博士論文

Studies on novel polyketides discovered via activation of cryptic gene clusters in *Streptomyces*

（放線菌の潜在的遺伝子クラスターの活性化により発見した新奇ポリケタイドに関する研究）

トン　ウェイ　リー

TABLE OF CONTENTS

ABBREVIATIONS

Chemicals

Apr	Apramycin
DMSO	Dimethyl sulfoxide
EDTA	Ethylenediaminetetraacetic acid
ODS	Octadecyl silane
SDS	Sodium dodecyl sulfate
TEAP	Triethyl amine phosphate
Tris	2-amino-2-hydroxymethyl-1,3-propanediol

Genes · Enzymes

IdmB	Isoindolinomycin biosynthesis
MBG	Methylbenzene gene
NRPS	Nonribosomal peptide synthase
PKS	Polyketide synthase
RNAP	RNA polymerase

Experimental procedures · Instruments

λ	Wavelength [nm]
BLAST	Basic local alignment search tool
COSY	Correlated spectroscopy
HMBC	Heteronuclear multiple bond correlation
HPLC	High performance liquid chromatography
HRESI-MS	High resolution electron spray ionization mass spectrometry
HSQC	Heteronuclear single quantum coherence
Hz	Hertz
Kb	Kilobase
LC-MS	Liquid chromatography-mass spectrometry
NMR	Nuclear magnetic resonance
NOESY	Nuclear Overhauser enhancement spectroscopy
Orf	Open reading frame
PCR	Polymerase chain reaction
UHPLC	Ultra high performance liquid chromatography

UV	Ultraviolet
Vis	Visual
kDa	Kilodalton
m/z	Mass-to-charge ratio
mult	Multiplicity
ppm	Part per million
vol.	Volume

Bacterial strains

E	*Escherichia*
S	*Streptomyces*

INTRODUCTION

INTRODUCTION

Microbial natural products

The search for bioactive natural products starting with the discovery of penicillin by Alexander Fleming and extending to the present day has been an exciting and challenging adventure. For the past 50 years, antibiotic research has provided us with thousands of microbial-derived bioactive compounds with numerous varieties of chemical structures such as tetracyclines, aminoglycosides, and macrolides (Figure 1). Of these bioactive compounds, many are isolated from actinomycetes, typically from members of the genus *Streptomyces.*[1] To date, actinomycetes remain an important natural source for the production of naturally derived antibiotics that are used clinically.[2] Despite the success of the microorganism as a source of natural products, the discovery of new bioactive metabolites from actinomycetes has been decreasing since the end of 1970s (Table 1).[3] Because many of the readily obtainable bioactive compounds have been identified, conventional screening for novel bioactive compounds from microorganisms often results in the rediscovery of known compounds. Consequently, antibiotic research from microbial origins has been stagnant, and efforts are instead focused on the generation of synthetic structural analogs. However, recent genome sequencing has revealed that actinomycetes are capable of providing more than what we have now.

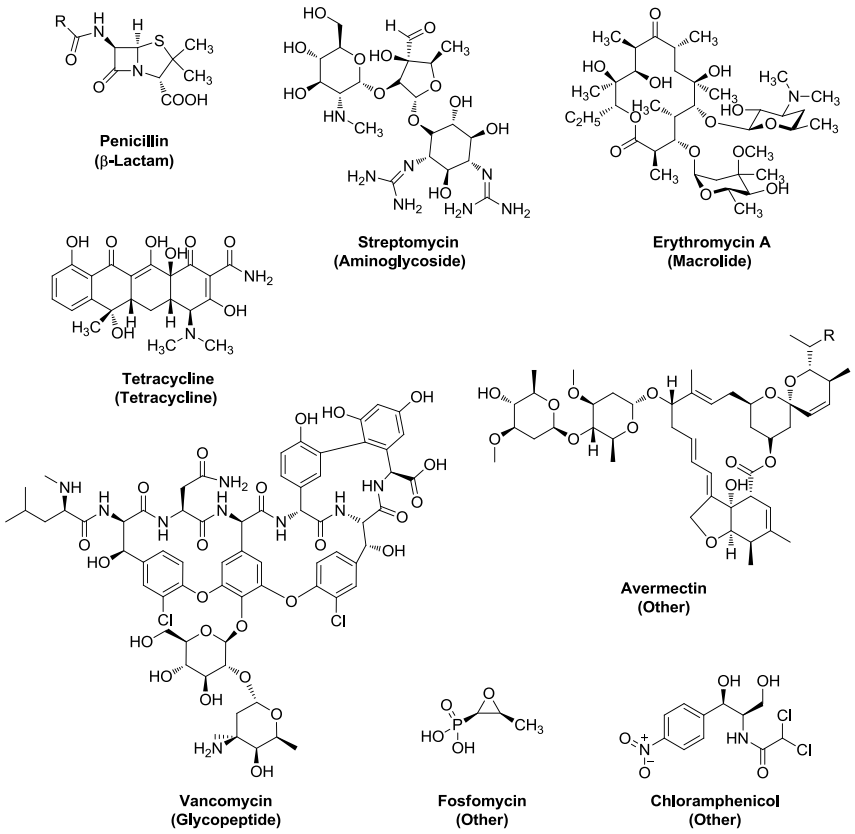


Figure 1. Types of microbial-derived bioactive compounds.

Table 1. Approximate number of bioactive metabolites from actinomycetes from 1940 to 2010

	1940-1974		1975-2000		2001-2010		
Periods	**Early years**	**%**	**Mid-era**	**%**	**New age**	**%**	**Total**
Actinobacteria	3,400	62[a]	7,200	42[a]	3,100	28.5[a]	13,700
***Streptomyces* sp.**	2,900		5,100		2,400		10,400
Other actinobacteria	500		2,100		700		3,300

[a] Percentages shown represent the amount of bioactive metabolites from actinomycetes compared with total metabolites from microbial producers.

Adapted from Berdy, J (2012).[3]

Genetic potential of actinomycetes

A typical actinomycete strain has more than 8 million base pairs of nucleotides in its chromosomes, which are estimated to encode more than 20 biosynthetic gene clusters for secondary metabolites [4-6]. Although these gene clusters are predicted to encode potential unprecedented secondary metabolites, only a fraction of the genes have been expressed and characterized. For example, *Streptomyces coelicolor* A(3)2, which is one of the best studied actinomycete strains, has biosynthetic gene clusters for several compounds including actinorhodin (Act), prodiginine, and calcium-dependent antibiotic (CDA). However, some other gene clusters found in the genome predicted to encode for type I, II, and III polyketide synthase (PKS), terpene cyclase, and non-ribosomal peptide synthase (NRPS) remain uncharacterized.[6] Genome sequencing of the streptomycin producer, *Streptomyces griseus* IFO 13350, has also revealed 36 biosynthetic gene clusters, of which only 9 have been characterized and their metabolites identified (Figure 2A). However, the remaining gene clusters predicted to produce metabolites, such as polyketide, terpene, non-ribosomal peptide, lantibiotic, and thiazolyl peptide, remain elusive (Figure 2B).[5, 7] Furthermore, the draft genome sequence of some actinomycete strains used in this study reveal that each strain has more than 30 secondary metabolite gene clusters. However, only one to two compounds were reported from each strain, indicating the possibility that more compounds could be produced by these strains (Table 2).

Why have these genetic potentials remained undiscovered? According to Bode et al. (2002), there are a few reasons: 1) analytical and detection limitations, 2) the gene clusters are incomplete or non-functional, or 3) gene clusters are "silenced" due to the standard laboratory cultivation conditions.[8] In most cases, it is the "silent" or "cryptic" gene clusters that produce unknown secondary metabolites. Therefore, the current main issue is to develop a method that

can unlock these cryptic genes to obtain enough metabolites for structure elucidation and biological testing.

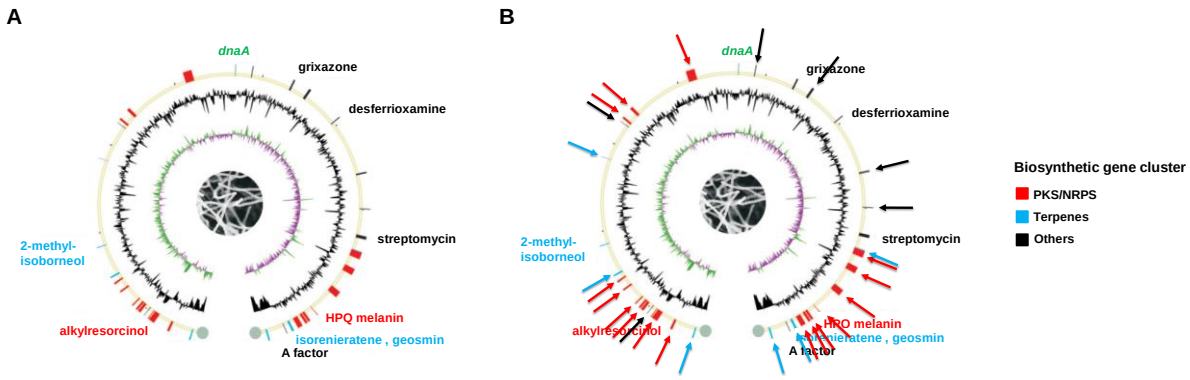


Figure 2. Representation of the chromosome of *Streptomyces griseus* IFO 13350 adapted from Nett et al. (2009).[7] Biosynthetic gene clusters encoding PKS or NRPS are labeled in red, terpenes in blue, and all other secondary metabolites in black. Identified products are labeled accordingly, and the arrows in (**B**) are gene clusters with unknown metabolites.

Table 2. Number of predicted gene cluster and isolated compound(s) in some actinomycete strains used in this study.

Actinomycetes	Isolated Compound(s)	Number of predicted gene cluster*
Streptomyces sp. CL190	Naphterpin [9], Lavanducyanin [10]	60
Streptomyces lysosuperificus AV791	Tunicamycin [11]	74
Streptomyces sp. SANK 60404	A-94964 [12]	65
Streptomyces sp. WK-2057	Phenazinomycin [13]	46
Streptomyces versipellis 4083-SVS6	Versipelostatin [14]	55
Saccharothrix sp. ST-888	Phosphonothrixin [15]	70
Streptomyces sp. RM72	Trichostatin [16]	57
Verrucosispora gifhornensis YM28-088	Gifhornenolones [17]	32
Streptomyces spectabilis ND90	SNF4435C and D [18]	75
Kitasatospora setae KM-6054	Bafilomycin B1[19]	51
Streptomyces sp. FGK858	Lavanduquinocin [20]	54

*Source, http://chemobio.cbrc.jp/natural/

Optimizing secondary metabolite production in actinomycetes

The most common method to improve the production of secondary metabolites in actinomycetes is the systematic modification of the culture conditions such as media constituents, temperature, pH, cultivation period, and oxygen supply. The logic is that to survive, an organism changes its metabolic profile according to the environment. Additionally, to accommodate these changes, different secondary metabolites might be produced.[21] However, due to our limited knowledge of the complex regulatory and biosynthetic crosstalk in actinomycetes,[8] this approach may be rather time-consuming and costly.

With the current advances in both sequencing technologies and the understanding of the metabolic pathways of actinomycetes, cryptic gene activation using genetic tools is a favorable strategy. Genome sequences have not only provided a platform for detailed bioinformatics analysis on potential gene clusters but have also opened the door for designing possible cryptic gene-activation approaches (Figure 3). MS/MS-based networking guided genome mining, the manipulation of pathway-specific regulators, and heterologous expression are among the methods governed by genome sequences.[22, 23]

MS/MS-based networking guided genome mining is a very recent method developed by Dorrestein and colleagues. Their approach is to correlate MS/MS fragmentations to gene clusters encoding the secondary metabolites, thus creating an MS/MS network.[24, 25] This MS/MS networking not only provides an overview of the type of metabolites produced by a bacterial strain but also allows the identification of the gene clusters responsible for their production.

On the other hand, the manipulation of pathway-specific regulators offers a more straightforward method to activate the expression of target gene clusters. Unlike global regulators that act on multiple metabolic pathways simultaneously, a pathway-specific

regulator resides in a biosynthetic gene cluster and acts only on its gene cluster. The overexpression of pathway-specific regulators has been demonstrated to trigger the generation of secondary metabolites. For example, in *Streptomyces globisporus* C-1027, the consecutive overexpression of the 3 regulatory genes *sgcR1*, *sgcR2*, and *sgcR3* increased the production of enediyne C-1027. Curiously, the concurrent overexpression of these 3 genes did not affect the production of C-1027. This result implies the complex interaction of these regulators that belong to the AraC/XylS family.[26, 27]

The method of heterologous expression relies heavily on genome sequence; a gene cluster of interest is cloned and expressed in a heterologous host. Heterologous expression is becoming an established method to unlock cryptic gene clusters where gene clusters that are not expressed in their native host are introduced to engineered "clean" hosts to generate the products.[28] Luo et al. assembled the entire cryptic polycyclic tetramate macrolactams (PTM) biosynthetic gene cluster from *Streptomyces griseus* and expressed it in *Streptomyces lividans.*[29] The heterologous expression led to the identification of 2 additional PTM compounds that are not detected in the native host.

While all these gene cluster-based methods have shown success, there are also broader, sequence-independent strategies for cryptic gene activation. Methods such as co-culturing and the selection of mutants conferring antibiotic resistance are 2 successful examples reported (Figure 3). Co-culturing utilizes the concept that interspecies interaction can induce the production of certain bioactive metabolites, often for defense purposes. The production of new compounds has been repeatedly discovered by co-culturing different bacterial species.[30, 31] Another effective example is the co-culturing of *Streptomyces endus* S-522 with a mycolic acid-containing bacteria *Tsukamurella pulmonis* TP-B0596, which induces the production of a novel antibiotic, alchivemycin A.[32]

The selection of mutants conferring antibiotic resistance is the method utilized in this study to explore new secondary metabolites. Here, actinomycete strains are grown on rifampicin to generate rifampicin-resistant mutants. The cryptic gene activation mechanism of this approach will be discussed in the next section.

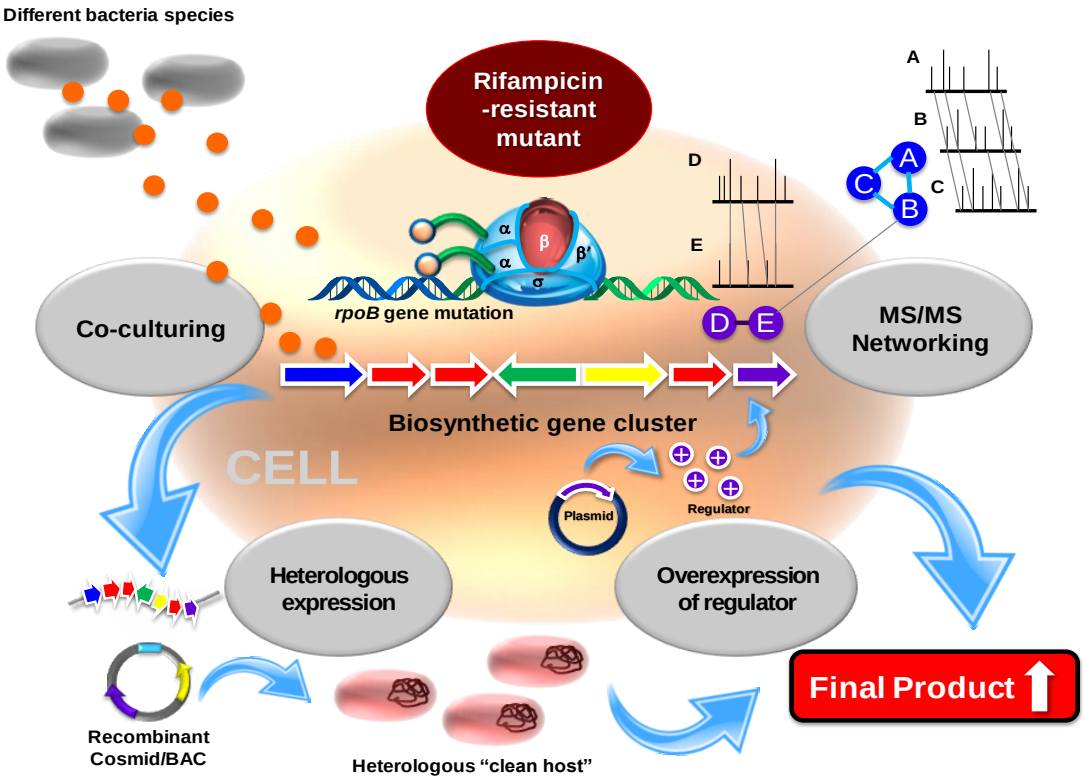


Figure 3. Cryptic gene activation strategies discussed in this chapter.

