin SA, and demycarosyl-mithramycin SK: Three new products generated in the mithramycin producer *Streptomyces argillaceus* through combinatorial biosynthesis, *J. Am. Chem. Soc. 125*, 5745-5753.
27. Hoffmeister, D., Wilkinson, B., Foster, G., Sidebottom, P. J., Ichinose, K., and Bechthold, A. (2002) Engineered urdamycin glycosyltransferases are broadened and altered in substrate specificity, *Chem. Biol. 9*, 287-295.
28. Vilches, C., Hernandez, C., Mendez, C., and Salas, J. A. (1992) Role of glycosylation and deglycosylation in biosynthesis of and resistance to oleandomycin in the producer organism, *Streptomyces antibioticus*, *J. Bacteriol. 174*, 161-165.
29. Bolam, D. N., Roberts, S., Proctor, M. R., Turkenburg, J. P., Dodson, E. J., Martinez-Fleites, C., Yang, M., Davis, B. G., Davies, G. J., and Gilbert, H. J. (2007) The crystal structure of two macrolide glycosyltransferases provides a blueprint for host cell antibiotic immunity, *Proc. Natl. Acad. Sci. USA 104*, 5336-5341.
30. Chen, Y. H., Wang, C. C., Greenwell, L., Rix, U., Hoffmeister, D., Vining, L. C., Rohr, J. R., and Yang, K. Q. (2005) Functional analyses of oxygenases in jadomycin biosynthesis and identification of JadH as a bifunctional oxygenase/dehydrase, *J. Biol. Chem. 280*, 22508-22514.
31. Zhang, C. S., Albermann, C., Fu, X., and Thorson, J. S. (2006) The in vitro characterization of the iterative avermectin glycosyltransferase AveBI reveals reaction reversibility and sugar nucleotide flexibility, *J. Am. Chem. Soc. 128*, 16420-16421.
32. Durr, C., Hoffmeister, D., Wohlert, S. E., Ichinose, K., Weber, M., von Mulert, U., Thorson, J. S., and Bechthold, A. (2004) The glycosyltransferase UrdGT2 catalyzes both C- and O-glycosidic sugar transfers, *Angew. Chem. Int. Ed. 43*, 2962-2965.
33. Liu, T., Kharel, M. K., Fischer, C., McCormick, A., and Rohr, J. (2006) Inactivation of *gilGT*, encoding a C-glycosyltransferase, and *gilOlll*, encoding a P450 enzyme, allows the details of the late biosynthetic pathway to gilvocarcin V to be delineated, *Chembiochem 7*, 1070-1077.

34.	Blanco, G., Patallo, E. P., Brana, A. F., Trefzer, A., Bechthold, A., Rohr, J., Mendez, C., and Salas, J. A. (2001) Identification of a sugar flexible glycosyltransferase from *Streptomyces olivaceus*, the producer of the antitumor polyketide elloramycin, *Chem. Biol. 8*, 253-263.

35.	Fischer, C., Rodriguez, L., Patallo, E. P., Lipata, F., Brana, A. F., Mendez, C., Salas, J. A., and Rohr, J. (2002) Digitoxosyltetracenomycin C and glucosyltetracenomycin C, two novel elloramycin analogues obtained by exploring the sugar donor substrate specificity of glycosyltransferase ElmGT, *J. Nat. Prod. 65*, 1685-1689.

36.	Lombo, F., Gibson, M., Greenwell, L., Brana, A. F., Rohr, J., Salas, J. A., and Mendez, C. (2004) Engineering biosynthetic pathways for deoxysugars: Branched-chain sugar pathways and derivatives from the antitumor tetracenomycin, *Chem. Biol. 11*, 1709-1718.

37.	Perez, M., Lombo, F., Zhu, L. L., Gibson, M., Brana, A. F., Rohr, R., Salas, J. A., and Mendez, C. (2005) Combining sugar biosynthesis genes for the generation of L- and D-amicetose and formation of two novel antitumor tetracenomycins, *Chem.Commun.*, 1604-1606.

38.	Rodriguez, L., Aguirrezabalaga, I., Allende, N., Brana, A. F., Mendez, C., and Salas, J. A. (2002) Engineering deoxysugar biosynthetic pathways from antibiotic-producing microorganisms: A tool to produce novel glycosylated bioactive compounds, *Chem. Biol. 9*, 721-729.

39.	Luzhetskyy, A., Mayer, A., Hoffmann, J., Pelzer, S., Holzenkamper, M., Schmitt, B., Wohlert, S. E., Vente, A., and Bechthold, A. (2007) Cloning and heterologous expression of the aranciamycin biosynthetic gene cluster revealed a new flexible glycosyltransferase, *Chembiochem 8*, 599-602.

40.	Olano, C., Abdelfattah, M. S., Gullon, S., Brana, A. F., Rohr, J., Mendez, C., and Salas, J. A. (2008) Glycosylated derivatives of steffimycin: Insights into the role of the sugar moieties for the biological activity, *Chembiochem 9*, 624-633.

41.	Williams, G. J., Goff, R. D., Zhang, C. S., and Thorson, J. S. (2008) Optimizing glycosyltransferase specificity via "Hot spot" saturation mutagenesis presents a catalyst for novobiocin glycorandomization, *Chem. Biol. 15*, 393-401.

42.	Griffith, B. R., Langenhan, J. M., and Thorson, J. S. (2005) 'Sweetening' natural products via glycorandomization, *Curr. Opin. Biotechnol. 16*, 622-630.

43.	Langenhan, J. M., Griffith, B. R., and Thorson, J. S. (2005) Neoglycorandomization and chemoenzymatic glycorandomization: Two complementary tools for natural product diversification, *J. Nat. Prod. 68*, 1696-1711.

44.	Baig, I., Kharel, M., Kobylyanskyy, A., Zhu, L. L., Rebets, Y., Ostash, B., Luzhetskyy, A., Bechthold, A., Fedorenko, V. A., and Rohr, J. (2006) On the acceptor substrate of C-Glycosyltransferase UrdGT2: three prejadomycin C-Glycosides from an engineered mutant of *Streptomyces globisporus* 1912 delta lndE(urdGT2, *Angew. Chem. Int. Ed. 45*, 7842-7846.

45.	Thibodeaux, C. J., Melancon, C. E., and Liu, H. W. (2008) Natural-product sugar biosynthesis and enzymatic glycodiversification, *Angew. Chem.-Int. Ed. 47*, 9814-9859.

46.	Bililign, T., Shepard, E. M., Ahlert, J., and Thorson, J. S. (2002) On the origin of deoxypentoses: evidence to support a glucose progenitor in the biosynthesis of calicheamicin, *Chembiochem 4*, 1143-1146.

47.	Hofmann, C., Boll, R., Heitmann, B., Hauser, G., Durr, C., Frerich, A., Weitnauer, G., Glaser, S. J., and Bechthold, A. (2005) Genes encoding enzymes responsible for biosynthesis of L-lyxose and attachment of eurekanate during avilamycin biosynthesis, *Chem. Biol. 12*, 1137-1143.

48.	Aparicio, J. F., Caffrey, P., Gil, J. A., and Zotchev, S. B. (2003) Polyene antibiotic biosynthesis gene clusters, *Appl. Microbiol. Biotechnol. 61*, 179-188.

49. He, X. M. M., and Liu, H. W. (2002) Formation of unusual sugars: Mechanistic studies and biosynthetic applications, *Annu. Rev. Biochem. 71*, 701-754.