Antibiotic production by rifampicin-resistant mutants

Ochi and co-workers have developed another straightforward yet effective method to unlock cryptic gene clusters by simply generating spontaneous antibiotic (targeting ribosome or RNA polymerase) resistant mutants. They found that when a *Streptomyces* strain acquired rifampicin resistance, the rifampicin-resistant (*rif)* mutants sometimes overproduced secondary metabolites compared with the parent strain (Figure 4).[33] Rifampicin targets a broad spectrum of bacteria by binding to their RNA polymerase (RNAP). *Rif* mutants have mutation(s) primarily in their *rpoB* gene, which encodes the RNAP β-subunit.[34] In *Bacillus subtilis* and *Escherichia coli*, these *rpoB* mutations were reported to affect a number of phenotypes such as sporulation, metabolism, stability of plasmids, and susceptibility to other inhibitors.[35-37] Ochi's research group demonstrated that *rif* mutants generated from actinomycetes could exhibit an up to 70-fold increase in the transcription of secondary metabolite genes.[38] The group rationalized that the activation mechanism might mimic a guanosine 5'-diphosphate 3'-diphosphate (ppGpp)-bound RNAP, which regulates the transcription of stable RNA.[39] ppGpp is a signal molecule responsible for growth rate control and is suggested to have a close relation to the onset of antibiotic production [40].

A recent paper reported by Hosaka et al. further described that this method is effective for the activation of silent genes.[41] They reported that a *rpoB* mutant KO-916 (H437L) and a *rpoB* (H437L) and *rpsL* (K88R) double mutant KO-923 from a non-producing strain, *Streptomyces mauvecolor* 631689, produced a cryptic new type of antibiotics called piperidamycins (Figure 5). Surface plasmon resonance analysis of the wild-type and *rpoB* mutant RNAPs from stationary-phase cells showed that the two mutant RNAPs have significantly higher affinity toward the silent gene promoters compared with wild-type RNAP.

Similarly, a *rif* mutant acquired when a deep sea-derived *Streptomyces somaliensis* SCSIO ZH66 was cultivated with 300 μg mL^{-1} rifampicin unlocked the production of an anti-cancer drug lead called fredericamycin. Using this method followed by media optimization, the group was able to increase fredericamycin from non-detectable production to titers of 679.5±15.8 mg L^{-1}.[42]

Considering these successful cases, the screening of mutants exhibiting rifampicin resistance was used in this study to identify potential novel compounds in actinomycetes.

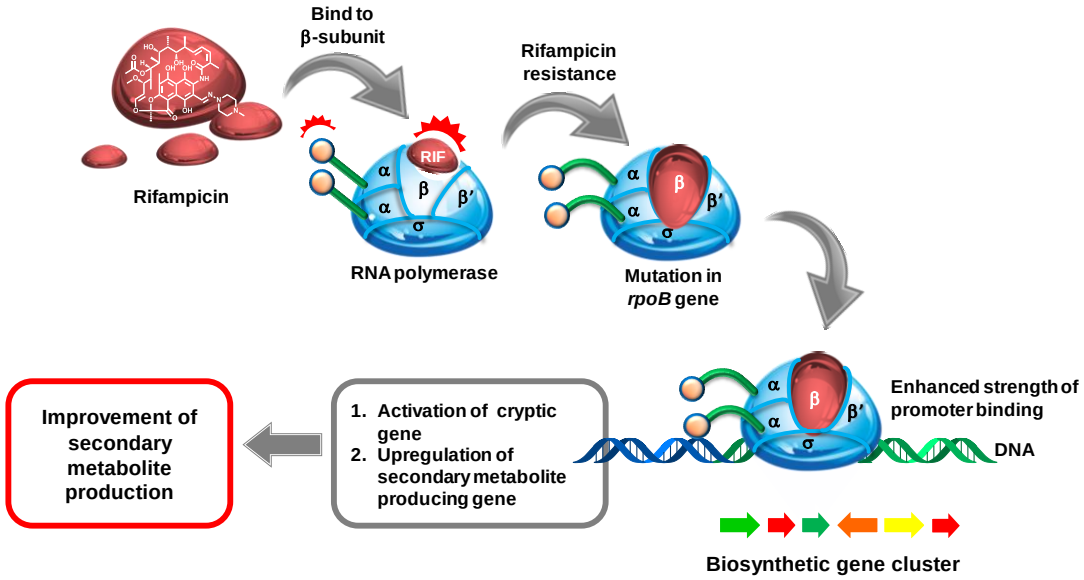


Figure 4. Cryptic gene activation by the generation of spontaneous *rif* mutant.

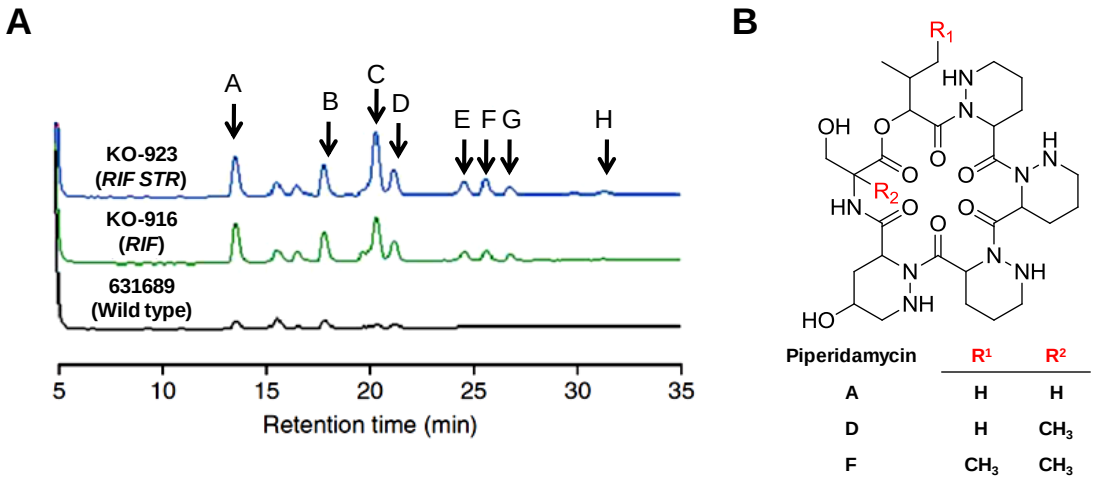


Figure 5. Detection of the new compounds produced by drug-resistant mutants. (**A**) Comparative HPLC profiles showing the overproduced products. (**B**) Structures of the cryptic new antibacterial compounds. Edited from Hosaka et al. (2009).[41]

The aims of my research are as follows:

1. Generate *rif* mutants from 20 actinomycete strains,
2. Screen for overproduced secondary metabolites in all *rif* mutants obtained,
3. Identify the overproduced secondary metabolites, and
4. Propose a biosynthetic pathway for the identified compounds.

CHAPTER 1

CHAPTER 1: Screening of rifampicin-resistant mutants and comparative metabolic analysis

Overview

The cryptic gene activation method used in this study is based on the screening of mutants that acquire rifampicin resistance. Rifampicin-resistant (*rif*) mutants emerge after 10^7-10^9 bacteria are grown on a plate containing a high concentration of rifampicin.[39] These *rif* mutants are reported to possess either a point mutation or a deletion mutation that is often in the *rif* cluster I region within the *rpoB* gene, which encodes the RNA polymerase (RNAP) β-subunit (Figure 1-1).[38, 43] The frequency of spontaneous *rpoB* mutation ranges from 10^{-6} to 10^{-8} depending on the bacterial strain.[39] The effect of the *rpoB* mutation on the activation of cryptic genes can be evaluated by comparative HPLC analysis of the culture extracts of *rif* mutants and control wild-type strains.

In this chapter, spontaneous *rif* mutant generation was performed on 20 actinomycete strains for which we have draft genome sequences (Table 1-1). The generation of *rif* mutants started with determining the minimum inhibitory concentration (MIC) for each strain used. The MICs obtained were then used as references to choose the concentration of rifampicin needed for *rif* mutant generation in each strain. *Rif* mutants that developed spontaneously after induction were isolated before being cultivated and extracted for HPLC analysis comparison with wild-type strains.

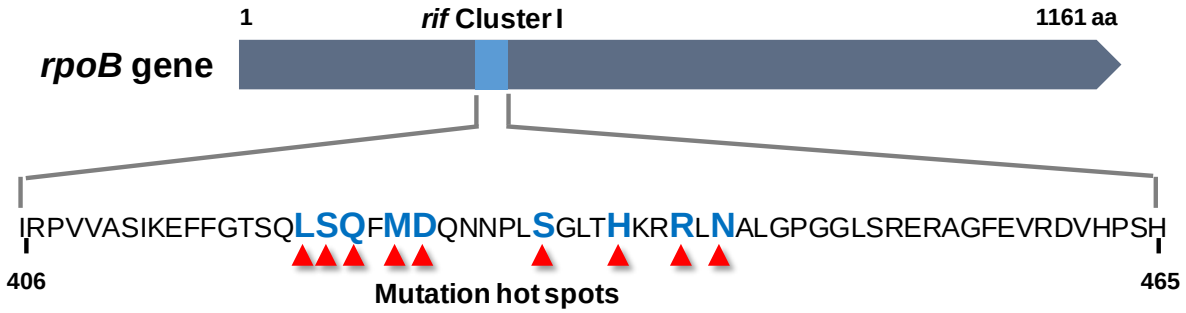


Figure 1-1. The location of *rif* cluster I within the *rpoB* gene, which encodes RNA polymerase β-subunit in *Streptomyces coelicolor* A3(2) according to Xu et al. (2002).[43] Hot spots for mutation found in *rif* mutants are indicated by arrows.[38] Numbering begins at the start codon in *Streptomyces coelicolor* A3(2).

Table 1-1. Actinomycete strains used in this study.

Actinomycetes	Abbreviation	References
Streptomyces sp. CL190	CL190	Shin-ya et al. (1990)[44]
		Imai et al. (1989)[10]
Streptomyces lysosuperificus AV791	AV791	Takatsuki et al. (1971)[11]
Streptomyces sp. WK-2057	WK-2057	Omura et al. (1989)[13]
Streptomyces sp. SANK 60404	SANK 60404	Murakami et al. (2008)[12]
Streptomyces versipellis 4083-SVS6	4083-SVS6	Ueda et al. (2008)[14]
Saccharothrix sp. ST-888	ST-888	Takahashi et al. (1995)[15]
Streptomyces sp. RM72	RM72	Hosoya et al. (2012)[16]
Verrucosispora gifhornensis YM28-088	YM28-088	Shirai et al. (2010)[17]
Streptomyces spectabilis ND90	ND90	Kurosawa et al. (2001)[18]
Kitasatospora setae KM-6054	KM-6054	Omura et al. (1981)[19]
Streptomyces sp. FGK858	FGK858	Shin-ya et al. (1995)[20]
Streptomyces sp. TB090715SC-03 (G9)	G9	Culture collections from AIST*
Streptomyces sp. Sp080902JE-04 (S25)	S25	Culture collections from AIST
Streptomyces sp. MSB090630SC-24 (S34)	S34	Culture collections from AIST
Streptomyces sp. TB090715LN-02 (S41)	S41	Culture collections from AIST
Streptomyces sp. TB090715SC-05 (S44)	S44	Culture collections from AIST
Streptomyces sp. TB090715SC-07 (S45)	S45	Culture collections from AIST
Streptomyces sp. SoE090715HV-08 (S47)	S47	Culture collections from AIST
Streptomyces sp. SoC090715LN-16 (S55)	S55	Culture collections from AIST
Streptomyces sp. SoC090715LN-17 (S56)	S56	Culture collections from AIST

*AIST, National Institute of Advanced Industrial Science and Technology.

1.1 Determination of minimum inhibitory concentration (MIC)

Methods

The wild-type strains used in this study, for which draft genome sequences are available, are listed in Table 1-1. The MIC was determined by streaking the actinomycete strain on a tryptone soya broth (TSB) agar plate containing 3% TSB (Oxoid, England) and 2% agar with different concentrations of rifampicin and incubated at 30°C for 3 days. The minimum drug concentration that fully inhibited the growth of the strain was defined as the MIC. This MIC was then used as the reference to choose the concentration of rifampicin needed for resistant mutant generation in each strain.

Results and discussion

The MICs of rifampicin for all the actinomycetes used are summarized in Table 1-2. Rifampicin is a potent broad spectrum antibiotic that inhibits bacterial DNA-dependent RNA synthesis.[45] Consistent with previous reports, rifampicin was potent against most of the strains used at 3.125 μg mL^{-1} and at lower concentrations for some. Particularly, the strains SANK 60404, ST-888, YM28-088, FGK858, and S56 were sensitive to the drug with MICs less than 1 μg mL^{-1}. However, strains RM72 and ND90 had high MIC values of 25 μg mL^{-1} and 100 μg mL^{-1}, respectively, indicating that they are most likely naturally resistant against rifampicin. Rifampicin binds to the pocket of the RNAP β-subunit adjacent to the RNAP active site and thereby prevents the elongation of RNA.[46] Therefore, RNAP of RM72 and ND90 may have low affinity to rifampicin. This possibility was supported by the alignment of the *rif* cluster I regions of the RNAP β-subunit using draft genome sequences with different types of bacteria; the strains RM72 and ND90 have altered amino acids at E457D and A443S, respectively (Figure 1-2). These variations may change structure of their RNAP β-subunits and reduce the binding affinity of rifampicin.

Table 1-2. MICs of rifampicin and concentrations of rifampicin used for mutant generation in each actinomycete strain.

Strain	MIC (μg mL^{-1})[a]	Concentration of rifampicin used for mutant generation (μg mL^{-1})
CL190	3.125	15, 30, 150
AV791	3.125	15, 30, 150
WK-2057	3.125	15, 30, 150
SANK 60404	0.195	1, 2, 10
4083-SVS6	3.125	15, 30, 150
ST-888	0.049	0.25, 0.5, 2.5
RM72	25	125, 250
YM28-088	0.195	1, 2, 10
ND90	100	500
KM-6054	6.25	30, 60, 300
FGK858	0.781	4, 8, 40
G9	3.125	15, 30, 150
S25	3.125	15, 30, 150
S34	1.562	8, 16, 80
S41	1.562	8, 16, 80
S44	3.125	15, 30, 150
S45	3.125	15, 30, 150
S47	1.562	8, 16, 80
S55	3.125	15, 30, 150
S56	<0.02	0.1, 0.2, 1
G9	3.125	15, 30, 150

[a] Determined after incubation on TSB agar medium at 30°C for 3 days.

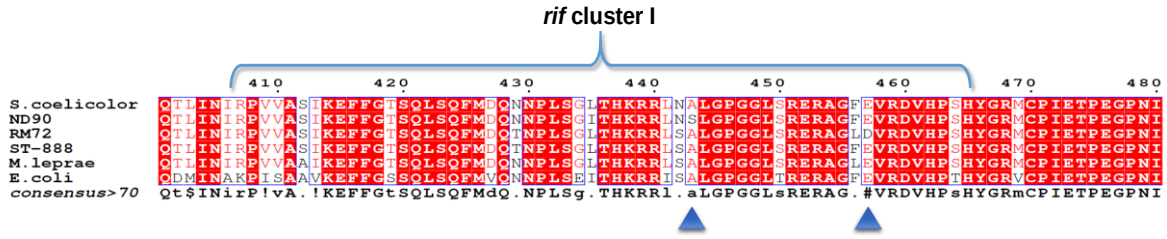


Figure 1-2. Alignments of *rif* cluster I in various bacteria according to Xu et al. (2002). [43] The arrows indicate the altered amino acids that may confer rifampicin resistance to the strains RM72 and ND90. Numbering was adjusted to the start codon in *Streptomyces coelicolor* A3(2). S. coelicolor, *Streptomyces coelicolor* A3(2) [NP_628815]; ST-888, *Saccharothrix* sp. ST-888 [WP_045304571.1]; M. leprae, *Mycobacterium leprae* [CAC30845.1]; E. coli: *Escherichia coli* [CAA23625.1].

1.2 Preparation of *rif* mutants.

Methods

To generate *rif* mutants, cells from each strain were harvested from 100 mL of culture by centrifugation at 3,600×*g* for 10 min, which was followed by two washes with autoclaved distilled water. The cells were then spread onto GYM agar containing different concentrations of rifampicin according to the MIC of each strain (Table 1-2). The induction of rifampicin resistance was performed at 5-fold, 10-fold, and 50-fold MICs of rifampicin for each actinomycete strain. Due to the possible natural resistance, mutant generation for the strains RM72 and ND90 was performed at lower 5- and 10-fold MICs for RM72 and at 5-fold MIC for ND90. Spontaneous *rif* mutants were obtained as colonies that grew within 4 – 7 days of incubation at 30°C. Single colonies were isolated from the growing colonies before storing for further studies.

Results and discussion

Spontaneous *rif* mutant generation in the 20 actinomycete strains used in this study resulted in the isolation of a total of 278 resistant mutants (Table 1-3). From all the strains used, only 16 were responsible for the mutants obtained. The number of *rif* mutants varied depending on the strain and seemed to be unrelated to the concentration of rifampicin used. The same concentration of rifampicin yielded 90 mutants from the strain WK-2057 but only one mutant from the strain AV791. Moreover, mutants derived from WK-2057 appeared to have obvious variation in their morphology at different concentrations of rifampicin (Figure 1-3). A previous study on *rif* mutants showed that *rif* mutations may contribute to the formation of different conformations and changes in the 3-dimensional structure of the β-subunit of RNAP and alter promoter affinity [33]. Thus, the variation in the appearance of the resistant mutants may be due to the changes in gene expression that affect the phenotypes of the mutants.

Additionally, no mutant was obtained from the strains RM72, ND90, KM-6054, or FGK858. As mentioned above, RM72 and ND90 have high MICs of 25 μg mL^{-1} and 100 μg mL^{-1} of rifampicin. These MIC values suggest that RM72 and ND90 are naturally rifampicin resistant, which explains why no *rif* mutant emerged after induction. In contrast, the strains FGK858 and S47 were sensitive to rifampicin. However, no *rif* mutant was acquired from these strains either. The strain FGK858 formed only spores on the rifampicin-containing agar plates, which may act as a defense mechanism against the drug (Figure 1-3).

In the next step, comparative metabolic screening was performed on all mutants obtained in strains CL190 through FGK585, but only on 50-fold MIC mutants for strains G9 through S56 (Table 1-5). This choice was made because *rif* mutants of the strains CL190 to FGK585 that developed from high concentrations of rifampicin were observed to have better overproduction compared with those developed from the lower concentrations.[47] Additionally, Ochi et al. (2004) reported that mutants resistant to high concentrations of drug are more stable in their phenotype.[39]

Table 1-3. Rifampicin-resistant mutants obtained at different folds of MIC.

Strain	Number of resistant mutants		
	(5× MIC)[a]	(10× MIC)[a]	(50× MIC)[a]
CL190	1	5	0
AV791	0	1	0
WK-2057	23	39	28
SANK 60404	7	7	14
4083-SVS6	8	6	10
ST-888	3	6	0
RM72	0	0	N/A[b]
YM28-088	0	0	6
ND90	0	N/A	N/A
KM-6054	0	0	0
FGK858	0	0	0
G9	0	3	0
S25	0	0	15
S34	0	15	0
S41	0	10	7
S44	0	10	2
S45	15	15	4
S47	0	0	0
S55	0	10	7
S56	0	0	1
Total	57	127	94

[a] Determined after single colony isolation on TSB agar medium containing rifampicin at MIC for each strain.

[b] N/A, Not applicable

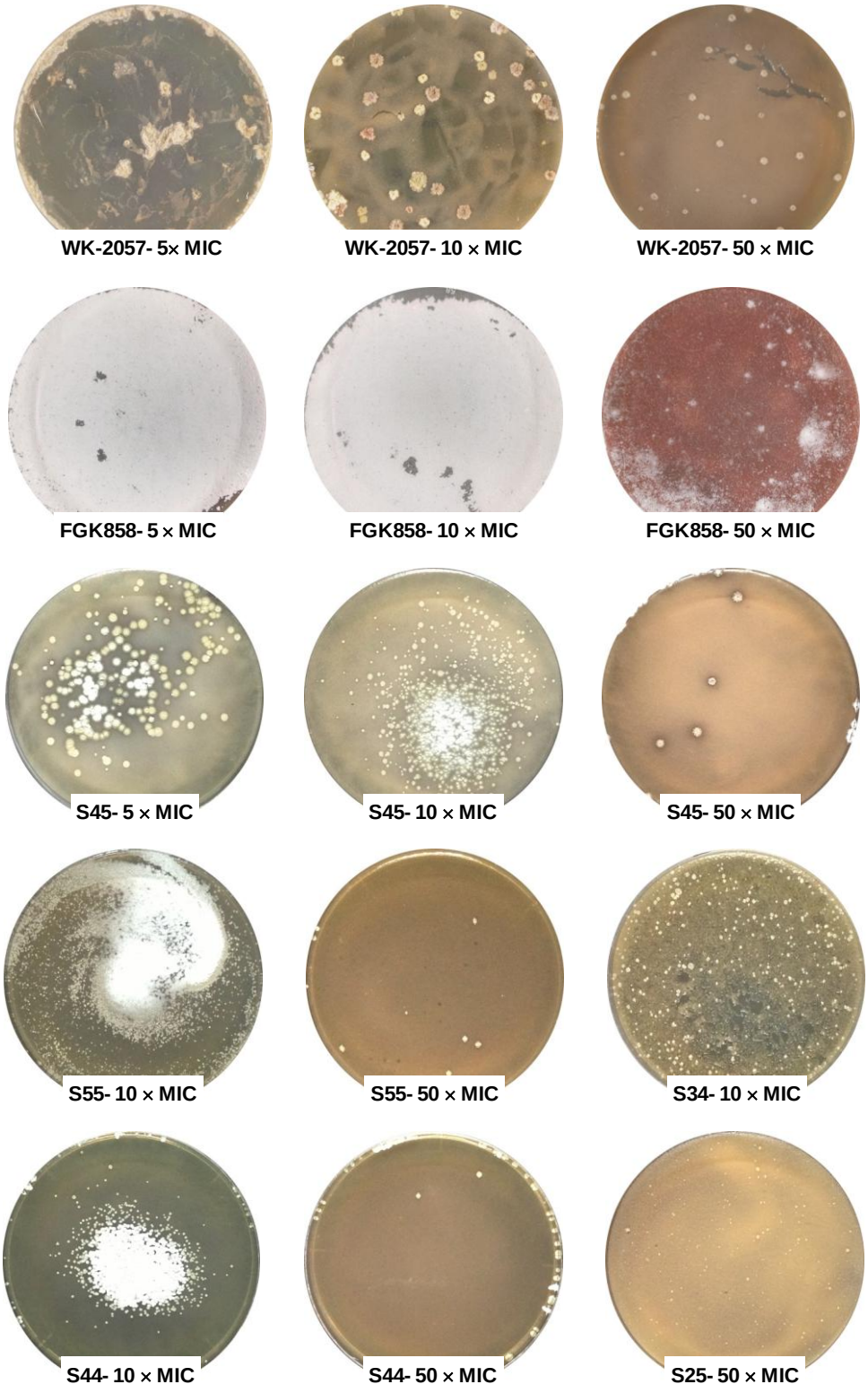


Figure 1-3. Some of the selected induction plates for rifampicin resistance. The morphology of some mutants varies according to the concentration of rifampicin. Colonies with heavy sporulation observed in strains FGK858 and S55-10×MIC are not *rif* mutants.

1.3 Comparative metabolic screening of *rif* mutants

Methods

Rif mutants and their wild-type strains, as listed in Table 1-5, were pre-cultured in 10 mL of 3% TSB (supplemented with rifampicin at the MIC for *rif* mutants) for 2 days before growing in a 50-mL test tube containing 10 mL of K medium and incubated on a reciprocal shaker (300 rpm) at 30°C. After 4 days of cultivation, the culture broths were extracted with 2 volumes of acetone, filtered, and evaporated *in vacuo* to remove the solvent. The resulting solutions were then extracted twice with 2 volumes of ethyl acetate before being evaporated to dryness. The residue was dissolved in 1 mL of methanol. The metabolite analysis was performed with UHPLC. Aliquots of 5 μL of each residue were subjected to UHPLC using the condition listed in Table 1-4. For cultivation using baffled flasks, 2 mL of pre-culture (prepared as mentioned above) was inoculated into a 500-mL baffled flask containing 100 mL of medium and cultivated at 27°C for 4 days on a rotary shaker (180 rpm). Extraction and HPLC analysis of the extracts were performed in the same manner as described above.

Table 1-4: UHPLC conditions.

Column	CAPCELLPAK C18 column (2.0 × 50 mm; Shiseido, Tokyo, Japan)
Solvent	Mobile phase A: Water + 0.1% formic acid
	Mobile phase B: acetonitrile + 0.1% formic acid
	10% B for 1 min, a linear gradient of 10%–90% B for 4 min, 90% B for 1 min, and 10% B for 5 min
Flow rate	0.4 mL min^{-1}
Detector	Photo Diode Array (JASCO X-LC 3110MD)

Table 1-5. Rifampicin-resistant mutants that were screened for their metabolite profiles.

Strain	Number of resistant mutants screened	Annotation[a]
CL190	6	CL190-5-1 CL190-10-1 to CL190-10-5
AV791	1	AV791-10-1
WK-2057	90	WK2057-5-1 to WK2057-5-23 WK2057-10-1 to WK2057-10-39 WK2057-50-1 to WK2057-50-28
SANK 60404	28	TW-R5-1 to TW-R5-7 TW-R10-1 to TW-R10-7 TW-R50-1 to TW-R50-14
4083-SVS6	24	SVS6-5-1 to SVS6-5-8 SVS6-10-1 to SVS6-10-6 SVS6-50-1 to SVS6-50-10
ST-888	9	ST888-5-1 to ST888-5-3 ST888-10-1 to ST888-10-6
YM28-088	6	YM28-50-1 to YM28-50-6
S25	15	S25-50-1 to S25-50-15
S41	7	S41-50-1 to S41-50-5
S44	2	S44-50-1 to S44-50-2
S45	4	S45-50-1 to S45-50-4
S55	7	S55-50-1 to S55-50-7
S56	1	S56-50-1

[a] Resistant mutants are annotated according to the degree of MIC used.

Results and discussion

Comparative metabolic screening was performed on the *rif* mutants listed in Table 1-5. The results discussed below are the representative data acquired from the screening.

Streptomyces sp. WK-2057

The cultivation of the *rif* mutants of WK-2057 along with the parent strain revealed that 2 mutants (WK2057-10-9 and WK2057-10-39) produced several additional metabolites with UV absorbance at a range of over 400 nm (Figure 1-4). However, these metabolites were found in low quantity at 10-mL scale culture. Thus, the mutants were cultivated in a larger scale (100 mL) of K medium.

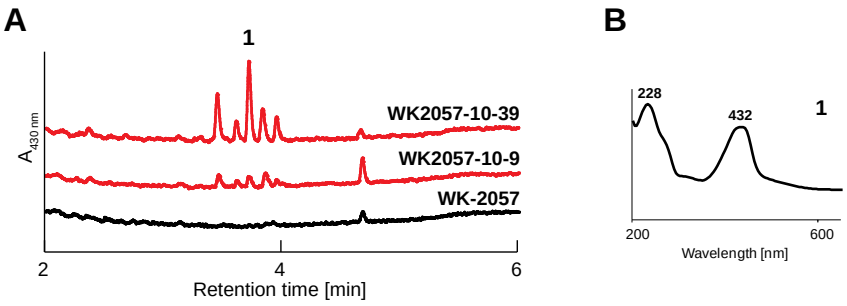


Figure 1-4. (**A**) Comparative HPLC analysis of metabolites in the culture of *Streptomyces* sp. WK-2057 and its *rif* mutants WK2057-10-9 and WK2057-10-39 in K medium. (**B**) UV spectrum of the major metabolite **1**.

As expected, the cultivation of WK2057-10-9 and WK2057-10-39 in baffled flasks confirmed the production of the additional metabolites (Figure 1-5A). Additionally, the mutant WK2057-10-39 showed significantly higher production of a known compound, phenazinomycin (Figure 1-5B).[13] Phenazinomycin was previously known in our laboratory to be poorly produced in K medium, and WK-2057 needed to be cultivated in its producing medium for detection. Additionally, to investigate whether the production of phenazinomycin can be further increased, I cultivated WK2057-10-39 in the phenazinomycin-producing

medium along with its wild-type strain. However, the production of phenazinomycin by WK2057-10-39 increased only slightly compared with the wild-type strain. This result may be due to possible limiting factors such as the quantities of biosynthetic precursors or self-defense against the cytotoxicity of phenazinomycin[13], which would prevent the mutant from producing additional compound (Figure 1-6).

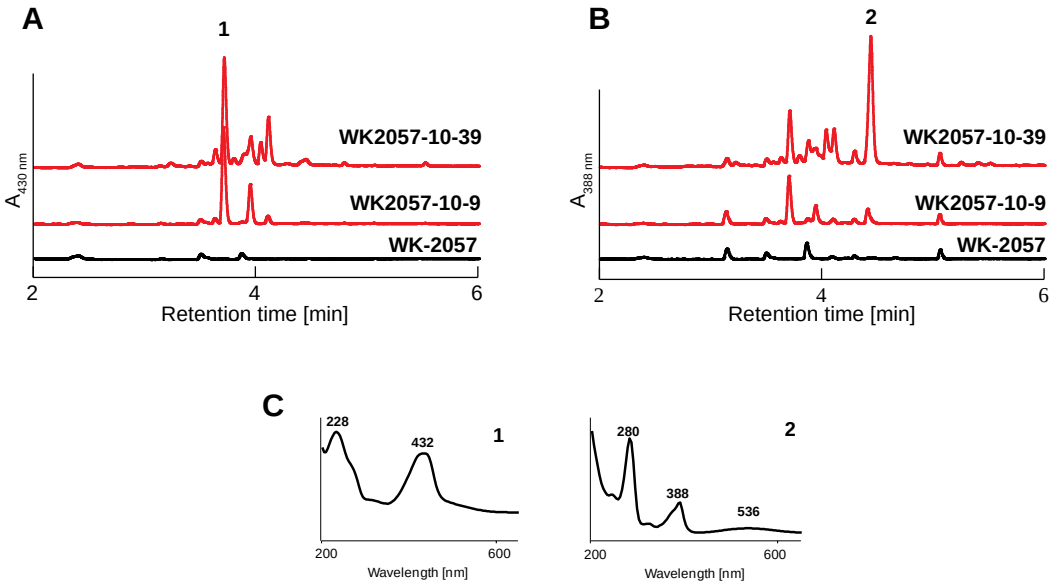


Figure 1-5. UHPLC analysis of the culture extracts from the strain WK-2057 and its mutants cultivated in 100 mL of K medium. (**A**) Chromatogram at 430 nm showing metabolite **1** as the major peak of the overproduced peaks. (**B**) Chromatogram at 388 nm with the overproduced metabolite **2**, which corresponds to phenazinomycin. (**C**) UV spectra of metabolites **1** and **2** produced in the mutant WK2057-10-39.

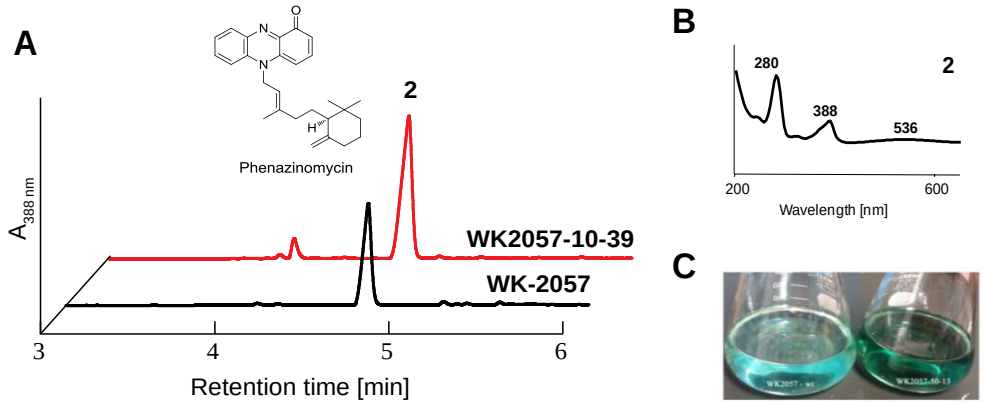


Figure 1-6. (**A**) The production of phenazinomycin in the culture extracts from the mutant WK2057-10-39 and its wild-type strain. (**B**) UV spectrum of phenazinomycin in its producing medium. (**C**) Ethyl acetate extracts from the wild-type strain and WK2057-10-39 showing the difference in the color intensity due to the higher productivity of phenazinomycin.

Streptomyces sp. SANK 60404

The induction of spontaneous rifampicin in *Streptomyces* sp. SANK 60404 resulted in the isolation of a total of 28 *rif* mutants (Table 1-3). The comparative UHPLC analysis of the *rif* mutants and their wild-type strains revealed a mutant (TW-R50-13) that had enhanced production in its metabolic profile when cultivated using K medium (Figure 1-7A). The enhanced metabolites, however, were also produced by the wild-type strain when the cultivation was performed at 100 mL scale (Figure 1-7B).

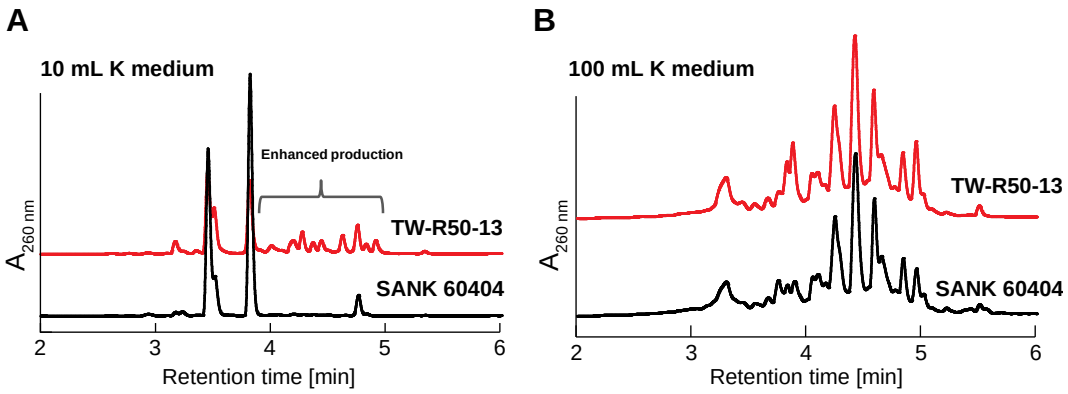


Figure 1-7. Comparative HPLC analysis of metabolites in the culture extracts of *Streptomyces* sp. SANK 60404 and its *rif* mutant TW-R50-13 in K medium, (**A**) 10 mL test tube scale; (**B**) 100 mL baffled flask scale.