50. Xue, Y. Q., Wilson, D., Zhao, L. S., Liu, H. W., and Sherman, D. H. (1998) Hydroxylation of macrolactones YC-17 and narbomycin is mediated by the *pikC*-encoded cytochrome P450 in *Streptomyces venezuelae*, *Chem. Biol. 5*, 661-667.

51. Lee, S. K., Park, J. W., Kim, J. W., Jung, W. S., Park, S. R., Choi, C. Y., Kim, E. S., Kim, B. S., Ahn, J. S., Sherman, D. H., and Yoon, Y. J. (2006) Neopikromycin and novapikromycin from the pikromycin biosynthetic pathway of *Streptomyces venezuelae*, *J. Nat. Prod. 69*, 847-849.

52. Yoon, Y. J., Beck, B. J., Kim, B. S., Kang, H. Y., Reynolds, K. A., and Sherman, D. H. (2002) Generation of multiple bioactive macrolides by hybrid modular polyketide synthases in *Streptomyces venezuelae*, *Chem. Biol. 9*, 203-214.

53. Desai, R. P., Rodriguez, E., Galazzo, J. L., and Licari, P. (2004) Improved bioconversion of 15-fluoro-6-deoxyerythronolide B to 15-fluoro-erythromycin A by overexpression of the *eryK* gene in *Saccharopolyspora erythraea*, *Biotechnol. Prog. 20*, 1660-1665.

54. Gibson, M., Nur-e-alam, M., Lipata, F., Oliveira, M. A., and Rohr, J. (2005) Characterization of kinetics and products of the Baeyer-Villiger oxygenase MtmOIV, the key enzyme of the biosynthetic pathway toward the natural product anticancer drug mithramycin from *Streptomyces argillaceus*, *J. Am. Chem. Soc. 127*, 17594-17595.

55. Menendez, N., Nur-e-Alam, M., Brana, A. F., Rohr, J., Salas, A. F., and Mendez, C. (2004) Biosynthesis of the antitumor chromomycin A(3) in *Streptomyces griseus*: Analysis of the gene cluster and rational design of novel chromomycin analogs, *Chem. Biol. 11*, 21-32.

56. Kharel, M. K., Zhu, L. L., Liu, T., and Rohr, J. (2007) Multi-oxygenase complexes of the gilvocarcin and jadomycin biosyntheses, *J. Am. Chem. Soc. 129*, 3780-3781.

57. Liu, T., Fischer, C., Beninga, C., and Rohr, J. (2004) Oxidative rearrangement processes in the biosynthesis of gilvocarcin V, *J. Am. Chem. Soc. 126*, 12262-12263.

58. Beam, M. P., Bosserman, M. A., Noinaj, N., Wehenkel, M., and Rohr, J. (2009) Crystal structure of Baeyer-Villiger monooxygenase MtmOIV, the key enzyme of the mithramycin biosynthetic pathway, *Biochemistry 48*, 4476-4487.

59. Kallio, P., Liu, Z. L., Mantsala, P., Niemi, J., and Metsa-Ketela, M. (2008) Sequential action of two flavoenzymes, PgaE and PgaM, in angucycline biosynthesis: chemoenzymatic synthesis of gaudimycin C, *Chem. Biol. 15*, 157-166.

60. Chung, J. Y., Fujii, I., Harada, S., Sankawa, U., and Ebizuka, Y. (2002) Expression, purification, and characterization of AknX anthrone oxygenase, which is involved in aklavinone biosynthesis in *Streptomyces galilaeus*, *J. Bacteriol. 184*, 6115-6122.

61. Rix, U., Wang, C. C., Chen, Y. H., Lipata, F. M., Rix, L. L. R., Greenwell, L. M., Vining, L. C., Yang, K. Q., and Rohr, J. (2005) The oxidative ring cleavage in jadomycin biosynthesis: a multistep oxygenation cascade in a biosynthetic black box, *Chembiochem 6*, 838-835.

62. Connors, N. C., and Strohl, W. R. (1993) Partial purification and properties of carminomycin 4-O-methyltransferase from *Streptomyces* sp. strain C5, *J. Gen. Microbiol. 139*, 1353-1362.

63. Madduri, K., Torti, F., Colombo, A. L., and Hutchinson, C. R. (1993) Cloning and sequencing of a gene encoding carminomycin 4-O-methyltransferase from *Streptomyces peucetius* and its expression in *Escherichia coli*, *J. Bacteriol. 175*, 3900-3904.

64. Chen, Z. G., Fujii, I., Ebizuka, Y., and Sankawa, U. (1992) Emodin O-methyltransferase from *Aspergillus terreus*, *Arch. Microbiol. 158*, 29-34.

65. Summers, R. G., Wendtpienkowski, E., Motamedi, H., and Hutchinson, C. R. (1992) Nucleotide sequence of the *tcmII-tcmIV* region of the tetracenomycin C biosynthetic gene

cluster of *Streptomyces glaucescens* and evidence that the *tcmN* gene encodes a multifunctional cyclase-dehydratase-O-methyl transferase, *J. Bacteriol. 174*, 1810-1820.

66. Dickens, M. L., Priestley, N. D., and Strohl, W. R. (1997) In vivo and in vitro bioconversion of epsilon-rhodomycinone glycoside to doxorubicin: functions of DauP, DauK, and DoxA, *J. Bacteriol. 179*, 2641-2650.

67. Doumith, M., Weingarten, P., Wehmeier, U. F., Salah-Bey, K., Benhamou, B., Capdevila, C., Michel, J. M., Piepersberg, W., and Raynal, M. C. (2000) Analysis of genes involved in 6-deoxyhexose biosynthesis and transfer in *Saccharopolyspora erythraea*, *Mol. Gen. Genet. 264*, 477-485.

68. Chang, C. W., Zhao, L. H., Yamase, H., and Liu, H. W. (2000) DesVI: A new member of the sugar N,N-dimethyltransferase family involved in the biosynthesis of desosamine, *Angew. Chem. Int. Ed. Engl. 39*, 2160-2163.

69. Shen, B., and Hutchinson, C. R. (1996) Deciphering the mechanism for the assembly of aromatic polyketides by a bacterial polyketide synthase, *Proc. Natl. Acad. Sci. USA. 93*, 6600-6604.

70. Krohn, K., and Rohr, J. (1997) Angucyclines: Total syntheses, new structures, and biosynthetic studies of an emerging new class of antibiotics, *Bioorganic Chemistry Deoxysugars, Polyketides and Related Classes: Synthesis, Biosynthesis, Enzymes 188*, 127-195.

71. Fischer, C., Lipata, F., and Rohr, J. (2003) The complete gene cluster of the antitumor agent gilvocarcin V and its implication for the biosynthesis of the gilvocarcins, *J. Am. Chem. Soc. 125*, 7818-7819.

72. Nakano, H., Matsuda, Y., Ito, K., Ohkubo, S., Morimoto, M., and Tomita, F. (1981) Gilvocarcins, new antitumor antibiotics. 1.Taxonomy, fermentation, isolation and biological activities., *J. Antibiot. 34*, 266-270.

73. Takahashi, K. Y., M.; Tomita, F.; Shirahata, K. (1981) Gilvocarcins, new antitumor antibiotics. 2. Structural elucidation, *J. Antibiot. 43*, 271-275.

74. Morimoto, M. O., S.; Tomita, F.; Marumo, H.;. (1981) Gilvocarcins, new antitumor antibiotics. 3. Antitumor activity, *J. Antibiot. 34*, 701-707.