Although no apparent changes were observed in K medium, the enhancement of some metabolites in the screening prompted the cultivation of TW-R50-13 in different media to investigate the metabolites produced in each medium. Thus, TW-R50-13 was cultivated in 4 different media (PC1, NMMP, GYM, and A-94964 producing medium) using baffled flasks (100 mL). PC1 and NMMP media were chosen based on their ingredients (types and composition of carbon and nitrogen source); GYM is the medium used for *rif* mutation

induction, and the A-94964 producing medium is a rich medium used for the isolation of the compound A-94964.

PC1 medium has an equal content of both complex carbon sources (starch and molasses) and nitrogen sources (meat extract and polypeptone). With these components, PC1 is a considerably rich medium. However, the cultivation of TW-R50-13 in this medium resulted in no detectable change. Interestingly, cultivating the *rif* mutant in the remaining media revealed metabolite **3** with significantly increased production compared with the parent strain (Figure 1-8).

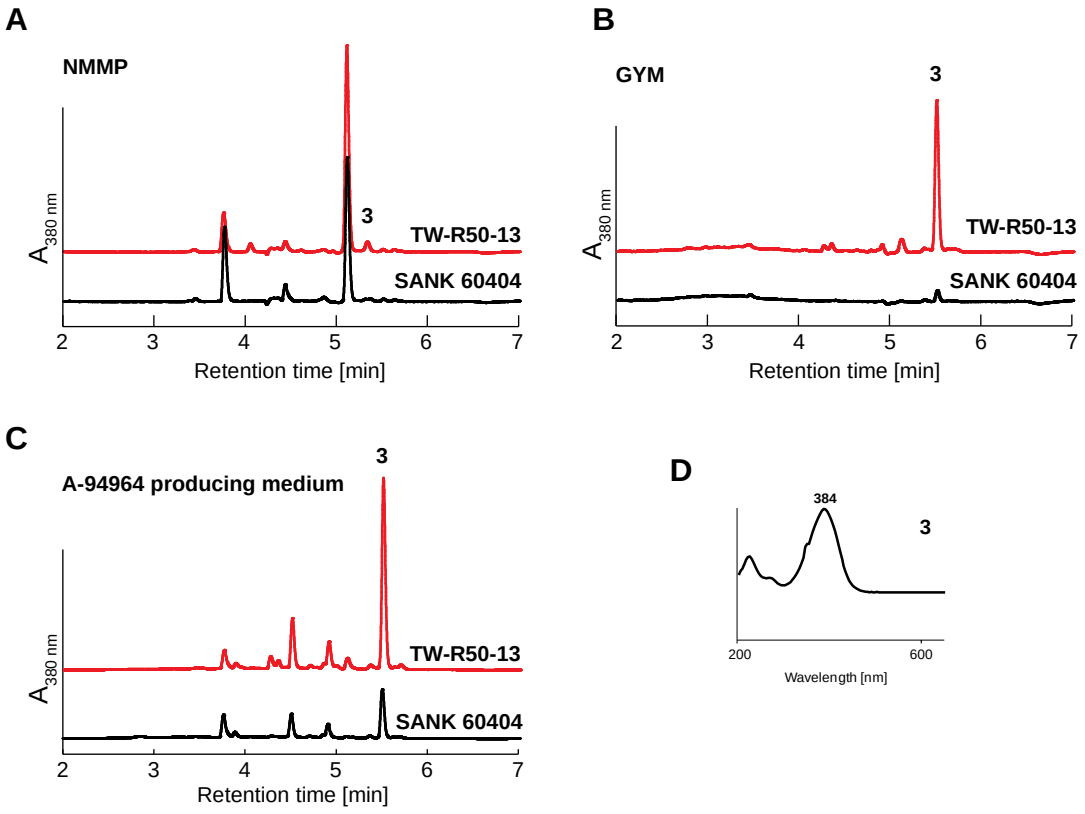


Figure 1-8. Comparative HPLC analysis of metabolites in the culture extract of TW-R50-13 and its wild-type strain in different media: (**A**) NMMP medium; (**B**) GYM medium; and (**C**) A-94964 producing medium. (**D**) UV spectrum of metabolite **3**.

Streptomyces sp. Sp080902JE-04 (S25)

The comparative HPLC analysis of *rif* mutants derived from *Streptomyces* sp. Sp080902JE-04 was performed only on mutants that emerged from selection with 50-fold MIC of rifampicin. Among the 15 *rif* mutants acquired, the mutant S25-50-1 showed significant overproduction of metabolite **4** in its HPLC profile compared with the wild-type strain (Figure 1-9A). However, metabolite **4** was not detected when S25-50-1 was cultivated at 100 mL scale (Figure 1-9B).

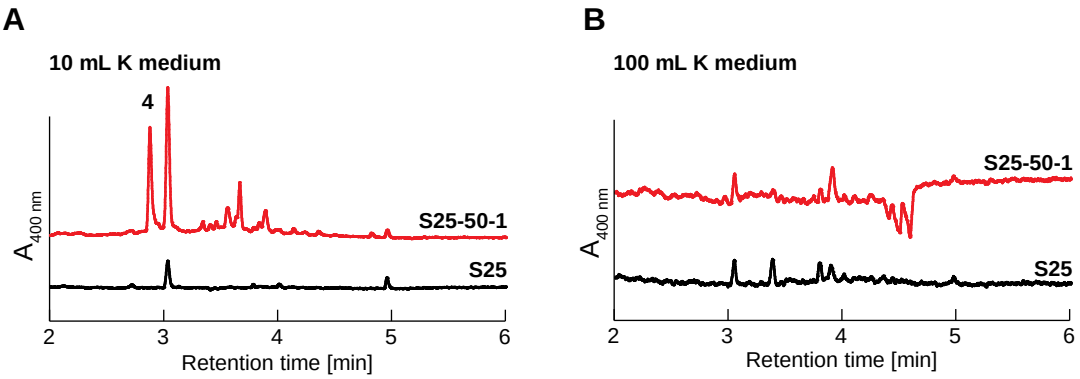


Figure 1-9. Comparative HPLC analysis of metabolites in the culture extracts of *Streptomyces* sp. Sp080902JE-04 (S25) and its *rif* mutant S25-50-1 in K medium. (**A**) 10 mL test tube scale; (**B**) 100 mL baffled flask scale.

To obtain metabolite **4** observed in the test tube scale, S25-50-1 was incubated in different conditions, which include the following: 1) cultivation in different media (A1, NMMP, GYM, A-94964 producing medium) in baffled flasks, 2) cultivation in K medium using a jar fermenter with an aeration level of 1.5 L min^{-1} and agitation at 200 rpm. All cultivations were maintained at 27°C, and time-course analyses were also performed for cultivation in the jar fermenter. Unfortunately, metabolite **4** was not observed in any of these cultivation conditions. Later, I found that the reproducibility of metabolite **4** was poor when I cultivated the *rif* mutant again at test tube scale with 3 repetitions (Figure 1-10).

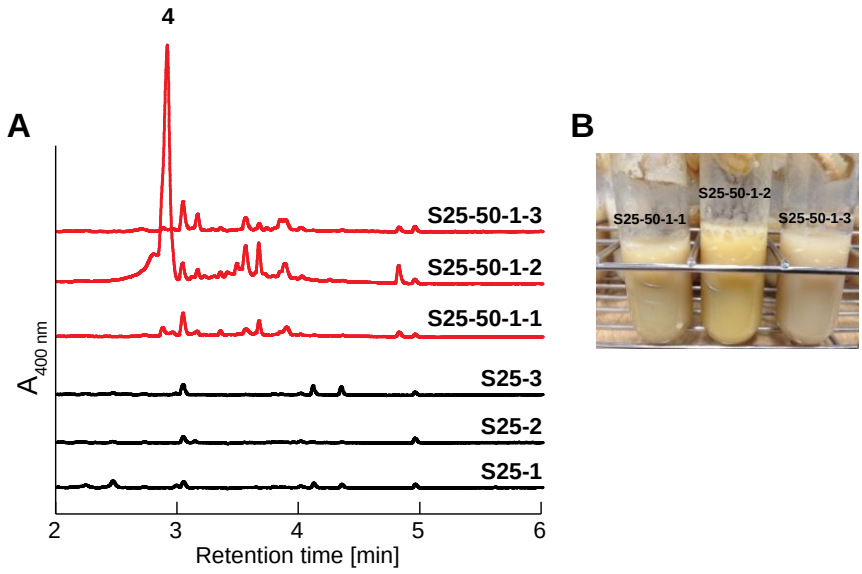


Figure 1-10. (**A**) HPLC analysis of S25-50-1 and its wild-type strain S25 in K medium (3 replicates). (**B**) Appearance of the S25-50-1 cultures after 4 days cultivation. Among the 3 replicates, only one test tube produced metabolite **4** (S25-50-1-2).

Streptomyces sp. TB090715SC-07 (S45)

As with the other *rif* mutants discussed above, comparative HPLC analysis of the mutants derived from S45 at 50-fold MIC yielded mutant S45-50-3, which overproduced a peak in its metabolic profile (Figure 1-11A). This metabolite **5**, however, was not produced when S45-50-3 was cultivated in a larger 100-mL scale. In the baffled flask cultivation, the mutant showed a slightly improved but similar metabolite profile to that of the wild-type strain (Figure 1-11B).

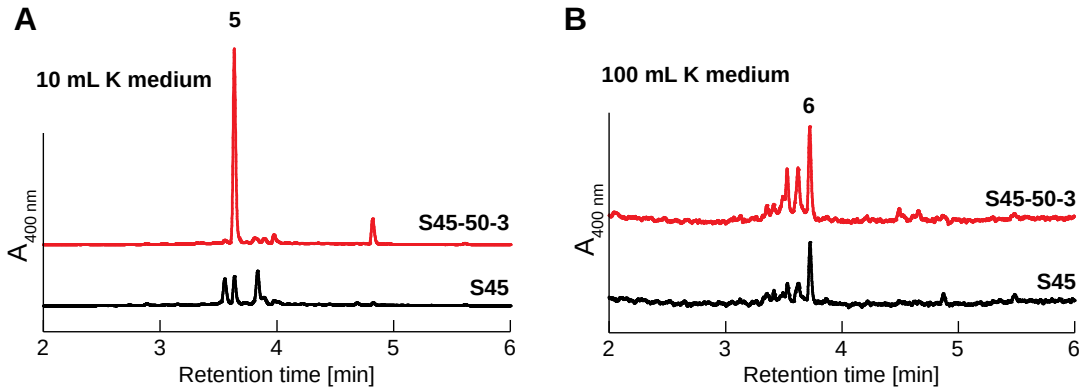


Figure 1-11. Comparative HPLC analysis of metabolites in the culture extracts of S46 and its *rif* mutant S45-50-3 in K medium, (**A**) 10 mL test tube scale; (**B**) 100 mL baffled flask scale.

Again, metabolite **5** failed to be reproduced in large scale even after a few attempts to cultivate S45-50-3 in 4 different media (A1, NMMP, GYM, and A-94964 producing medium) in baffled flasks or using K medium in a jar fermenter (aeration level 1.5 L min^{-1} and agitation at 200 rpm). The reproducibility of metabolite **5** in test tube scale, however, was good; production was detected in all 3 replicates (Figure 1-12). This observation indicated that metabolite **5** may only be produced in test tube scale.

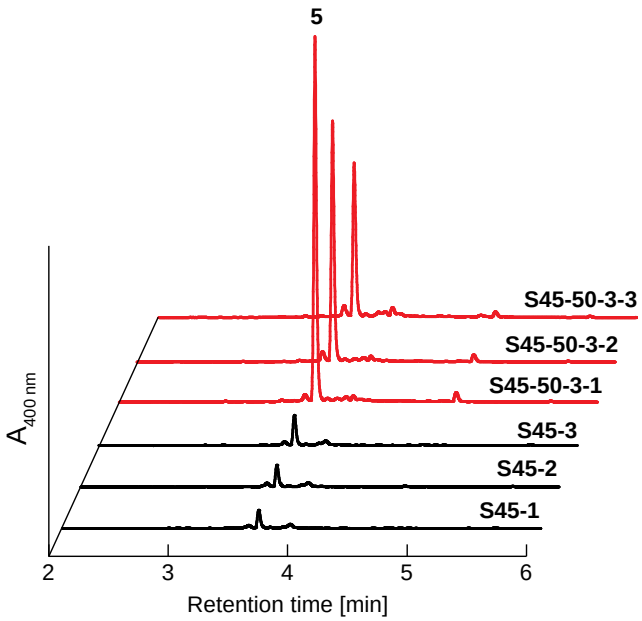


Figure 1-12. HPLC analysis of S45-50-3 and its wild-type strain S45 in test-tube scale using K medium (3 replicates).

Streptomyces sp. SoC090715LN-16 (S55)

Seven spontaneous rifampicin-resistant mutants developed from S55 upon induction with 50-fold MIC. Primary comparative screening of the *rif* mutants using K medium in both test tubes and baffled flasks revealed no apparent difference between the mutants and their wild-type S55. Unexpectedly, cultivation of the *rif* mutants in different media (100 mL) revealed a new metabolite (**7**) with UV/VIS (methanol) λ_{max} 280 nm and 376 nm that was produced only by the mutant S55-50-5 (Figure 1-12).

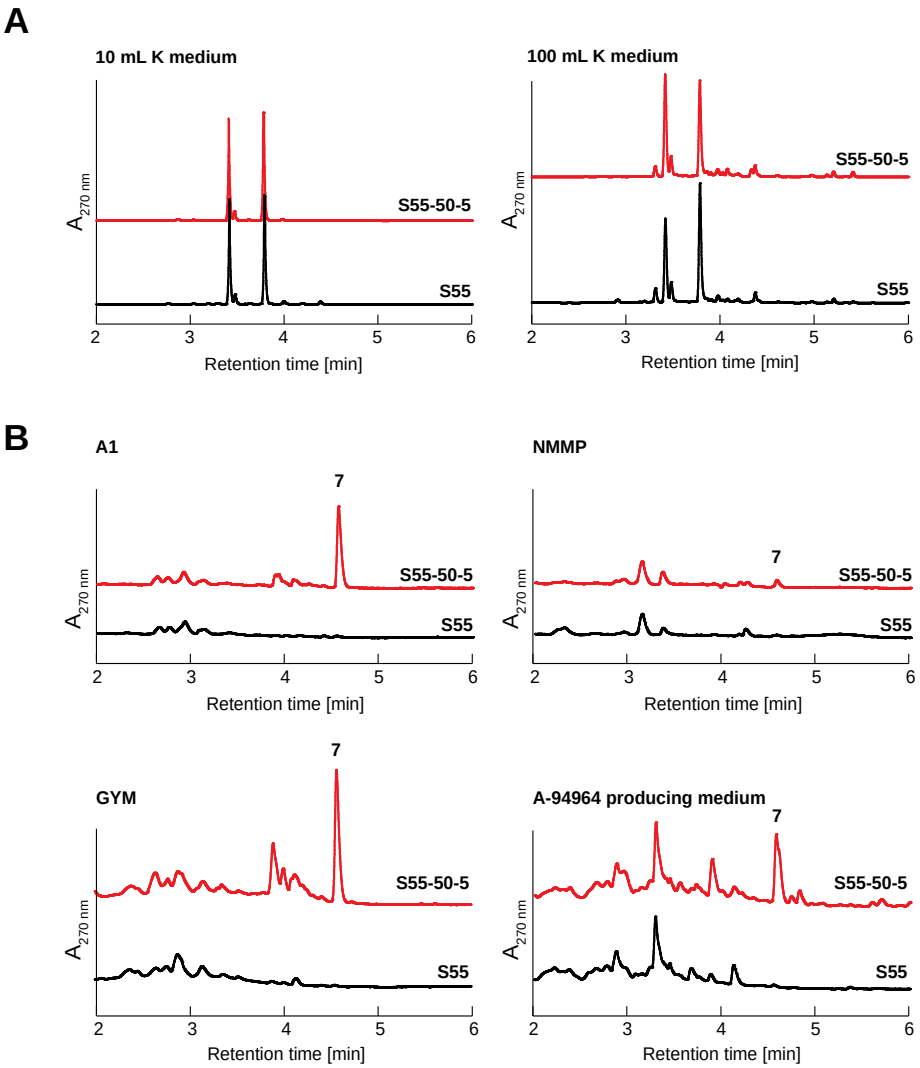


Figure 1-12. Comparative HPLC analysis of the *rif* mutant S55-50-5 along with its wild-type strain S55. (**A**) HPLC chromatograms of the mutant and S55 in K medium, showing no apparent changes. (**B**) HPLC chromatograms of the mutant in 4 different media showing overproduced metabolite **7**.

1.4 Conclusion

Conventional screening for novel bioactive compounds in actinomycetes often results in the rediscovery of known compounds. In contrast, recent genome sequencing has revealed that most of the predicted gene clusters for secondary metabolisms are not expressed under standard cultivation conditions. To explore the potential metabolites produced by these gene clusters, a cryptic gene activation strategy was implemented by screening mutants that acquire resistance to rifampicin.

Comparative metabolic analyses were performed on the selected 200 *rif* mutants. The screening revealed improved production of several metabolites by the mutants. Among them, I selected 4 overproduced metabolites (metabolites **1**, **3**, **6**, and **7**) detected in *rif* mutants WK2057-10-39, TW-R50-13, S45-50-3, and S55-50-5, respectively, for isolation and structural elucidation. The preliminary structure searches of these metabolites on Scifinder (American Chemical Society) and the Dictionary of Natural Products on DVD ver. 21:1 (CRC press) suggested that they are potential novel compounds.

CHAPTER 2

CHAPTER 2: Identification of the overproduced secondary metabolites

Overview

The comparative metabolic screening in chapter 1 produced a number of potential metabolites overproduced by the *rif* mutants. This chapter describes the isolation and structural elucidation of the 4 overproduced metabolites (metabolites **1**, **3**, **6**, and **7**) from their producing *rif* mutants WK2057-10-39, TW-R50-13, S45-50-3, and S55-50-5, respectively. Biological activity was also tested for novel compounds, and *rpoB* mutation analysis was performed on *rif* mutants producing the novel compounds.

2.1 Isolation and structural elucidation of metabolite 1 produced by *rif* mutant WK2057-10-39

Methods

To isolate metabolite **1**, *rif* mutant WK2057-10-39 was routinely grown at 30°C on a TSB agar plate containing 3% TSB and 2% agar. For fermentation, the *rif* mutant was pre-cultured in 10 mL of a medium containing 3% TSB supplemented with rifampicin (at MIC) in 50-mL test tubes on a reciprocal shaker at 30°C for 2 days. The seed medium (2 mL) was then inoculated into a 500-mL baffled flask containing 100 mL of K medium. WK2057-10-39 was cultivated in 4 L (100 mL × 40) of K medium at 27°C for 4 days on a rotary shaker (180 rpm). After fermentation, the culture was separated into broth and mycelia cake by centrifugation at 3900×*g* for 10 min. The broth was extracted with ethyl acetate, and the organic layer was dried over anhydrous Na_2SO_4 before concentration *in vacuo.* The mycelia cake was extracted with acetone, and after removal of the solvent, the residual solution was extracted with ethyl acetate

in the same manner as describe above. The combined extract (1.79 g) was then dissolved in methanol and passed through a Diaion® HP20 column (Mitsubishi Chemical, Tokyo). The metabolites were then eluted successively with 10% to 100% methanol. The 100% methanol fraction containing the target compound (170 mg) was separated using preparative HPLC at an isocratic elution of 30% acetonitrile with a flow rate of 8 mL min^{-1}. The separation was monitored at 280 nm. The purified samples were analyzed by NMR and by HRESI-MS using the conditions listed in Table 2-1 .

Table 2-1. HRESI-MS conditions.

Column	CAPCELLPAK C18 column (2.0 × 50 mm; Shiseido, Tokyo, Japan)
Solvent	Mobile phase A: Water + 0.1% formic acid Mobile phase B: acetonitrile + 0.1% formic acid 10% B for 1 min, a linear gradient of 10%–90% B for 4 min, 90% B for 1 min, and 10% B for 5 min
Flow rate	0.4 mL min^{-1}
Detector	Photo Diode Array (SPD-M20A, Shimadzu, Kyoto, Japan) High-resolution MS (Triple TOF® 5600, AB Sciex, Tokyo, Japan)
Ionization mode	ESI (+)
TOF accumulation time	0.1 s
Collision energy	40 V
MS2 accumulation time	0.04 s

Results and discussion

In the process of purifying metabolite **1,** 4 reddish-orange compounds were obtained. NMR analysis revealed that these compounds were previously reported by Kinjo et al. (1987) [48], Gerber. N (1967)[49] and Liu et al. (2008) [50]. The structures of **1** and its derivatives (**1a, 1b**, and **1c)** are shown in Figure 2-1, and their ^{1}H-NMR spectral data are summarized in Table 2-2 to 2-5.

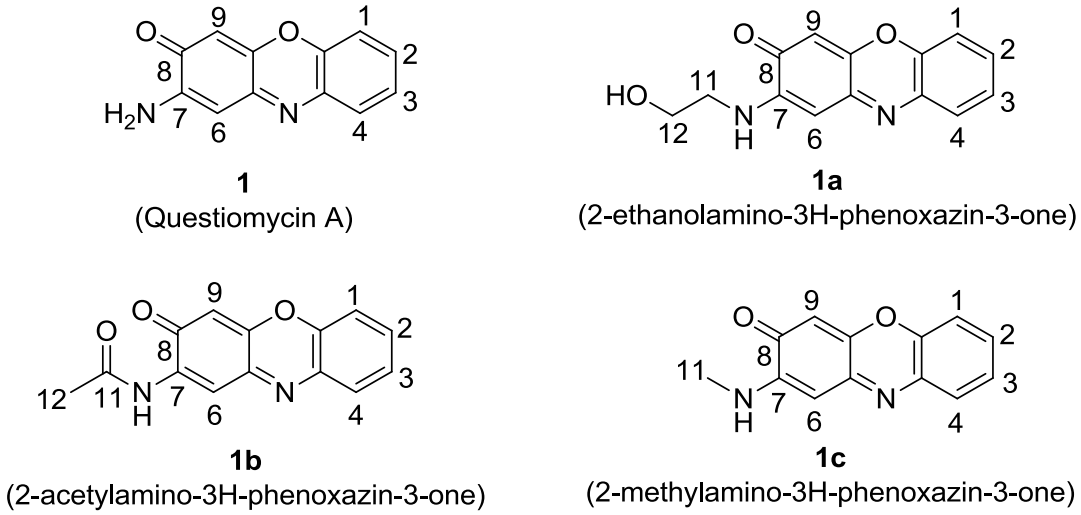


Figure 2-1. Structures of metabolite **1** and its derivatives.

Table 2-2. ^{1}H NMR spectral data of metabolite **1**[*].

Position	Found	Literature[51]
	$\delta_{H,\ mult}$ (*J* in Hz)[**]	$\delta_{H,\ mult}$ (*J* in Hz)[**]
1	7.47, dd (1.2, 8.4)	7.50, dd (1.9, 8.2)
2	7.43, ddd (1.2, 7.2, 8.4)	7.45, ddd (1.5, 6.8, 8.2)
3	7.35, ddd (1.2, 7.8, 8.4)	7.38, ddd (1.9, 6.8, 7.8)
4	7.67, dd (1.2, 7.8)	7.70, dd (1.5, 7.8)
6	6.34, s	6.36, s
9	6.33, s	6.34, s
11	6.80, brs	6.83, s

[*]red amorphous solid. UV/VIS (methanol) λ_{max} 228 nm and 438 nm. HRESI-MS *m/z* 213.0659 $[M+H]^+$ (calculated for $C_{12}H_9N_2O_2^+$: 213.0659).

[**]Measured in DMSO-d_6

Table 2-3. ^{1}H NMR spectral data of metabolite **1a***.

Position	Found	Literature[49]
	$\delta_{H,\ mult}$ (*J* in Hz)[**]	$\delta_{H,\ mult}$ (*J* in Hz)
1	7.37, m	N/A
2	7.41, ddd (1.2, 7.2, 7.8)	N/A
3	7.35, m	N/A
4	7.72, dd (1.2, 7.8)	N/A
6	6.24 s	N/A
9	6.40 s	N/A
NH	6.18, m	N/A
11	3.44, td (5.4, 10.8)	N/A
12	3.93, t (5.4)	N/A

[*]reddish-orange amorphous solid. UV/VIS (methanol) λ_{max} 224 nm and 444 nm. HRESI-MS *m/z* 257.0921 $[M+H]^+$ (calculated for $C_{14}H_{13}N_2O_3^+$: 257.0919).

[**]Measured in chloroform-*d*

N/A, Not available. Structure was elucidated by comparing with authentic sample synthesized from 2-hydroxy-3*H*-phenoxazin-3-one.

Table 2-4. ^{1}H NMR spectral data of metabolite **1b***.

Position	Found	Literature[50]
	$\delta_{H,\ mult}$ (*J* in Hz)**	δ_{H} (*J* in Hz)**
1	7.53, d (8.4)	7.42 (N/A)
2	7.60, ddd (1.8, 7.8, 8.4)	7.56 (N/A)
3	7.45, ddd (1.2, 7.8, 8.4)	7.44 (N/A)
4	7.83, dd (1.2, 7.8)	7.89 (N/A)
6	8.24, s	8.44 (N/A)
9	6.45, s	6.46 (N/A)
NH	9.68, s	8.60 (N/A)
12	2.19, s	2.29 (N/A)

*yellowish-brown needle. UV/VIS (methanol) λ_{max} 220 nm and 404 nm. HRESI-MS *m/z* 255.0764 $[M+H]^+$ (calculated for $C_{14}H_{11}N_2O_3^+$: 255.0764).

**Measured in DMSO-d_6

N/A, Not available

Table 2-5.^{1}H NMR spectral data of metabolite **1c**.

Position	Found	Literature[50]
	$\delta_{H,\ mult}$ (*J* in Hz)	δ_{H} (*J* in Hz)**
1	7.48, dd (1.2, 7.8)	7.50 (N/A)
2	7.43, ddd (1.2, 7.2, 7.8)	7.46 (N/A)
3	7.37, ddd (1.8, 7.2, 8.4)	7.40 (N/A)
4	7.66, dd (1.2, 7.8)	7.70 (N/A)
6	6.35, s	6.37 (N/A)
9	6.06, s	6.09 (N/A)
NH	7.19, m	7.19 (N/A)
11	2.18, d (5.4)	2.85 (N/A)

*brownish-red amorphous solid. UV/VIS (methanol) λ_{max} 220 nm and 440 m. HRESI-MS *m/z* 227.0815 $[M+H]^+$ (calculated for $C_{13}H_{11}N_2O_2^+$: 227.0815).

**Measured in DMSO-d_6

N/A, Not available

2.2 Isolation and structural elucidation of metabolite 3 produced by *rif* mutant TW-R50-13

Methods

The *rif* mutant TW-R50-13 was pre-cultured using 10 mL of medium containing 3% TSB supplemented with 0.2 μg mL^{-1} rifampicin in 50-mL test tubes on a reciprocal shaker at 30°C for 2 days. The seed medium (2 mL) was then inoculated into a 500-mL baffled flask containing 100 mL of A-94964 producing medium. The TW-R50-13 was cultivated in 6 L (100 mL × 60) of the producing medium at 27°C for 4 days on a rotary shaker (180 rpm). After fermentation, the culture was extracted as described above (Unit 2.1). The combined extract (1.16 g) was then applied to a silica gel column (Wakogel C-200) and eluted with chloroform–methanol (50:1). The eluate containing the target compounds (395 mg) was further purified using preparative HPLC with a Senshu Pak PEGASIL ODS column (20 × 250 mm, Senshu Scientific, Tokyo, Japan) using an isocratic elution of 70% acetonitrile at a flow rate of 8 mL min^{-1} and monitored at 365 nm to yield compound **3** (3 mg). Compound **3a** was found in one of the fractions during the purification of compound **3**. After an additional purification of the fraction by preparative HPLC using an isocratic elution of 70% methanol containing 0.1% trifluoroacetic acid (TFA) with a flow rate of 8 mL min^{-1}, compound **3a** was isolated as a colorless amorphous solid with a yield of 2 mg. Structural elucidation of **3** and **3a** were performed using NMR spectroscopy and HRESI-MS (condition listed in Table 2-1).

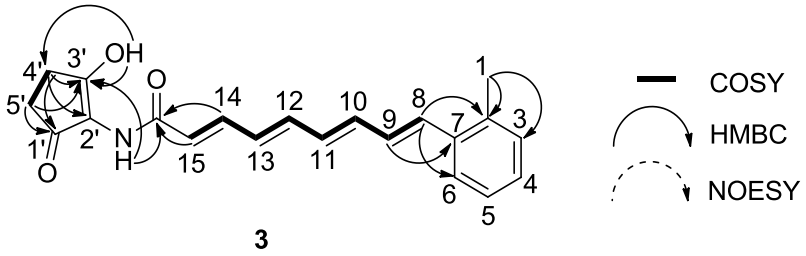
Results and discussion

Metabolite **3** was isolated as a yellow amorphous solid with UV absorption maximum at 384 nm. The structure was elucidated by HR/MS spectrometry and extensive NMR spectral analyses (Figure S1-S6, Table 2-6). Its molecular formula was established as $C_{21}H_{21}NO_3$ by HRESI-MS (*m/z* 336.1594 $[M+H]^+$, calculated for $C_{21}H_{22}NO_3^+$: 336.1600). The ^{1}H-NMR spectrum (600 MHz, chloroform-*d*) of **3** showed 17 signals, which were assigned to a singlet methyl signal (δ 2.37), 4 aromatic protons (δ 7.15 – 7.51), 8 olefinic protons (δ 6.00 – 7.38), 1 exchangeable NH signal (δ 7.56), 1 exchangeable OH signal (δ 13.70), and 2 methylene signals (δ 2.53 and 2.60). The ^{13}C-NMR and HSQC spectra confirmed the presence of 21 carbons, including a methyl (C-1), 6 aromatic carbons (C-2 to C-7), 8 olefinic carbons (C-8 to C-15), 2 carbonyls (C-16 and C-1'), 2 non-protonated *sp2* carbons (C-2' and C-3'), and 2 methylenes (C-4' and C-5').

The ^{1}H-^{1}H COSY correlations established that C-8 to C-15 form a conjugated polyene (Figure 2-2). These double bonds were determined to be in a *trans* configuration based on the coupling constants, J = 14.4 – 15.6 Hz. The HMBC correlations from H-8 (δ 6.91) and H-9 (δ 6.79) to C-6 (δ 125.2) and C-7 (δ 135.7), respectively, established the connection of the conjugated chain to an aromatic ring. The position of the methyl group in the aromatic ring was determined by the HMBC correlations observed from H-1 (δ 2.36) to C-3 (δ 130.7) and C-2 (δ 136.1), which indicated a methylbenzene moiety. The correlations of H-14 (δ 7.38) and H-15 (δ 6.00) to a C-N carbon C-16 (δ 165.9) showed a possible linkage of the unsaturated chain to an amide function. The presence of an amide was then confirmed based on the molecular formula and HMBC correlation from the amide proton (δ 7.56) to C-16.

The observation of a highly deshielded exchangeable 3'-OH proton (δ 13.70) and C-1' (δ 197.3), C-2' (δ 115.2), C-3' (δ 174.2), C-4' (δ 25.8), and C-5' (δ 32.2) indicated the presence of a 2-amino-3-hydroxycyclopent-2-enone (C_5N) moiety, which is a typical feature in several other compounds such as asukamycin and manumycin.[52, 53] The HMBC correlation of the

amide proton (δ 7.56) to a non-protonated *sp2* carbon C-3' (δ 174.2) connected the conjugated chain to the C_5N moiety.



3

Figure 2-2. Key correlations in the ^{1}H-^{1}H COSY and HMBC spectra of compound **3**.

While isolating compound **3**, I also attempted to purify other metabolites in the extract and identified compound **3a** in one of the purification fractions. Compound **3a** has a UV absorption maximum at 248 nm, and its molecular formula was determined to be $C_{17}H_{19}NO_3$ by HRESI-MS (*m/z* 286.1440 $[M+H]^+$, calculated for $C_{17}H_{20}NO_3^+$: 286.1438). A comparison of the NMR data of compound **3** and **3a** indicated that **3a** is most likely an analog of **3** that differs by the length of the conjugated chain (Figure S7-S12, Table 2-6). In the ^{1}H-NMR spectrum of **3a**, only 2 olefinic protons (δ 6.03 and 6.66) and an additional 4 methylene signals (δ 2.56 – 2.60) were observed.

The ^{1}H-^{1}H COSY spectrum established a conjugated double bond for C-8 and C-9, as shown in Figure 2-3. The double bond geometry was determined as a *trans* configuration based on the coupling constant between H-8 and H-9 (J = 15.6 Hz). The connection of C-9 and C-10 was determined by ^{1}H-^{1}H COSY.

The linkage between C-10 and C-11 was established by HMBC because of the crowded ^{1}H-^{1}H COSY signals in the methylene region. Furthermore, no HMBC signal was found to indicate the connection between the C_5N moiety and the carboxamide chain. To confirm this connection, a NOESY experiment was performed. As expected, NOESY correlation was observed from the amide proton (δ 7.54) to H-11 (δ 2.56) and the highly deshielded

exchangeable proton (δ 13.06). Moreover, NOESY correlation was observed between H-1 (δ 2.31) and H-8 (δ 6.66), suggesting that the methylbenzene moiety has a fixed rotation.