75. Lytle, C. D., Routson, L. B., and Prodouz, K. N. (1994) Herpes virus infection and repair in cells pretreated with gilvocarcin V or merocyanine 540 and radiation, *J. Photochem. Photobiol. B-Biology 23*, 57-62.

76. Lytle, C. D., Wagner, S. J., and Prodouz, K. N. (1993) Antiviral activity of gilvocarcin V plus UVA radiation, *Photochem. Photobiol. 58*, 818-821.

77. Oyola, R., Arce, R., Alegria, A. E., and Garcia, C. (1997) Photophysical properties of gilvocarcins V and M and their binding constant to calf thymus DNA, *Photochem. Photobiol. 65*, 802-810.

78. Knobler, R. M., Radlwimmer, F. B., and Lane, M. J. (1993) Gilvocarcin V exhibits both equilibrium DNA binding and UV light induced DNA adduct formation which is sequence context dependent (Vol. 20, pp. 4553, 1992), *Nucleic Acids Res. 21*, 3920-3920.

79. Knobler, R. M., Radlwimmer, F. B., and Lane, M. J. (1992) Gilvocarcin V exhibits both equilibrium DNA binding and UV light induced DNA adduct formation which is sequence context dependent, *Nucleic Acids Res. 20*, 4553-4557.

80. Matsumoto, A., and Hanawalt, P. C. (2000) Histone H3 and heat shock protein GRP78 are selectively cross-linked to DNA by photoactivated gilvocarcin V in human fibroblasts, *Cancer Res. 60*, 3921-3926.

81. Matsumoto, T., Hosoya, T., and Suzuki, K. (1992) Total synthesis and absolute stereochemical assignment of gilvocarcin M, *J. Am. Chem. Soc. 114*, 3568-3570.

82. Farr, R. N., Kwok, D. I., and Daves, G. D. (1992) 8-Ethenyl-l-hydroxy-4-beta-D-ribofuranosylbenzo[*d*]naptho[1,2-*b*]pyran-6-one and 8-Ethenyl-1-hydroxy-4-(2'-deoxy-beta-D-ribofuranosyl)benzo[*d*]naptho[1,2-*b*-pyran]-6-one. synthetic C-glycosides related

to the gilvocarcin, ravidomycin, and chrysomycin antibiotics, *J. Org. Chem. 57*, 2093-2100.

83. Cordero-Vargas, A., Quiclet-Sire, B., and Zard, S. Z. (2005) Model studies towards the synthesis of gilvocarcin M, *Org. Biomol. Chem. 3*, 4432-4443.

84. Hosoya, T., Takashiro, E., Matsumoto, T., and Suzuki, K. (1994) Total synthesis of the gilvocarcins, *J. Am. Chem. Soc. 116*, 1004-1015.

85. Yamashita, N., Shin-Ya, K., Furihata, K., Hayakawa, Y., and Seto, H. (1998) New ravidomycin analogues, FE35A and FE35B, apoptosis inducers produced by Streptomyces rochei, *J. Antibiot. 51*, 1105-1108.

86. Kojiri, K., Arakawa, H., Satoh, F., Kawamura, K., Okura, A., Suda, H., and Okanishi, M. (1991) New antitumor substances, BE-12406A and BE-12406B, produced by a streptomycete. I. Taxonomy, fermentation, isolation, physico-chemical and biological properties, *J. Antibiot. 44*, 1054-1060.

87. Nakajima, S., Kojiri, K., Suda, H., and Okanishi, M. (1991) New antitumor substances, BE-12406A and BE-12406B, produced by a *streptomycete*. II. Structure determination, *J. Antibiot. 44*, 1061-1064.

88. Nakashima, T., Tadashi, F., Kazuya, S., Tomohiro, S., Hiroyuki, K., and Takeo, Y. (1999) Chrysomycin derivative compounds and use as antitumor agents, p 9, Mercian Corporation, *United States Patent 09/308710*

89. Li, Y. Q., Huang, X. S., Ishida, K., Maier, A., Kelter, G., Jiang, Y., Peschel, G., Menzel, K. D., Li, M. G., Wen, M. L., Xu, L. H., Grabley, S., Fiebig, H. H., Jiang, C. L., Hertweck, C., and Sattler, I. (2008) Plasticity in gilvocarcin-type C-glycoside pathways: discovery and antitumoral evaluation of polycarcin V from *Streptomyces polyformus*, *Org. Biomol. Chem. 6*, 3601-3605.

90. Singh, K. (1984) Studies on the mechanism of action of ravidomycin (AY-25,545), *J. Antibiot. 37*, 71-73.

91. Rakhit, S., Eng, C., Baker, H., and Singh, K. (1983) Chemical modification of ravidomycin and evaluation of biological activities of its derivatives, *J. Antibiot. 36*, 1490-1494.

92. Greenstein, M., Monji, T., Yeung, R., Maiese, W. M., and White, R. J. (1986) Light-dependent activity of the antitumor antibiotics ravidomycin and desacetylravidomycin, *Antimicrob. Agents Chemother. 29*, 861-866.

93. Rakhit, S., Eng, C., Baker, H., and Singh, K. (1983) Chemical modification of ravidomycin and evaluation of biological activities of its derivatives, *J. Antibiot. 36*, 1490-1494.

94. Lorico, A., and Long, B. H. (1993) Biochemical characterization of elsamicin and other coumarin-related antitumor agents as potent inhibitors of human topoisomerase II, *Eur. J. Cancer 29A*, 1985-1991.

95. Chakrabarti, A. C., Clarklewis, I., Harrigan, P. R., and Cullis, P. R. (1992) Uptake of basic amino acids and peptides into liposomes in response to transmembrane pH gradients, *Biophys. J. 61*, 228-234.

96. Carter, G. T., Fantini, A. A., James, J. C., Borders, D. B., and White, R. J. (1984) Biosynthesis of ravidomycin- use of ^{13}C-^{13}C double quantum NMR to follow precursor incorporation, *Tetrahedron Lett. 25*, 255-258.

97. Takahashi, K., and Tomita, F. (1983) Gilvocarcins, new antitumor antibiotics. 5. Biosynthesis of gilvocarcins: Incorporation of ^{13}C-labeled compounds into gilvocarcin aglycones, *J. Antibiot. 36*, 1531-1535.

98. Carter, G. T., Fantini, A. A., James, J. C., Borders, D. B., and White, R. J. (1985) Biosynthesis of chrysomycin-A and chrysomycin-B origin of the chromophore, *J. Antibiot. 38*, 242-248.

99.   Syvitski, R. T., Borissow, C. N., Graham, C. L., and Jakeman, D. L. (2006) Ring-opening dynamics of jadomycin A and B and dalomycin T, *Org. Lett. 8*, 697-700.

100.   Rix, U., Zheng, J. T., Rix, L. L. R., Greenwell, L., Yang, K. Q., and Rohr, J. (2004) The dynamic structure of Jadomycin B and the amino acid incorporation step of its biosynthesis, *J. Am. Chem. Soc. 126*, 4496-4497.

101.   Liu, T. (2006) Generating novel derivatives of the anticancer agent gilvocarcin by combinatorial biosynthesis, *dissertation, University of Kentucky, College of Pharmacy*.

102.   Liu, T., Kharel, M. K., Zhu, L., Bright, S., and Rohr, J. (2009) Inactivation of the ketoreductase *gilU* gene of the gilvocarcin biosynthetic gene cluster yields new analogues with partly improved biological activity., *Chembiochem 10*, 278-286.