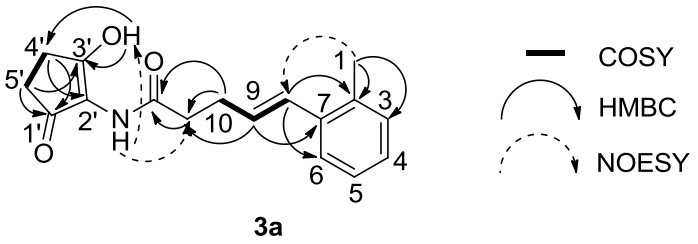


Figure 2-3. Key correlations in the ^{1}H-^{1}H COSY, HMBC, and NOESY spectra of compound **3a**.

To the best of my knowledge, compounds **3** and **3a** represent the first isolation of natural products having both the methylbenzene and C_5N moieties. Structure searches on SciFinder (American Chemical Society) returned results indicating that a 5-(2-methylphenyl)-4-pentenoic acid (**3b**) isolated from a terrestrial *Streptomyces* has close similarity to compound **3a** (Figure 2-4).[54] We believe that the observed overproduction most likely results from the increase in the transcription of genes involved in the biosynthesis of compounds **3** and **3a**. This assumption is based on a report discussing the effect of *rpoB* mutations on the transcription of genes involved in the secondary metabolite biosynthetic gene clusters.[38]

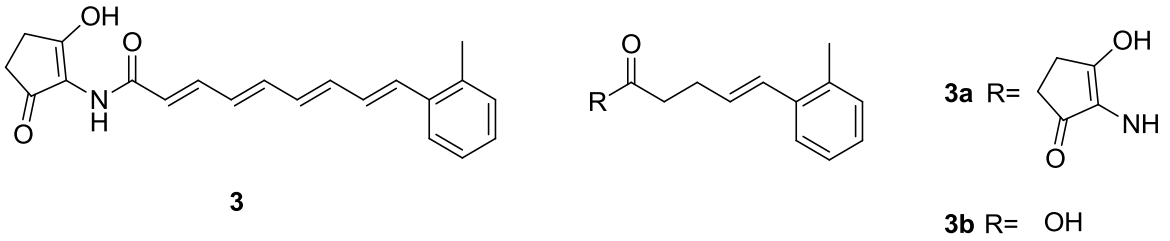


Figure 2-4. Structures of compounds **3** and **3a** that were identified in this study. 5-(2-Methylphenyl)-4-pentenoic acid (**3b**) is the previously reported compound.[54]

Table 2-6. ^{1}H (600 MHz) and ^{13}C (150 MHz) NMR spectral data for **3** and **3a** in chloroform-*d*.

	3		**3a**	
Position	**δ_C**	**δ_H, mult (*J* in Hz)**	**δ_C**	**δ_H, mult (*J* in Hz)**
1	19.9	2.37, s	19.8	2.31, s
2	136.1		135.2	
3	130.7	7.15, m	130.3	7.12, m
4	128.2	7.15, m	127.8	7.12, m
5	126.3	7.15, m	126.1	7.12, m
6	125.2	7.51, dd (2.4, 6.0)	125.7	7.37, dd (3.6, 6.5)
7	135.7		136.3	
8	133.2	6.91, d (15.6)	130.0	6.66, d (15.6)
9	129.6	6.79, dd (10.8, 15.6)	128.7	6.03, td (6.6, 15.6)
10	138.6	6.64, dd (10.8, 14.4)	29.1	2.60, m
11	132.0	6.43, dd (10.8, 14.4)	35.8	2.56, m
12	142.1	6.69, dd (10.8, 14.4)	173.0	
13	129.6	6.38, dd (11.4, 14.4)		
14	144.5	7.38, dd (11.4, 14.4)		
15	119.7	6.00, d (14.4)		
16	165.9			
NH		7.56, s		7.54, s
1'	197.3		197.4	
2'	115.2		114.8	
3'	174.2		173.4	
3'-OH		13.70, s		13.06, s
4'	25.8	2.60, m	25.6	2.60, m
5'	32.2	2.53, m	32.1	2.50, m

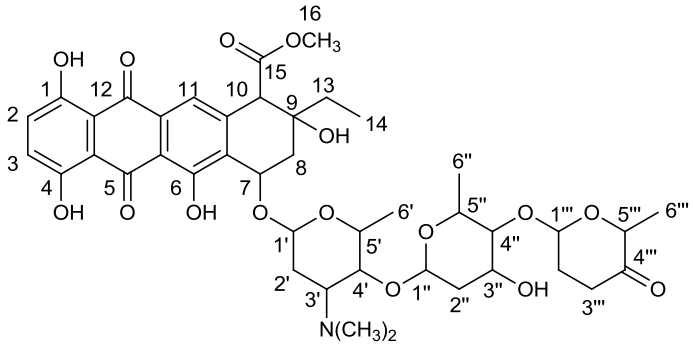
2.3 Isolation and structural elucidation of metabolite 6 produced by *rif* mutant S45-50-3

Methods

The *rif* mutant S45-50-3 was pre-cultured as mentioned above. Then, the pre-culture broth (2 mL) was inoculated into a 500-mL baffled flask containing 100 mL of K medium. S45-50-3 was cultivated in 3 L (100 mL × 30) of the medium at 27°C on a rotary shaker (180 rpm). After 4 days cultivation, the culture was extracted as described above (Unit 2.1). The combined ethyl acetate extract (1 g) was passed through a Diaion® HP20 column (Mitsubishi Chemical, Tokyo). The column was eluted with 100% water followed by 30%, 50%, 60%, 70%, 80%, 90%, and 100% aqueous methanol. The 90% and 100% methanol fractions containing the target compound were combined (89.7 mg) and further purified using preparative HPLC with a Senshu Pak PEGASIL ODS column (20 × 250 mm, Senshu Scientific, Tokyo, Japan). The purification was performed using an isocratic elution of 65% aqueous methanol containing 0.1% formic acid at a flow rate of 8 mL min^{-1} and monitored at 255 nm to yield compound **6**. Compound **6** was analyzed by NMR spectroscopy and HRESI-MS (conditions listed in Table 2-1).

Results and discussion

The molecular formula of compound **6** was determined as $C_{42}H_{53}NO_{16}$ by positive HRESI-MS (*m/z* 828.3425 $[M+H]^+$, calculated for $C_{42}H_{54}NO_{16}^+$: 828.3437). Preliminary structure search on SciFinder (American Chemical Society) using the molecular formula suggested that it is potentially an anthracycline. NMR analysis indicated that it is most likely a known compound called cinerubin A (Figure 2-5).[55] The ^{1}H-NMR spectrum (600 MHz) of **6** (Table 2-7) was compared with 2 reports: (1) Nakata et al. (1992) for the 4-ring aglycone and (2) Johdo et al. (1991) for the sugar moieties. [55, 56] Most of the proton signals recorded for **6** (chloroform-*d*) matched well with the reported data. However, the H-3' and N-$(CH_3)_2$ protons in the rhodosamine were not detected in the NMR spectra acquired for **6**. NMR analysis of compound **6** in methanol-d_4 revealed a broad singlet proton (δ 2.60) that correlated to a carbon with the chemical shift of δ 41.5 in HSQC, confirming the presence of the dimethylamino moiety. The H-3' adjacent to the dimethylamino moiety, however, was still not detected in NMR spectra obtained in methanol-d_4. Because this phenomenon may be due to the crowded environment of the H-3' and all other NMR data fit well with the reported data, I concluded that metabolite **6** is cinerubin A.



6
(Cinerubin A)

Figure 2-5. Structure of compound **6**.

Table 2-7. ^{1}H NMR spectral data for metabolite **6**.

Position	Found δ_H, mult (*J* in Hz)*	Literature[55, 56] δ_H, mult (*J* in Hz)*
1-OH	12.95, s	12.96, s
2	7.30, d (9.29)	7.31, d (9)
3	7.30, d (9.29)	7.29, d (9)
4-OH	12.82, s	12.80, s
6-OH	12.22, s	12.23, s
7	5.27, brs	5.27, brd (4.0)
8a	2.53, dd (15, 4.2)	2.52, dd (15.5, 4.0)
8b	2.26, d (15)	2.30, brd (15.5)
10	4.11, s	4.12, s
11	7.72, s	7.72, s
13a	1.74, m	1.75, m
13b	1.50, m	1.51, m
14	1.07, t (7.8)	1.09, t, (7.5)
16	3.69, s	3.70, s
1'	5.54, brs	5.52 brs
2'	1.91, brs	1.8 (N/A)
3'	N/A	2.15 (N/A)
4'	3.84, brs	3.74, brs
5'	4.00, q (6.6)	4.02, q (7)
6'	1.31, m	1.28, d (7)
N-$(CH_3)_2$	N/A	2.18, s
1''	5.06, t (6.6)	5.04, brd
2''a	2.10, m	2.10, dd (12, 5)
2''b	1.86, m	1.8 (N/A)
3''	4.21, m	4.12, br
4''	3.72, s	3.67 brs
5''	4.63, m	4.56 q (7)
6''	1.16, d (6.6)	1.17, d (7)
1'''	5.12, m	5.08, t (6)
2'''a	2.42, m	2.4, m
2'''b	2.12, m	2.1 (N/A)
3'''	2.46, m	2.46 (N/A)
5'''	4.48, q (6.6)	4.49, q (7)
6'''	1.31, m	1.32, d (7)

*Measured in chloroform-*d*, N/A, Not available

2.4 Isolation and structural elucidation of metabolite 7 produced by *rif* mutant S55-50-5

Methods

As described above, the *rif* mutant S55-50-5 was pre-cultured, and then the pre-culture broth (2 mL) was inoculated into a 500-mL baffled flask containing 100 mL of A-94964 producing medium. S55-50-5 was cultivated in 8 L (100 mL × 80) of the producing medium at 27°C on a rotary shaker (180 rpm). After 4 days cultivation, the culture was extracted as described above (Unit 2.1). The combined ethyl acetate extract was passed through a Diaion® HP20 column (Mitsubishi Chemical, Tokyo). The column was eluted with 100% water followed by 30%, 50%, 70%, 80%, 90%, and 100% aqueous methanol. The target compound **7** was detected in the 100% methanol fraction (78.7 mg), and the fraction was further purified using preparative HPLC with a Senshu Pak PEGASIL ODS column (20 × 250 mm, Senshu Scientific, Tokyo, Japan). The purification was performed using an isocratic elution of 75% aqueous methanol containing 0.1% formic acid at a flow rate of 8 mL min^{-1} and monitored at 280 nm to yield compound **7** (4.0 mg). Compound **7** was analyzed by NMR spectroscopy and HRESI-MS (conditions listed in Table 2-1).

Results and discussion

Metabolite **7** was isolated as a reddish-orange amorphous solid with the molecular formula $C_{24}H_{26}NO_8Cl$ as determined by HRESI-MS (*m/z* 492.1417 $[M+H]^+$, calculated for $C_{24}H_{27}NO_8Cl^+$: 492.1420). The presence the M+2 peak in the mass spectrum indicated the presence of a chlorine atom (Figure 2-6).

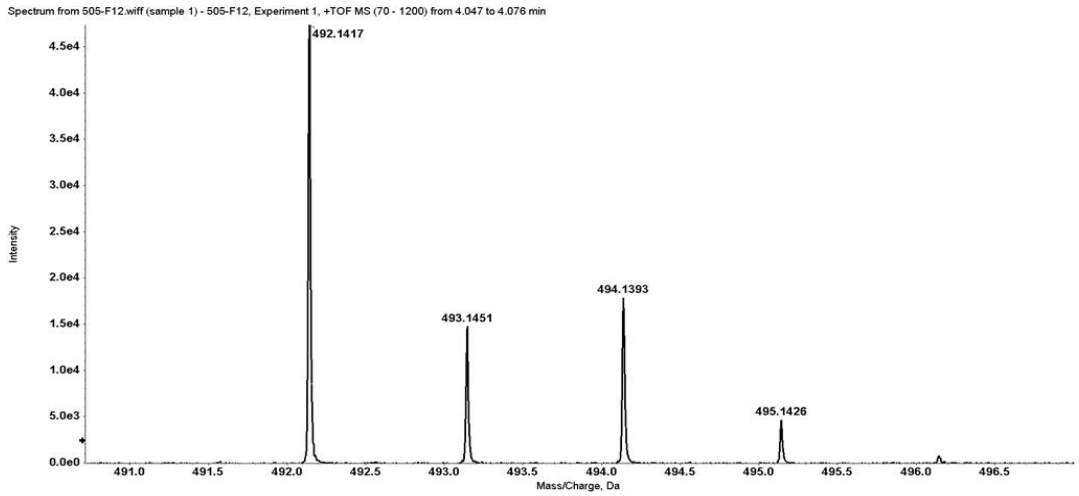


Figure 2-6. HRESI (+)-MS spectrum of compound **7**.

The structure of **7** was then elucidated by extensive NMR spectral analyses (Figure S13-S18, Table 2-8). The ^{1}H-NMR spectrum (600 MHz, chloroform-*d*) of **7** showed 17 signals, which were assigned to 3 aliphatic methylene units, 4 methine signals (δ 2.67, 4.89 and 2.89, 3.72), an aromatic proton (δ 7.63), an exchangeable NH (δ 6.79), a hydrogen-bonded phenolic OH group (δ 14.37), 2 aliphatic methyl signals (δ 0.98 and 1.41), an anomeric proton (δ 4.74), an aliphatic OH group (δ 4.22), and a methoxy signal (δ 3.40). The ^{13}C-NMR and HSQC spectra indicated the presence of 24 carbons, which include 2 carbonyls (C-1 and C-9), an aromatic methine (C-5), and 9 aromatic quaternary carbons, 2 of which (C-10 and C-11) could be assigned to aromatic carbons attached to oxygen atoms. In the aliphatic region, 3 methylenes (C-2, C-7 and C2'), 4 methine carbons (C-3, C-4, C-4', and

C-5'), 2 methyl carbons (C-12 and C-6'), an anomeric carbon (C-1'), a carbon connected to oxygen atom (C-3'), and a methoxy carbon (C-7') were detected.

The structure of **7** was then determined by 2D NMR experiments (COSY, HSQC, HMBC, and NOESY). Analysis of the ^{1}H-^{1}H COSY spectrum showed the correlations of a proton sequence from H-1' to H-6' indicating that it is a sugar moiety (Figure 2-7A). The HMBC correlations from H-4' to the methoxy carbon C-7' revealed that the sugar moiety has a structure similar to the 4-*O*-methyl-α-D-oliose (chromose A) in chromomycin.[57]

Another fragment of H-2/H-3/H-4 and H-3/H12 correlations was also observed in the ^{1}H-^{1}H COSY spectrum. The HMBC correlations of this partial structure to a carbonyl group (H-2 to C-1), an aromatic quaternary carbon (H-2 and H-4 to C-11a), and an aromatic methine (H-4 to C-5) indicated that it is attached to an aromatic ring. The presence of the aromatic ring was confirmed by the HMBC correlations observed from H-5 to C-11a and C-10a. HMBC correlations were also observed from the highly deshielded 11-OH to C-10a, C-11, and C-11a, indicating that it is a phenolic ring. However, the HMBC connectivity observed in compound **7** ended at the correlation from H-5 to C-6 (Figure 2-7A). Furthermore, no correlation was observed from H-7, leaving the structure partially elucidated.

To solve this problem, NMR analyses for compound **7** in different solvents were performed. As a result, the missing correlations from the H-7 signal were detected in NMR analyses taken in acetone-d_6 (Figure S19-S24, Table 2-8). The HMBC correlations from H-7 to C-6, C-6a, C-9, and C-9a established a 5-membered ring fused with the other 2 aromatic rings (Figure 2-7B). Furthermore, the relatively downshifted chemical shifts in both the ^{1}H and ^{13}C-NMR of the methylene C-7 indicated that it is attached to an amine moiety in the ring system.

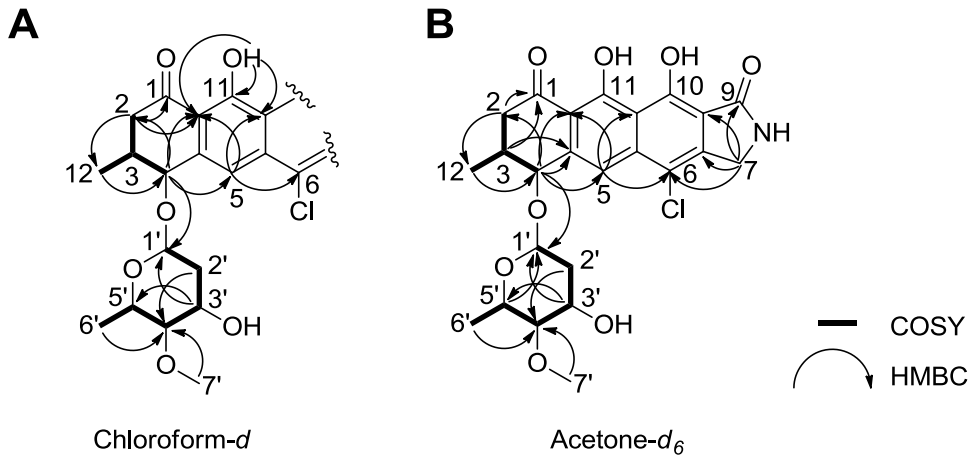


Figure 2-7. Key correlations in the ^{1}H-^{1}H COSY and HMBC spectra of compound **7** in (**A**) chloroform-*d* and (**B**) acetone-d_6.