103.   Ichinose, K., Ozawa, M., Itou, K., Kunieda, K., and Ebizuka, Y. (2003) Cloning, sequencing and heterologous expression of the medermycin biosynthetic gene cluster of Streptomyces sp. AM-7161: towards comparative analysis of the benzoisochromanequinone gene clusters, *Microbiol.-SGM 149*, 1633-1645.

104.   Bierman, M., Logan, R., Obrien, K., Seno, E. T., Rao, R. N., and Schoner, B. E. (1992) Plasmid Cloning vectors for the conjugal transfer of DNA from *E. coli* to *Streptomyces* spp., *Gene 116*, 43-49.

105.   Decker, H., Gaisser, S., Pelzer, S., Schneider, P., Westrich, L., Wohlleben, W., and Bechthold, A. (1996) A general approach for cloning and characterizing dNDP-glucose dehydratase genes from actinomycetes, *FEMS Microbiol. Lett. 141*, 195-201.

106.   Decker, H., and Haag, S. (1995) Cloning and characterization of a polyketide synthase gene from *Streptomyces fradiae* Tu2717, which carries the genes for biosynthesis of the angucycline antibiotic urdamycin A and a gene probably involved in its oxygenation, *J. Bacteriol. 177*, 6126-6136.

107.   Han, L., Yang, K. Q., Ramalingam, E., Mosher, R. H., and Vining, L. C. (1994) Cloning and characterization of polyketide synthase genes for jadomycin B biosynthesis in *Streptomyces venezuelae* IS P52 30, *Microbiol-UK 140*, 3379-3389.

108.   Tang, Y., Lee, T. S., Kobayashi, S., and Khosla, C. (2003) Ketosynthases in the initiation and elongation modules of aromatic polyketide synthases have orthogonal acyl carrier protein specificity, *Biochemistry 42*, 6588-6595.

109.   Jiralerspong, S., Rangaswamy, V., Bender, C. L., and Parry, R. J. (2001) Analysis of the enzymatic domains in the modular portion of the coronafacic acid polyketide synthase, *Gene 270*, 191-200.

110.   Raty, K., Kantola, J., Hautala, A., Hakala, J., Ylihonko, K., and Mantsala, P. (2002) Cloning and characterization of Streptomyces galilaeus aclacinomycins polyketide synthase (PKS) cluster, *Gene 293*, 115-122.

111.   Hutchinson, C. R. (1997) Biosynthetic studies of daunorubicin and tetracenomycin C, *Chem. Rev. 97*, 2525-2535.

112.   Trefzer, A., Pelzer, S., Schimana, J., Stockert, S., Bihlmaier, C., Fiedler, H. P., Welzel, K., Vente, A., and Bechthold, A. (2002) Biosynthetic gene cluster of simocyclinone, a natural multihybrid antibiotic, *Antimicrob.Agents Chemother. 46*, 1174-1182.

113.   Tsukamoto, N., Fujii, I., Ebizuka, Y., and Sankawa, U. (1994) Nucleotide-sequence of the *aknA* region of the aklavinone biosynthetic gene cluster of *Streptomyces galilaeus*, *J. Bacteriol. 176*, 2473-2475.

114.   Basnet, D. B., Oh, T. J., Vu, T. T. H., Sthapit, B., Liou, K., Lee, H. C., Yoo, J. C., and Sohng, J. K. (2006) Angucyclines Sch 47554 and Sch 47555 from *Streptomyces* sp SCC-2136: cloning, sequencing, and characterization, *Mol. Cells 22*, 154-162.

115.   Lombo, F., Brana, A. F., Salas, J. A., and Mendez, C. (2004) Genetic organization of the biosynthetic gene cluster for the antitumor angucycline oviedomycin in *Streptomyces antibioticus* ATCC 11891, *Chembiochem 5*, 1181-1187.

116. Yu, T. W., Bibb, M. J., Revill, W. P., and Hopwood, D. A. (1994) Cloning, sequencing, and analysis of the griseusin polyketide synthase gene-cluster from *Streptomyces griseus*, *J. Bacteriol. 176*, 2627-2634.

117. Summers, R. G., Ali, A., Shen, B., Wessel, W. A., and Hutchinson, C. R. (1995) Malonyl-coenzyme A:acyl carrier protein acyltransferase of *Streptomyces glaucescens*: a possible link between fatty acid and polyketide biosynthesis, *Biochemistry 34*, 9389-9402.

118. Metsa-Ketela, M., Ylihonko, K., and Mantsala, P. (2004) Partial activation of a silent angucycline-type gene cluster from a rubromycin beta producing Streptomyces sp PGA64, *J. Antibiot. 57*, 502-510.

119. Novakova, R., Bistakova, J., Homerova, D., Rezuchova, B., and Kormanec, J. (2002) Cloning and characterization of a polyketide synthase gene cluster involved in biosynthesis of a proposed angucycline-like polyketide auricin in *Streptomyces aureofaciens* CCM 3239, *Gene 297*, 197-208.

120. Bibb, M. J., Sherman, D. H., Omura, S., and Hopwood, D. A. (1994) Cloning, sequencing and deduced functions of a cluster of Streptomyces genes probably encoding biosynthesis of the polyketide antibiotic frenolicin, *Gene 142*, 31-39.

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

122. Grimm, A., Madduri, K., Ali, A., and Hutchinson, C. R. (1994) Characterization of the *Streptomyces peucetius* ATCC 29050 genes encoding doxorubicin polyketide synthase, *Gene 151*, 1-10.

123. Lombo, F., Abdelfattah, M. S., Brana, A. F., Salas, J. A., Rohr, J., and Mendez, C. (2009) Elucidation of oxygenation steps during oviedomycin biosynthesis and generation of derivatives with increased antitumor activity, *Chembiochem 10*, 296-303.

124. Trefzer, A., Pelzer, S., Schimana, J., Stockert, S., Bihlmaier, C., Fiedler, H. P., Welzel, K., Vente, A., and Bechthold, A. (2002) Biosynthetic gene cluster of simocyclinone, a natural multihybrid antibiotic, *Antimicrob. Agents Chemother. 46*, 1174-1182.

125. Palmu, K., Ishida, K., Mantsala, P., Hertweck, C., and Metsa-Ketela, M. (2007) Artificial reconstruction of two cryptic anglucycline antibiotic Biosynthetic pathways, *Chembiochem 8*, 1577-1584.

126. Kharel, M. K., Pahari, P., Lian, H., and Rohr, J. (2009) GilR, an unusual lactone-forming enzyme involved in gilvocarcin biosynthesis, *Chembiochem 10*, 1305-1308.

127. Lombo, F., Siems, K., Brana, A. F., Mendez, C., Bindseil, K., and Salas, J. A. (1997) Cloning and insertional inactivation of *Streptomyces argillaceus* genes involved in the earliest steps of biosynthesis of the sugar moieties of the antitumor polyketide mithramycin, *J. Bacteriol. 179*, 3354-3357.

128. Distler, J., A. Ebert, K. Mansouri, K. Pissowotzki, M. Stockmann, and W., and Piepersberg. (1987) Gene cluster for streptomycin biosynthesis in *Streptomyces griseus*: nucleotide sequence of three genes and analysis of transcriptional activity., *Nucleic Acids Res. 15*, 8041-8056.

129. Vorholter, F. J., Niehaus, K., and Puhler, A. (2001) Lipopolysaccharide biosynthesis in *Xanthomonas campestris* pv. *campestris*: a cluster of 15 genes is involved in the biosynthesis of the LPS O-antigen and the LPS core, *Mol. Genet. Genomics 266*, 79-95.

130. Karray, F., Darbon, E., Oestreicher, N., Dominguez, H., Tuphile, K., Gagnat, J., Blondelet-Rouault, M. H., and Pernodet, C. (2007) Organization of the biosynthetic gene cluster for the macrolide antibiotic spiramycin in Streptomyces ambofaciens, *Microbiol-Sgm 153*, 4111-4122.

131. Melancon, C. E., Hong, L., White, J. A., Liu, Y. N., and Liu, H. W. (2007) Characterization of TDP-4-keto-6-deoxy-D-glucose-3,4-ketoisomerase from the D-mycaminose biosynthetic pathway of *Streptomyces fradiae*: In vitro activity and substrate specificity studies, *Biochemistry 46*, 577-590.

132. Pfoestl, A., Hofinger, A., Kosma, P., and Messner, P. (2003) Biosynthesis of dTDP-3-acetamido-3,6-dideoxy-alpha-D-galactose in *Aneurinibacillus thermoaerophilus* L420-91(T), *J. Biol. Chem. 278*, 26410-26417.

133. Dhillon, N., Hale, R. S., Cortes, J., and Leadlay, P. F. (1989) Molecular characterization of a gene from *Saccharopolyspora erythraea* (*Streptomyces erythraeus*) which is involved in erythromycin biosynthesis, *Mol. Microbiol. 3*, 1405-1414.

134. Quiros, L. M., Aguirrezabalaga, I., Olano, C., Mendez, C., and Salas, J. A. (1998) Two glycosyltransferases and a glycosidase are involved in oleandomycin modification during its biosynthesis by *Streptomyces antibioticus*, *Mol. Microbiol. 28*, 1177-1185.

135. Zhao, L. S., Que, N. L. S., Xue, Y. Q., Sherman, D. H., and Liu, H. W. (1998) Mechanistic studies of desosamine biosynthesis: C-4 deoxygenation precedes C-3 transamination, *J. Am. Chem. Soc. 120*, 12159-12160.

136. Olano, C., Rodriguez, A. M., Michel, J. M., Mendez, C., Raynal, M. C., and Salas, J. A. (1998) Analysis of a *Streptomyces antibioticus* chromosomal region involved in oleandomycin biosynthesis, which encodes two glycosyltransferases responsible for glycosylation of the macrolactone ring, *Mol. Gen. Genet. 259*, 299-308.

137. Torkkell, S., Kunnari, T., Palmu, K., Mantsala, P., Hakala, J., and Ylihonko, K. (2001) The entire nogalamycin biosynthetic gene cluster of *Streptomyces nogalater*: characterization of a 20-kb DNA region and generation of hybrid structures, *Mol. Genet. Genomics 266*, 276-288.

138. Gu, L. C., Geders, T. W., Wang, B., Gerwick, W. H., Hakansson, K., Smith, J. L., and Sherman, D. H. (2007) GNAT-like strategy for polyketide chain initiation, *Science 318*, 970-974.

139. Luzhetskyy, A., Taguchi, T., Fedoryshyn, M., Durr, C., Wohlert, S. E., Novikov, V., and Bechthold, A. (2005) LanGT2 catalyzes the first glycosylation step during landomycin a biosynthesis, *Chembiochem 6*, 1406-1410.

140. Kunzel, E., Faust, B., Oelkers, C., Weissbach, U., Bearden, D. W., Weitnauer, G., Westrich, L., Bechthold, A., and Rohr, J. (1999) Inactivation of the urdGT2 gene, which encodes a glycosyltransferase responsible for the C-glycosyltransfer of activated D-olivose, leads to formation of the novel urdamycins I, J, and K, *J. Am. Chem. Soc. 121*, 11058-11062.

141. Fischbach, M. A., Lin, H. N., Liu, D. R., and Walsh, C. T. (2005) In vitro characterization of IroB, a pathogen-associated C-glycosyltransferase, *Proc. Nati. Acad. Sci. U.S.A. 102*, 571-576.

142. Erb, A., Luzhetskyy, A., Hardter, U., and Bechthold, A. (2009) Cloning and sequencing of the biosynthetic gene cluster for saquayamycin Z and galtamycin B and the elucidation of the assembly of their saccharide chains, *Chembiochem 10*, 1392-1401.

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

144. Sambrook, J., and Russell, D. W. (2001) *Molecular Cloning: a laboratory manual Cold Spring Harbor Laboratory Press, New York.*

145. Macneil, D. J., Occi, J. L., Gewain, K. M., Macneil, T., Gibbons, P. H., Ruby, C. L., and Danis, S. J. (1992) Complex organization of the *Streptomyces avermitilis* genes encoding the avermectin polyketide synthase, *Gene 115*, 119-125.

146. Johnson, D. A. L., H.-W. (1999) *Occurrence, genetics, and mechanisms of biosynthesis, in Comprehensive Natural Products Chemistry (Barton, D. H. R., and Nakanishi, K., Eds.) Vol. 3, pp. 311-365, Elsevier, Amsterdam.*

147. Kren, V., and Martinkova, L. (2001) Glycosides in medicine: "The role of glycosidic residue in biological activity", *Curr. Med. Chem. 8*, 1303-1328.

148. WeymouthWilson, A. C. (1997) The role of carbohydrates in biologically active natural products, *Nat. Prod. Rep. 14*, 99-110.

149. Thibodeaux, C. J., Melancon, C. E., and Liu, H. W. (2007) Unusual sugar biosynthesis and natural product glycodiversification, *Nature 446*, 1008-1016.

150. Trefzer, A., Salas, J. A., and Bechthold, A. (1999) Genes and enzymes involved in deoxysugar biosynthesis in bacteria, *Nat. Prod. Rep. 16*, 283-299.

151. Blanchard, S., and Thorson, J. S. (2006) Enzymatic tools for engineering natural product glycosylation, *Curr. Opin. Chem. Biol. 10*, 263-271.

152. Gantt, R. W., Goff, R. D., Williams, G. J., and Thorson, J. S. (2008) Probing the Aglycon Promiscuity of an Engineered Glycosyltransferase, *Angew.Chem. Int. Ed. 47*, 8889-8892.

153. Zhang, C. S., Fu, Q., Albermann, C., Li, L. J., and Thorson, J. S. (2007) The in vitro characterization of the erythronolide mycarosyltransferase EryBV and its utility in macrolide diversification, *Chembiochem 8*, 385-390.

154. Losey, H. C., Jiang, J. Q., Biggins, J. B., Oberthur, M., Ye, X. Y., Dong, S. D., Kahne, D., Thorson, J. S., and Walsh, C. T. (2002) Incorporation of glucose analogs by GtfE and GtfD from the vancomycin biosynthetic pathway to generate variant glycopeptides, *Chem. Biol. 9*, 1305-1314.

155. Park, S. H., Park, H. Y., Sohng, J. K., Lee, H. C., Liou, K., Yoon, Y. J., and Kim, B. G. (2009) Expanding substrate specificity of GT-B fold glycosyltransferase via domain swapping and high-throughput screening, *Biotechnol. Bioeng. 102*, 988-994.

156. Zhang, C. S., Moretti, R., Jiang, J. Q., and Thorson, J. S. (2008) The in vitro characterization of polyene glycosyltransferases AmphDI and NysDI, *Chembiochem 9*, 2506-2514.