The stereochemistry of compound **7** was determined relatively with a phase-sensitive NOESY experiment combined with the coupling constants of protons involved (Figure 2-8). The NOESY correlations were interpreted using the spectrum recorded in chloroform-*d* due to the overlapping of the acetone-d_6 solvent peak in the aliphatic region. To determine the absolute stereochemistry, the hydrolysis of compound **7** is necessary to obtain its aglycone followed by Mosher's method, targeting C-4. On the other hand, the absolute stereochemistry of the oliose sugar moiety could be determined by methylation at the hydroxyl substituents before comparing with a known standard. Due to the low yield of **7**, absolute stereochemistry determination is still in progress.

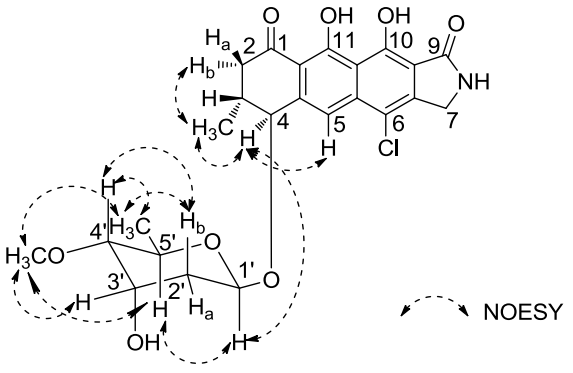


Figure 2-8. Key correlations in the NOESY spectrum of compound **7** in chloroform-*d*.

To the best of my knowledge, compound **7** is the first isolation of natural product having a tetracyclic structure with an isoindolinone skeleton (Figure 2-9A). Structure searches on SciFinder (American Chemical Society) returned results indicating that a compound called lactonamycin isolated from *Streptomyces rishiriensis* MJ773-88K4 has the closest similarity to compound **7** (Figure 2-9B).[58] As mentioned above, the sugar moiety found in compound **7** has a structure similar to that of chromose A in chromomycin.

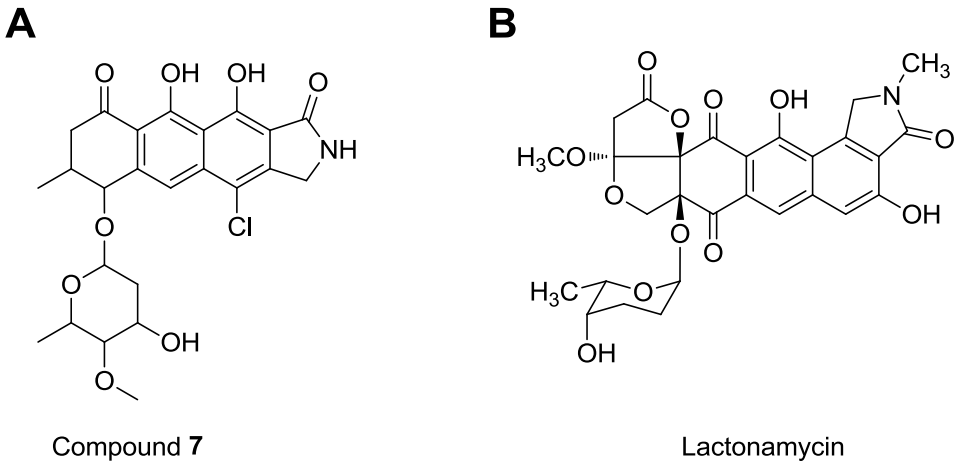


Figure 2-9. (A) Structure of compound **7** elucidated in this study. **(B)** Lactonamycin is a previously reported compound with the closest similarity to compound **7**.[58]

Table 2-8. ^{1}H (600 MHz) and ^{13}C (150 MHz) NMR spectral data for **7** in chloroform-*d* and acetone-d_6.

	Chloroform-*d*		Acetone-d_6	
Position	**δ_C**	**δ_H, mult (*J* in Hz)**	**δ_C**	**δ_H, mult (*J* in Hz)**
1	204.3		204.3	
2a	40.3	3.28, dd (4.8, 18)	40.8	3.17, dd (4.8, 18)
2b		2.43, dd (2.4, 18)		2.49, dd (5.4, 18)
3	33.8	2.67, m	34.2	2.57, m
4	76.2	4.89, d (3.6)	76.7	4.91, d (5.4)
4a	138.6		138.5	
5	116.6	7.63, s	115.3	7.68, s
5a	136.7		138.2	
6	113.5		112.0	
6a	143.0		143.7	
7	48.6	4.93, m	48.4	4.96, s
NH		6.79 brs		8.31 brs
9	171.7		171.2	
9a	117.5		118.2	
10	152.3		152.6	
10a	115.9		115.1	
11	164.0		163.3	
11a	109.4		109.3	
11-OH		14.37, s		
12	17.1	0.98, d (7.2)	16.7	1.03, d (7.2)
1'	94.0	4.74, brd (9)	95.1	4.86, dd (1.8, 9.6)
2'a	36.6	2.02, brd (13.8)	37.3	1.97, ddd (1.8, 3.6, 13.8)
2'b		1.69, ddd (3, 9, 13.8)		1.66, ddd (3, 9.6 13.8)
3'	64.0	4.22, dd (3, 3)	63.0	4.26, dd (3, 3)
4'	82.5	2.89, dd (3, 9)	82.8	2.87, dd (3, 9)
5'	68.2	3.72, dq (6, 9)	68.2	3.76, dq (6, 9)
6'	18.4	1.41, d (6)	17.9	1.31, d (6)
7'	57.4	3.40, s	56.0	3.32, s

2.5 Biological activity of the novel compounds

Methods

All biological assays discussed in this study were kindly conducted by Dr. Miho Izumikawa (National Institute of Advanced Industrial Science and Technology (AIST)). The antimicrobial activity of **3**, **3a**, and **7** was determined against *Escherichia coli*, *Micrococcus luteus*, *Staphylococcus aureus*, and *Bacillus subtilis*. Different concentrations of the novel compounds dissolved in DMSO were added to a 384-well microplate (Greiner, #781186) at 0.2 μL/well. Then, the culture solution for each test strain (cultivated up to OD_{620} ~0.5) was diluted 2,000-fold with LB medium, and 20 μL/well were added to each sample-containing well. After cultivation at 37°C for 24 h, the viability of cells in the culture was determined by measuring absorbance at 620 nm using an EnVision® microplate reader (PerkinElmer). The cytotoxicity of **3**, **3a**, and **7** against human ovarian adenocarcinoma SKOV-3, malignant pleural mesothelioma Meso-1, and T lymphoma Jurkat cells was evaluated using a WST-8 (2-(2-methoxy-4-nitrophenyl)-3-(4-nitrophenyl)-5-(2,4-disulfophenyl)-2*H*-tetrazolium, monosodium salt) colorimetric assay (Cell Counting Kit-8, Dojindo; Kumamoto, Japan) according to a previously described procedure.[59]

Results and discussion

Compounds **3** and **3a**

Compounds **3** and **3a** were tested for their cytotoxicity against SKOV3 (human ovarian carcinoma cell line), Meso-1 (malignant pleural mesothelioma), and Jurkat (T lymphoma) cells and their antimicrobial activity against several Gram-positive and Gram-negative bacteria. However, neither **3** nor **3a** exhibited any biological activity in these assays, even at 100 μM (Figure 2-10).

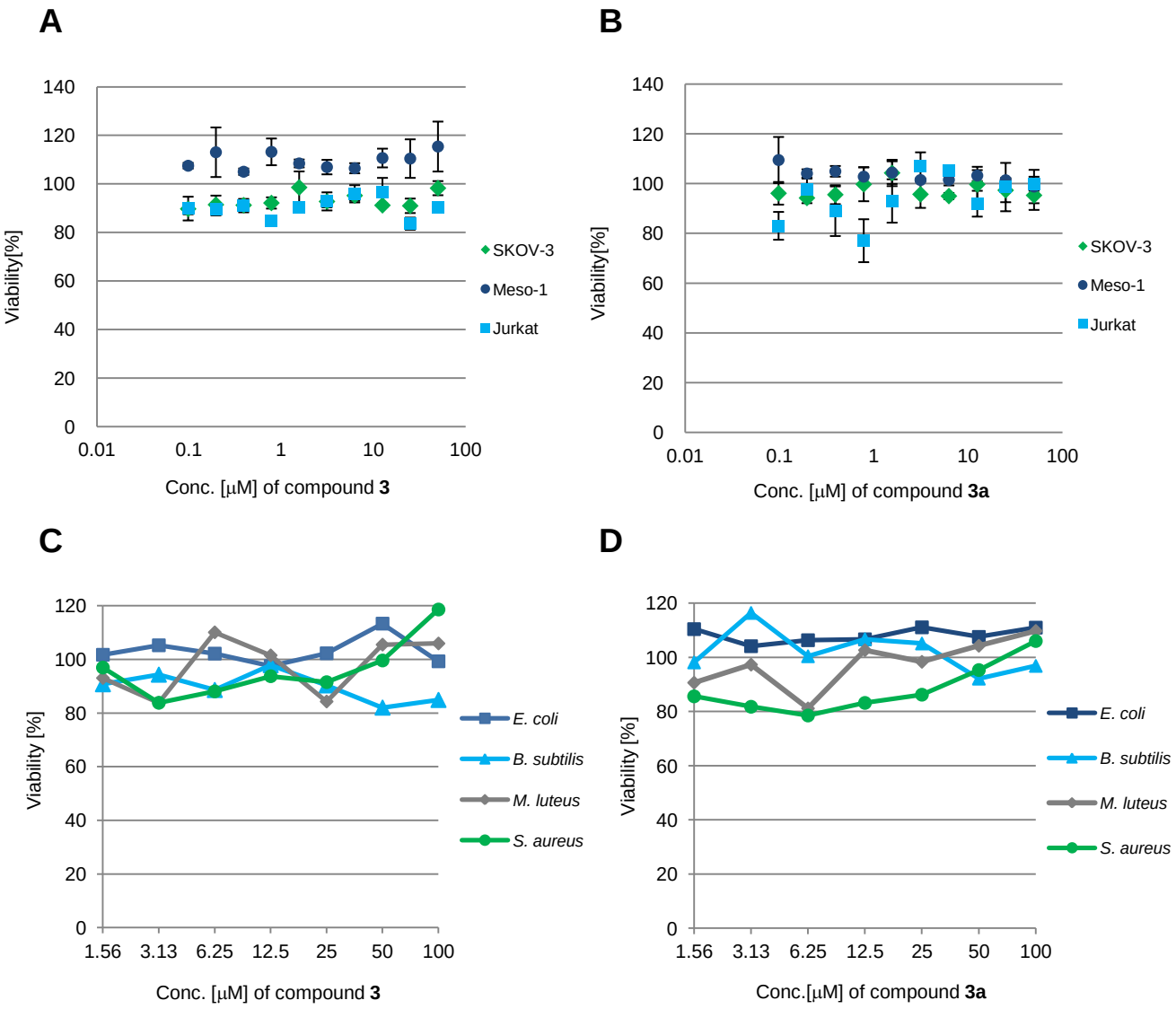


Figure 2-10. Cytotoxicity against different cell lines of (**A**) compound **3** and (**B**) compound **3a**. Antimicrobial activity of (**C**) compound **3** and (**D**) compound **3a**.

Compound **7**

As above, compound **7** was tested for cytotoxicity against SKOV-3, Meso-1, and Jurkat cells and for antimicrobial activity. In the cytotoxicity assay, compound **7** showed activity against all cell lines tested with IC_{50}s of 19.2 μM (SKOV-3), 7.35 μM (Meso-1), and 8.33 μM (Jurkat) (Figure 2-11A). Compound **7** also had inhibitory activity against *Staphylococcus aureus* with an IC_{50} of 11.6 μM (Figure 2-11B and C).

Tetracyclics, such as anthracycline and tetracycline, are well known classes of antibiotics. Despite the potency of these drugs that bind specifically to the bacterial ribosomal subunits, drug resistance (efflux pump type) has developed as a consequence of their overuse.[60] Compound **7** has a unique tetracyclic skeleton that may offer new binding possibilities to the active sites of drug targets. Because compound **7** exhibited potent biological activities, I name this compound "isoindolinomycin."

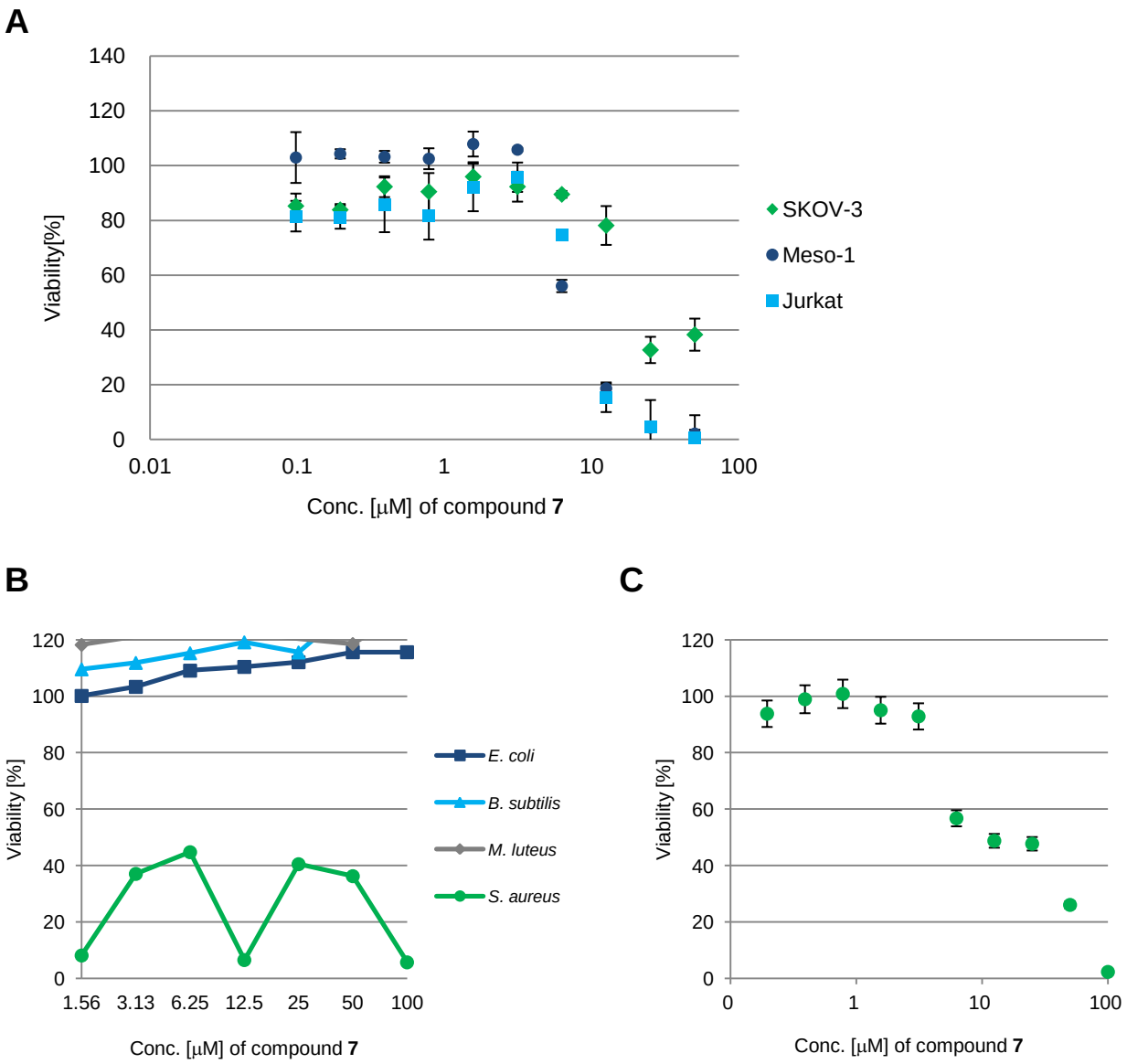


Figure 2-11. Biological assays tested on compound **7**. (**A**) Cytotoxicity against different cell lines. (**B**) Antimicrobial activity against Gram-positive and Gram-negative bacteria. (**C**) Antimicrobial activity against *Staphylococcus aureus* in 3 replicates.