157. Takahashi, H., Liu, Y. N., and Liu, H. W. (2006) A two-stage one-pot enzymatic synthesis of TDP-L-mycarose from thymidine and glucose-1-phosphate, *J. Am. Chem. Soc. 128*, 1432-1433.

158. Findlay, J. A., Liu, J. S., Radics, L., and Rakhit, S. (1981) The structure of ravidomycin, *Can. J. Chem. 59*, 3018-3020.

159. Davis, M. L., Thoden, J. B., and Holden, H. M. (2007) The X-ray structure of dTDP-4-keto-6-deoxy-D-glucose-3,4-ketoisomerase, *J. Biol. Chem. 282*, 19227-19236.

160. Oh, J., Lee, S. Q., Kim, B. G., Sohng, J. K., Liou, K., and Lee, H. C. (2003) One-pot enzymatic production of dTDP-4-keto-6-deoxy-D-glucose from dTMP and glucose-1-phosphate, *Biotechnol. Bioeng. 84*, 452-458.

161. Wang, Q., Ding, P., Perepelov, A. V., Xu, Y. L., Wang, Y., Knirel, Y. A., Wang, L., and Feng, L. (2008) Characterization of the dTDP-d-fucofuranose biosynthetic pathway in *Escherichia coli* O52, *Mol. Microbiol. 70*, 1358-1367.

162. Perez, M., Lombo, F., Baig, I., Brana, A. F., Rohr, J., Salas, J. A., and Mendez, C. (2006) Combinatorial biosynthesis of antitumor deoxysugar pathways in *Streptomyces griseus*: Reconstitution of "unnatural natural gene clusters" for the biosynthesis of four 2,6-D-dideoxyhexoses, *Appl. Environ. Microbiol. 72*, 6644-6652.

163. Yoshida, Y., Nakano, Y., Nezu, T., Yamashita, Y., and Koga, T. (1999) A novel NDP-6-deoxyhexosyl-4-ulose reductase in the pathway for the synthesis of thymidine diphosphate-D-fucose, *J. Biol. Chem. 274*, 16933-16939.

164. Kang, Y. B., Yang, Y. H., Lee, K. W., Lee, S. G., Sohng, J. K., Lee, H. C., Liou, K., and Kim, B. G. (2006) Preparative synthesis of dTDP-L-rhamnose through combined enzymatic pathways, *Biotechnol. Bioeng. 93*, 21-27.

165. Melancon, C. E., and Liu, H. W. (2007) Engineered biosynthesis of macrolide derivatives bearing the non-natural deoxysugars 4-epi-D-mycaminose and 3-N-monomethylamino-3-deoxy-D-fucose, *J. Am. Chem. Soc. 129*, 4896-+.

166.   Chen, H. W., Yamase, H., Murakami, K., Chang, C. W., Zhao, L. S., Zhao, Z. B., and Liu, H. W. (2002) Expression, purification, and characterization of two N,N-dimethyltransferases, TylM1 and DesVI, involved in the biosynthesis of mycaminose and desosamine, *Biochemistry 41*, 9165-9183.

167.   Chung, Y. S., Kim, D. H., Seo, W. M., Lee, H. C., Liou, K., Oh, T. J., and Sohng, J. K. (2007) Enzymatic synthesis of dTDP-4-amino-4,6-dideoxy-D-glucose using GerB (dTDP-4-keto-6-deoxy-D-glucose aminotransferase), *Carbohydr. Res. 342*, 1412-1418.

168.   Hsu, D. S., Matsumoto, T., and Suzuki, K. (2005) Efficient synthetic route to ravidosamine derivatives, *Synlett*, 801-804.

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

170.   Hopwood, D. A. H., G.; Kieser, T.; Wright, H. M. (1984) Integrated DNA sequences in three streptomyces form related autonomous plasmids after transfer to *Streptomyces lividans*, *Plasmid 11*, 1-16.

171.   Rix, U., Remsing, L. L., Hoffmeister, D., Bechthold, A., and Rohr, J. (2003) Urdamycin L: A novel metabolic shunt product that provides evidence for the role of the *urdM* gene in the urdamycin A biosynthetic pathway of *Streptomyces fradiae* TU 2717, *Chembiochem 4*, 109-111.

172.   Mayer, A., Taguchi, T., Linnenbrink, A., Hofmann, C., Luzhetskyy, A., and Bechthold, A. (2005) LanV, a bifunctional enzyme: aromatase and ketoreductase during landomycin A biosynthesis, *Chembiochem 6*, 2312-2315.

173.   Zhu, L. L., Ostash, B., Rix, U., Nur-e-Alam, M., Mayers, A., Luzhetskyy, A., Mendez, C., Salas, J. A., Bechthold, A., Fedorenko, V., and Rohr, J. (2005) Identification of the function of gene *lndM2* encoding a bifunctional oxygenase-reductase involved in the biosynthesis of the antitumor antibiotic landomycin E by *Streptomyces globisporus* 1912 supports the originally assigned structure for landomycinone, *J. Org. Chem. 70*, 631-638.

174.   Metsa-Ketela, M., Palmu, K., Kunnari, T., Ylihonko, K., and Mantsala, P. (2003) Engineering anthracycline biosynthesis toward angucyclines, *Antimicrob. Agents Chemother. 47*, 1291-1296.

175.   Cone, M. C., Yin, X. H., Grochowski, L. L., Parker, M. R., and Zabriskie, T. M. (2003) The blasticidin S biosynthesis gene cluster from *Streptomyces griseochromogenes*: sequence analysis, organization, and initial characterization, *Chembiochem 4*, 821-828.

176.   Pfeifer, B. A., Admiraal, S. J., Gramajo, H., Cane, D. E., and Khosla, C. (2001) Biosynthesis of complex polyketides in a metabolically engineered strain of *E.coli*, *Science 291*, 1790-1792.

177.   Kulowski, K., Wendt-Pienkowski, E., Han, L., Yang, K. Q., Vining, L. C., and Hutchinson, C. R. (1999) Functional characterization of the *jadI* gene as a cyclase forming angucyclinones, *J. Am. Chem. Soc. 121*, 1786-1794.

178.   Henkel, T., Rohr, J., Beale, J. M., and Schwenen, L. (1990) Landomhycins, new angucycline antibiotcs from *Streptomyces* sp.1. Structural studies on landomycin-A, landomycin-B, landomycin-C and landomycin-D, *J. Antibiot. 43*, 492-503.

179.   Gould, S. J., Melville, C. R., Cone, M. C., Chen, J., and Carney, J. R. (1997) Kinamycin biosynthesis. Synthesis, isolation, and incorporation of stealthin C, an aminobenzo[*b*]fluorene, *J. Org. Chem. 62*, 320-324.

180.   Yang, K. Q., Han, L., Ayer, S. W., and Vining, L. C. (1996) Accumulation of the angucycline antibiotic rabelomycin after disruption of an oxygenase gene in the jadomycin B biosynthetic gene cluster of *Streptomyces venezuelae*, *Microbiol. 142*, 123-132.

181. Liu, W. C., Parker, W. L., Slusarch.D. S., Greenwoo.G. l., Graham, S. F., and Meyers, E. (1970) Isolation, characterization, and structure of rabelomycin, a new antibiotic, *J. Antibiot. 23*, 437-441.

182. Faust, B., Hoffmeister, D., Weitnauer, G., Westrich, L., Haag, S., Schneider, P., Decker, H., Kunzel, E., Rohr, J., and Bechthold, A. (2000) Two new tailoring enzymes, a glycosyltransferase and an oxygenase, involved in biosynthesis of the angucycline antibiotic urdamycin A in *Streptomyces fradiae* Tu2717, *Microbiol. 146*, 147-154.

183. Krohn, K., Ballwanz, F., and Baltus, W. (1993) Total synthesis of angucyclines. 1. Synthesis of daunomycinone-rabelomycin hybrid, *Lieb. Ann. Chem.*, 911-913.

184. Krohn, K., Boker, N., Florke, U., and Freund, C. (1997) Synthesis of angucyclines .8. Biomimetic-type synthesis of rabelomycin, tetrangomycin, and related ring B aromatic angucyclinones, *J. Org. Chem. 62*, 2350-2356.

185. Krohn, K., and Khanbabaee, K. (1994) Synthetic angucyclines .2. First total synthesis of (+/-)-rabelomycin, *Angew. Chem. Intl. Ed. 33*, 99-100.

186. Krohn, K., Khanbabaee, K., Florke, U., Jones, P. G., and Chrapkowski, A. (1994) Synthetic angucyclines, 3. First total synthesis of rac-rabelomycin by Diels-Alder reaction, *Lieb. Anal. Chem.*, 471-477.

187. Valderrama, J. A., Arayamaturana, R., Gonzalez, M. F., Tapia, R., Farina, F., and Paredes, M. C. (1991) Studies on quinones. Part 21. Regioselective synthesis of tetracyclic quinones related to rabelomycin, *J. Chem. Soc.-Perkin Trans.1*, 555-559.

188. James, C. A., and Snieckus, V. (2009) Combined directed remote metalation-transition metal catalyzed cross coupling strategies: the total synthesis of the aglycones of the gilvocarcins V, M, and E and arnottin I, *J. Org. Chem. 74*, 4080-4093.

189. Patra, A., Pahari, P., Ray, S., and Mal, D. (2005) A brief and convergent synthetic route to defucogilvocarcin M chromophore: the formal synthesis of WS-5995 A and C, *J. Org. Chem. 70*, 9017-9020.

190. Takemura, I., Imura, K., Matsumoto, T., and Suzuki, K. (2004) Concise three-component synthesis of defucogilvocarcin M, *Org. Lett. 6*, 2503-2505.

191. Hertzberg, R. P., Caranfa, M. J., and Hecht, S. M. (1989) On the mechanism of topoisomerase I inhibition by camptothecin: evidence for binding to an enzyme-DNA complex, *Biochemistry 28*, 4629-4638.

192. Jaxel, C., Kohn, K. W., Wani, M. C., Wall, M. E., and Pommier, Y. (1989) Structure-activity study of the actions of camptothecin derivatives on mammalian topoisomerase I: evidence for a specific receptor site and a relation to antitumor activity *Cancer Res. 49*, 1465-1469.

193. Kharel, M. K., Zhu, L. L., Liu, T., and Rohr, J. (2007) Multi-oxygenase complexes of the gilvocarcin and jadomycin biosyntheses, *J. Am. Chem. Soc. 129*, 3780-+.

194. Ogasawara, Y., Katayama, K., Minami, A., Otsuka, M., Eguchi, T., and Kakinuma, K. (2004) Cloning, sequencing, and functional analysis of the biosynthetic gene cluster of macrolactam antibiotic vicenistatin in *Streptomyces halstedii*, *Chem. Biol. 11*, 79-86.

195. Xue, Y. Q., Wilson, D., and Sherman, D. H. (2000) Genetic architecture of the polyketide synthases for methymycin and pikromycin series macrolides, *Gene 245*, 203-211.

196. Tsai, S. C., Miercke, L. J. W., Krucinski, J., Gokhale, R., Chen, J. C. H., Foster, P. G., Cane, D. E., Khosla, C., and Stroud, R. M. (2001) Crystal structure of the macrocycle-forming thioesterase domain of the erythromycin polyketide synthase: Versatility from a unique substrate channel, *Proc. Natl. Acad. Sci. USA 98*, 14808-14813.

197. Kearney, A. S., Crawford, L. F., Mehta, S. C., and Radebaugh, G. W. (1993) The interconversion kinetics, equilibrium, and solubilities of the lactone and hydroxyacid forms of the HMG-CoA reductase inhibitor, CI-981 *Pharm. Res. 10*, 1461-1465.

198. You, Z., Omura, S., Ikeda, H., and Cane, D. E. (2006) Pentalenolactone biosynthesis. Molecular cloning and assignment of biochemical function to PtlH, a non-heme iron dioxygenase of *Streptomyces avermitilis*, *J. Am. Chem. Soc. 128*, 6566-6567.

199. R. E. Garrett, C. M. G. (2005) *Biochemistry Thomson Learning, Belmont*, p.727.

200. Strelitz, F. F., H; Asheshov, I. N. (1955) Chrysomycin: a new antibiotic substance for bacterial viruses, *J. Bacteriol. 69*, 280-283.

201. Zhang, W. J., Li, Y. R., and Tang, Y. (2008) Engineered biosynthesis of bacterial aromatic polyketides in *Escherichia coli*, *Proc. Natl. Acad. Sci. USA 105*, 20683-20688.

202. Xiang, L. K., Kalaitzis, J. A., Nilsen, G., Chen, L., and Moore, B. S. (2002) Mutational analysis of the enterocin Favorskii biosynthetic rearrangement, *Org. Lett. 4*, 957-960.

203. Gregory, M. A., Petkovic, H., Lill, R. E., Moss, S. J., Wilkinson, B., Gaisser, S., Leadlay, P. F., and Sheridan, R. M. (2005) Mutasynthesis of rapamycin analogues through the manipulation of a gene governing starter unit biosynthesis, *Angew. Chem. Intl. Ed. 44*, 4757-4760.

204. Levengood, M. R., Knerr, P. J., Oman, T. J., and van der Donk, W. A. (2009) In vitro mutasynthesis of lantibiotic analogues containing nonproteinogenic amino acids, *J. Am. Chem. Soc. 131*, 12024-12025.

Vita (Madan Kumar Kharel)

Date Of birth: 10/31/1975

Place of birh: Maharani Jhoda-9, Jhapa, Mechi, Nepal

Academic degrees

Ph. D. (Biochemistry; August, 2001, - August, 2004) Sun Moon University, Republic of Korea).

MS (Organic Chemistry, 1996-1998) Tribhuvan University, Department of Chemistry, Nepal).

BS (1992-1996): Biology, Tribhuvan University, Nepal

Professional experience

Teaching assistant (January, 2006 - May, 2006), University of Kentucky, Department of Pharmaceutical Sciences

Post-Doctoral Fellow (Oct, 2007- Dec, 2004): College of pharmacy, University of Kentucky, Prof. Jürgen Rohr's lab.

Awards and honors

General travel award, ASP annual meeting (**2009**)

Elizabeth-Hilton Travel award, University of Kentucky, College of Pharmacy (**2009**)

Institute of Biomolecule Reconstruction (IBR) research award, Republic of Korea (**2004**)

Narayan Prasad Pandit Medal, Nepal, (**2001**)

Scientific research publications

1. [a]**Kharel, M. K.**; [a]Pahari, P, Lian, H. and Rohr, J. (**2010**). Enzymatic total synthesis of rabelomycin (**drafted**).
2. **Kharel, M. K.**; Lian, H. and Rohr, J. (**2010**). Characterization of TDP-D-ravidosamine biosynthetic enzymes *in vitro* (**in preparation**).
3. Shepherd, M.D.; **Kharel, M. K** and Rohr, J. (**2010**). Delineating the earliest steps of gilvocarcin biosynthesis; role of GilP and GilQ is starter unit specificity (**drafted**).
4. [a]**Kharel, M. K.**; [a]Nybo, S.E.; [a]Shepherd, M.D. and Rohr, J. (**2010**). Cloning and characterization of the ravidomycin and chrysomycin biosynthetic gene clusters. ***Chembiochem*** (**in press**).
5. **Kharel, M. K.**; Pahari, P.; Lian, H. and Rohr, J. (**2009**). GilR, an unusual lactone-forming enzyme involved in gilvocarcin biosynthesis. ***Chembiochem* 10,** 1305-1308.

6. [a]Liu, T.; [a]**Kharel, M. K.**; Zhu, L.; Bright, S. A.; Mattingly, C.; Adams, V. R. and Rohr, J. (**2009**). Inactivation of the ketoreductase *gilU* Gene of the gilvocarcin biosynthetic gene cluster yields new analogues with partly improved biological activity. ***Chembiochem*** **10**, 278-286.

7. [a]Abdelfattah, M. S.; [a]**Kharel, M. K.;** Hitron, J. A.; Baig, I. and Rohr, J. (**2008**). Moromycins A and B, isolation and structure elucidation of C-glycosylangucycline-type antibiotics from *Streptomyces sp.* KY002. ***J Nat Prod.*** **71**, 1569-1573.

8. **Kharel, M. K.**; Zhu, L.; Liu, T. and Rohr, J. (**2007**). Multi-oxygenase complexes of the gilvocarcin and jadomycin biosyntheses. ***J. Am. Chem. Soc.*** **129**, 3780-3781.

9. [a]Baig I.; [a]**Kharel M.**; [a]Kobylyanskyy, A.; [a]Zhu, L.; Rebets, Y.; Ostash, B.; Luzhetskyy, A.; Bechthold, A.; Fedorenko, V.A. and Rohr, J. (**2006**). On the acceptor substrate of C-glycosyltransferase UrdGT2: three prejadomycin C-Glycosides from an engineered mutant of *Streptomyces globisporus* 1912 DeltalndE(urdGT2). ***Angew. Chem. Int. Ed. Engl.*** **27,** 7842-7846.

10. Liu, T.; **Kharel, M. K.;** Fischer, C.; McCormick, A. and Rohr, J. (**2006**). Inactivation of *gilGT* encoding a C-glycosyltransferase and *gilOIII* encoding a P540-enzyme, allows to delineate details of the late biosynthetic pathway to gilvocarcin V. ***ChemBiochem*** 7, 1070-1077.

11. Subba, B.; **Kharel, M. K.;** Lee, H. C.; Liu, K.; Kim, B.-G. and Sohng, J.K. (**2005**). The ribostamycin biosynthetic gene cluster in *Streptomyces ribosidificus*: comparison with butirosin biosynthesis. ***Mol. Cells.*** **20**, 90-96.

12. **Kharel M. K.;** Subba, B.; Lee, H.C.; Liou, K.; Woo, J. S. and Sohng, J. K. (**2005**). Characterization of L-glutamine:2-deoxy-*scyllo*-inosose aminotransferase (*tbmB*) from *Streptomyces tenebrarius*. ***Bioorg. Med.Chem. Lett.*** **15,** 89-92.

13. Mai, L. T.; **Kharel, M. K.**; Lamichhane, R.; Lee, H. C.; Suh, J.-W.; Liu, K. and Sohng, J. K. (**2005**). Expression of 2-deoxy-*scyllo*-inosose synthase (*kanA*) from kanamycin gene cluster in *Streptomyces lividans*. ***Biotechnol. Lett.*** 27, 465-470.

14. **Kharel M. K.;** Basnet, D. B.; Lee, H. C.; Liou, K.; Woo, J. S. and Sohng, J. K. (**2004**). A gene cluster for biosynthesis of Kanamycin in *S. kanamycetiucs*: comparison with gentamicin biosynthetic gene cluster. ***Arch. Biochem. Biophy*.** **429**, 204-214.

15. **Kharel M. K.**; Basnet, D. B; Lee H. C.; Liou, K.; Woo, J. S.; Kim, B. G. and Sohng, J. K. (**2004**). Isolation and characterization of tobramycin biosynthetic gene cluster from *Streptomyces tenebrarius*. ***FEMS Microbiol. Lett.*** **230**, 185-190.

16. **Kharel, M. K.**; Basnet, D. B.; Lee, H. C.; Liou, K., K.; Woo, J., S.; Kim, D. H.; Moon Y.-H. and Sohng, J. K. (**2004**). Molecular cloning and characterization of 2-deoxystreptamine biosynthetic gene cluster from *Micromonospora echinospora*. ***Mol. Cells.*** **18,** 71-78**.**

17. **Kharel, M. K.**; Lee, H. C.; Liou, K. K. and Sohng, J. K. (**2004**) Immobilization of dTDP-glucose 4,6-dehydratase for the enhanced production of dTDP-4-keto-6-deoxy glucose. ***J. Microbiol. Biotechnol*.** **14**, 297-301.

18. **Kharel, M. K.**; Subba B.; Lee, H. C.; Liou, K. K.; Woo, J., S.; Kim, D. H.; Moon Y.-H. and Sohng, J. K. (**2003**). Identification of 2-deoxy-*scyllo*-inosose synthase in aminoglycoside producer *Streptomyces*. ***J. Microbiol. Biotechnol.*** **13**, 828-831.

19. **Kharel M. K.**; Subba, B.; Lee, H. C.; Liou, K. K.; Woo, J. S. and Sohng, J. K. (**2003**). 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.*** **25**, 2041-2047.

20. **Kharel, M. K.**; Lee, H. C.; Liou, K. K. and Sohng, J. K. (**2002**). Statistical optimization of medium components for the improved production of cystocin by Streptomyces sp.GCA001. ***J. Ind. Eng. Chem.*,** **8**, 427-431.

21. Subba B., **Kharel, M. K.;**Woo, J. S.; Lee, H. C.; Liu, K.; Sohng, J. K. (**2003**). An

approach to clone the biosynthetic genes of 2-deoxystreptamine-containing antibiotics. ***Appl. Chem. 7,*** 101-104.

22. Ta, T.T.T.; **Kharel M. K**.; Heide, L.; Kim, C. G.; Lee, H. C.; Liu, K. K.; Sohng, J. K. (**2004**) Functional characterization of *novS* and *novU* involved in novobiocin producer *S. spheroids* ***Appl. Chem. 8,*** 482-485.

23. Subba, B; **Kharel, M. K.**; Liu, K, Lee, H. C.; Sohng, J. K. (2004). Isolation and characterization of biosynthetic gene clusters for 4,5-disubstituted aminoglycoside antibiotics from actinomcetes. ***Appl. Chem.*** **8** (1) 156-158.

24. **Kharel, M. K.**; Rai, N. P.; Manandhar, M. D.; Elix, J. A and Wardlaw, J. H. (**2000**). Dehydrocollatolic acid, a new depsidone from the lichen *Parmotrema nilgherrense*. ***Aust. J. Chem.***, **53**, 891-892.

[a] Authors contributed equally to the work.