2.6 Mutation analysis of *rif* mutants producing the novel compounds

Methods

Mutation analysis of the *rpoB* gene was performed in three parts using 3 sets of different primers (Table 2-7). PCR amplification was performed with the Expand High Fidelity PCR System (Roche Diagnostics) using genomic DNA as a template. The amplified PCR products were then purified with the QIAquick Gel Extraction Kit (Qiagen) and ligated into pT7 Blue vector using Ligation High (Toyobo, Osaka, Japan). *Escherichia coli* DH5α were transformed with the resulting plasmids before being purified with a GenElute Plasmid Miniprep Kit (Sigma-Aldrich), and sequencing of the insert DNAs was outsourced (Greiner Bio-One, Tokyo).

Table 2-7. Primers used for *rpoB* gene mutation analysis.

Gene	Primer	Oligonucleotide sequence (5' to 3')
rpoB-P1	Forward	GCCCTCGGAAGGACCCCCTC
	Reverse	CGGCGTGATCGCCTCGACGTC
rpoB-P2	Forward	TCGGCGAGCTCATCCAGAACC
	Reverse	AGGACCTCCTCGGAGACGTTC
rpoB-P3	Forward (SANK 60404)	GTCGACGCCCGTGACACCAAGCTC
	Reverse (SANK 60404)	GGAGGCCCCGACCAGACCCG
	Forward (S55)	GTCGACGCCCGTGACACCAAGCTG
	Reverse (S55)	GTGGCGCAGGTCTCCGGACTCC

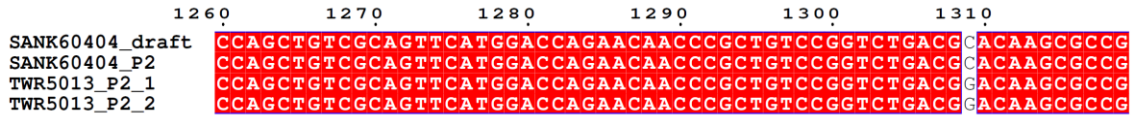
Results and discussion

Sequence analysis of the *rpoB* gene of TW-R50-13 identified a single C1309G mutation (Figure 2-12A), which causes a change of the amino acid at position 437 from histidine to aspartic acid (Figure 2-12B). This point mutation has been reported by Hosaka et al. (2009); a *rif* mutant (KO-915) with the same C1309G mutation demonstrated the increased production of antibacterial compounds that were poorly produced by the wild-type strain (*Streptomyces* sp. 631689).[41]

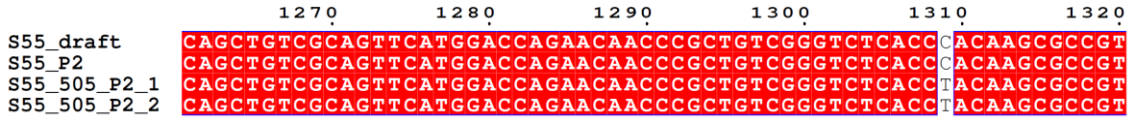
Mutation analysis of the *rpoB* gene of S55-50-5 also identified a point mutation that occurred at position 1309. The mutation was a C1309T (Figure 2-12A), which corresponds with a histidine to tyrosine alteration at position 437 (Figure 2-12B). In a recent report by Tanaka et al. (2013), a list of *rpoB* mutations that are effective for antibiotic overproduction has been summarized.[38] In the report, they emphasize that mutations at the amino acid residue 437, notably H437Y, are often effective in a wide variety of actinomycetes.

A

TW-R50-13

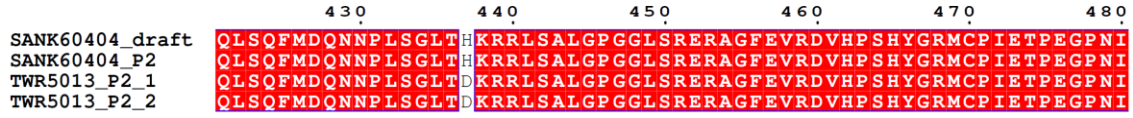


S55-50-5



B

TW-R50-13



S55-50-5

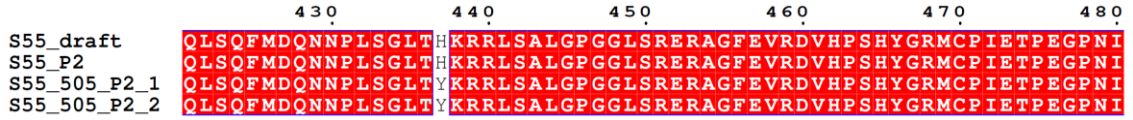


Figure 2-12. The position of *rpoB* gene mutation of the *rif* mutants TW-R50-13 and S55-50-5. (**A**) Nucleotide sequence (**B**) Amino acid sequence. Draft, Draft genome sequence; P2, Sequence of the wild-type strain amplified using *rpoB*-P2 primer; P2_1 and P2_2, Sequence of the *rif* mutant amplified using *rpoB*-P2 primer. Numbering begins at the start codon in *Streptomyces coelicolor* A3(2) [NP_628815].

2.7 Conclusion

After the comparative metabolic screening performed in chapter 1, isolation and structural elucidation were performed for the selected metabolites **1**, **3**, **6**, and **7**. The isolation of metabolite **1** from WK2057-10-39 resulted in 4 known phenoxazinones (compound **1-1c**). The biosynthesis of phenoxazinone is catalyzed by a GriF (tyrosinase), which oxidizes 3-amino-4-hydroxybenzaldehyde (3,4-AHBAL) to an *o*-quinone imine that then non-enzymatically reacts with a second *o*-quinone imine to produce a phenoxazinone.[61] The overproduction of phenoxazinones observed in WK2057-10-39 may be due to the up-regulation in the transcription of genes involved in producing 3,4-AHBAL.

Purification from a 6-L culture of TW-R50-13 yielded 2 novel compounds (compound **3** and **3a**) consisting of the C_5N and methylbenzene moieties. The C_5N ring formation has been described previously; the biosynthetic genes involved are 5-aminolevulinate (ALA) synthase and acyl-CoA ligase.[62] In a recent report, ALA synthase transcription was upregulated in the *rif* mutants derived from *Streptomyces griseus* IFO13189.[38] This result strongly indicated that the enhanced production of compounds **3** and **3a** is most likely due to the higher expression of ALA synthase in TW-R50-13.

Metabolite **6** was identified as cinerubin A, whereas metabolite **7** was an unknown compound with a novel skeleton. Both compounds **6** and **7** were tetracyclic compounds; type-II polyketide synthase catalyzed the formation of **6**[63] and most likely compound **7**. Compound **7** showed potent cytotoxicity against all cell lines tested and antimicrobial activity against *Staphylococcus aureus*. Due to its bioactivity, compound **7** was named isoindolinomycin. In contrast, no biological activity was detected for compound **3** or **3a** in any assays performed. Both mutants TW-R50-13 and S55-50-5 have a point mutation in their *rpoB* gene at position 437, which is a known mutation point associated with the activation of secondary metabolite production.

In the next chapters, the biosynthesis of compounds **3**, **3a** (chapter 3), and isoindolinomycin (chapter 4) will be discussed. Although the biosynthesis of the C_5N unit in **3** and **3a** has been reported previously, the formation of the methylbenzene moiety is intriguing because it is rarely found in natural products, and our knowledge of biosynthetic pathways for aromatic ring in actinomycetes is limited. In addition, the biosynthesis of isoindolinomycin requires attention. The formation of the 5-membered ring and the attachment of the methyl moiety C-12 to the tetracyclic core structure prompted me to investigate their biosynthetic mechanism.

CHAPTER 3

CHAPTER 3: Biosynthesis of the new compounds **3** and **3a** isolated from the *rif* mutant TW-R50-13

Overview

There are two distinct features present in both compounds **3** and **3a**: the C_5N and methylbenzene moieties (Figure 3-1).

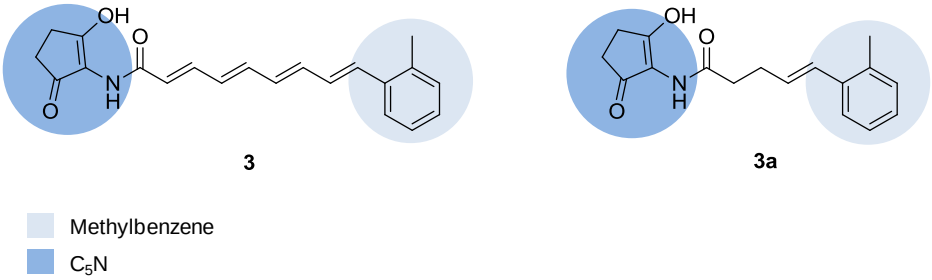


Figure 3-1. The structures of compound **3** and **3a** consisting of 2 distinct moieties.

The biosynthetic pathway for the C_5N (2-amino-3-hydroxycyclopent-2-enone) ring has been elucidated.[62] In this pathway, a PLP-dependent 5-aminolevulinate (ALA) synthase generates ALA from succinyl-CoA and glycine, and then ALA is converted to ALA-CoA by an acyl-CoA ligase and cyclized by ALA synthase to produce the C_5N moiety (Figure 3-2).

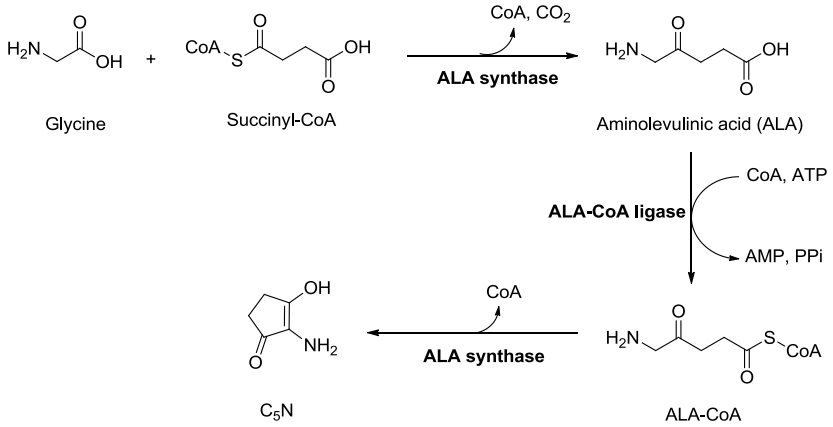


Figure 3-2. Biosynthetic pathway of the C_5N ring.

The formation of the methylbenzene moiety, however, remains poorly understood. The aromatic moiety and the unsaturated chain attached to compounds **3** and **3a** suggested that the methylbenzene moiety is a product of polyketide synthase (PKS). However, to the best of my knowledge, no currently known bacterial PKS is capable of producing this methylbenzene moiety. The orsellinic acid moiety in avilamycin is the structure of PKS origin with the closest structural similarity to the methylbenzene moiety in **3** and **3a** (Figure 3-3).[64] Therefore, in this chapter, biosynthetic studies on the novel compounds **3** and **3a** are described using *in silico* genome analysis, ^{13}C-labeling experiments, and gene inactivation.

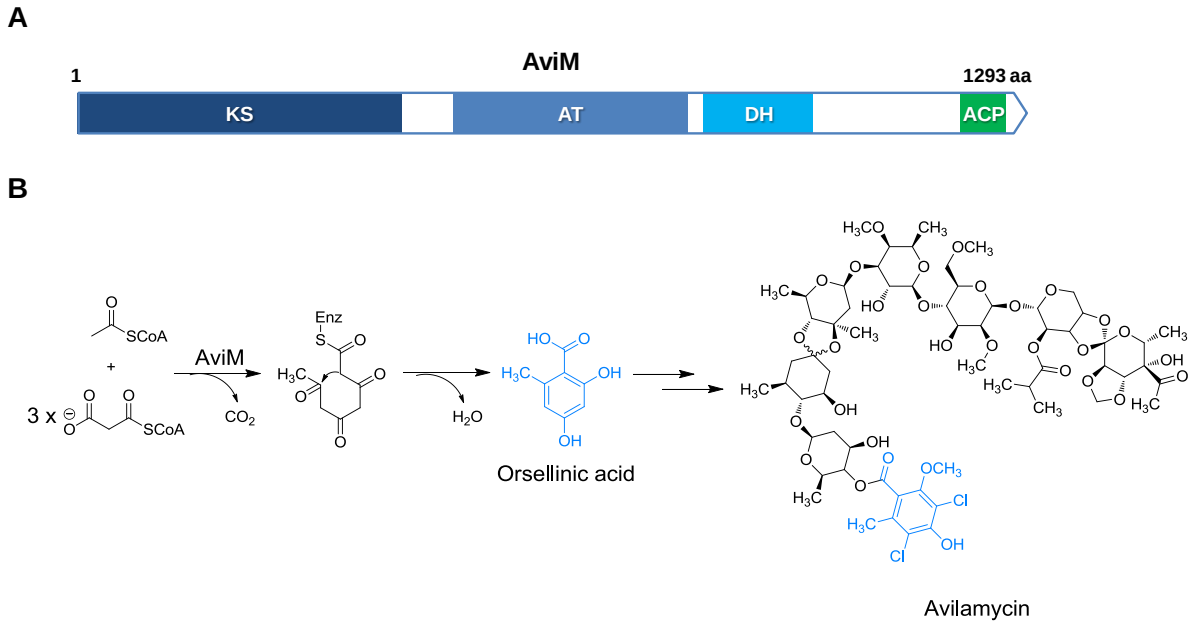


Figure 3-3. Mechanism of iterative type I PKS for aromatic polyketide biosynthesis. **(A)** Domain organization of AviM. **(B)** Proposed biosynthetic mechanism of orsellinic acid catalyzed iteratively by AviM.

3.1 *In silico* genome analysis and gene inactivation: Identification of the C_5N gene cluster

Methods

In silico genome analysis

The draft genome sequence of *Streptomyces* sp. SANK 60404 was acquired by using the Genome Analyzer IIx system (Illumina) at Takara Bio, Kyoto.[65] The sequence reads were assembled with the Velvet *de novo* short-reads assembler.[66] Genome sequences were annotated with protein BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi), FramePlot 4.0 (2007) (http://nocardia.nih.go.jp/fp4/)[67] and 2ndFind (http://biosyn.nih.go.jp/2ndfind/). A graphical representation of the PKS and NRPS domain organization was obtained from amino acid sequences using free databases such as PKS/NRPS-domain finder ver. 1.0 (http://avermitilis.ls.kitasato-u.ac.jp/pfam-pks/).

Inactivation of *orf1* in TW-R50-13 by in-frame deletion

The disruption of *orf1* in TW-R50-13 was accomplished by in-frame deletion (Figure 3-4). The construction of the *orf1* (1,584 bp) disruption plasmid was performed by amplifying a 2,001-bp DNA fragment containing the upstream region of *orf1* and a 2,025-bp DNA fragment containing the downstream region of *orf1*. The primers used for the amplification are listed in Table 3-1.

Table 3-1. Primer sets used to construct the *orf1* disruption plasmid.

Primer	Oligonucleotide sequence (5' to 3')*	Restriction enzyme
Del1-up-Fw	GGGAAGCTTCTGATTGGGGCAGTCGAGGC	*Hin*d III
Del1-up-Rv	GGGTCTAGACCCCTTGCCTACTCCGCG	*Xba* I
Del1-dw-Fw	GGGTCTAGAGGAGACATCGGCTGTCTCGAC	*Xba* I
Del1-dw-Rv	GGGAAGCTTGGTGAGGAGCACCGTGC	*Hin*d III

*Restriction enzyme site underlined.

Each fragment was cloned into the pBluescript II KS vector to produce Del1up-Pblue and Del1dw-Pblue, and the plasmids were sequenced (Greiner Bio-One, Tokyo). Then, the Del1up-Pblue plasmid was digested with *Hin*d III and *Xba* I, whereas Del1dw-Pblue was digested with *Hin*d III and *Kpn* I (restriction enzyme site on the vector). First, both fragments were simultaneously cloned into pBluescript II KS digested with *Hin*d III and *Kpn* I to give Del1updw-Pblue. This plasmid was then digested with *Hin*d III, and the fragment was cloned into pUC118apr[68] to give pUC118aprΔ*orf1*.

To obtain the *orf1* disruptant, TW-R-50-13 was transformed with pUC118aprΔ*orf1*. Single-crossover transformants were obtained as colonies that grew on R2YE plates containing 25 μg mL^{-1} apramycin. The single-crossover transformants obtained were confirmed by PCR, and one was used for the protoplast preparation. The prepared protoplasts were regenerated on R2YE plates. Each regenerated colony was streaked on TSB plates with or without apramycin, and an apramycin-sensitive colony was selected to obtain the *orf1* knockout mutant TW-R50-13Δ*orf1*. Successful deletion was confirmed by the PCR amplification (Go Taq® Green Master Mix, Promega) of the *orf1* gene using Del1-check-Fw (5'-TTGGTGGCGTGCAAGTCAGAGTTATGTC-3') and Del1-check-Rv (5'-CTAGGAAAATTGCGACTTCAGGAGCACC-3'). To assess the production of compound **3** and **3a**, TW-R50-13Δ*orf1* deletion mutants were grown in the A-94964 producing medium (100 mL) at 27°C for 4 days before being extracted and analyzed by HRESI-MS as described in Unit 1.3 and Table 2-1, respectively.

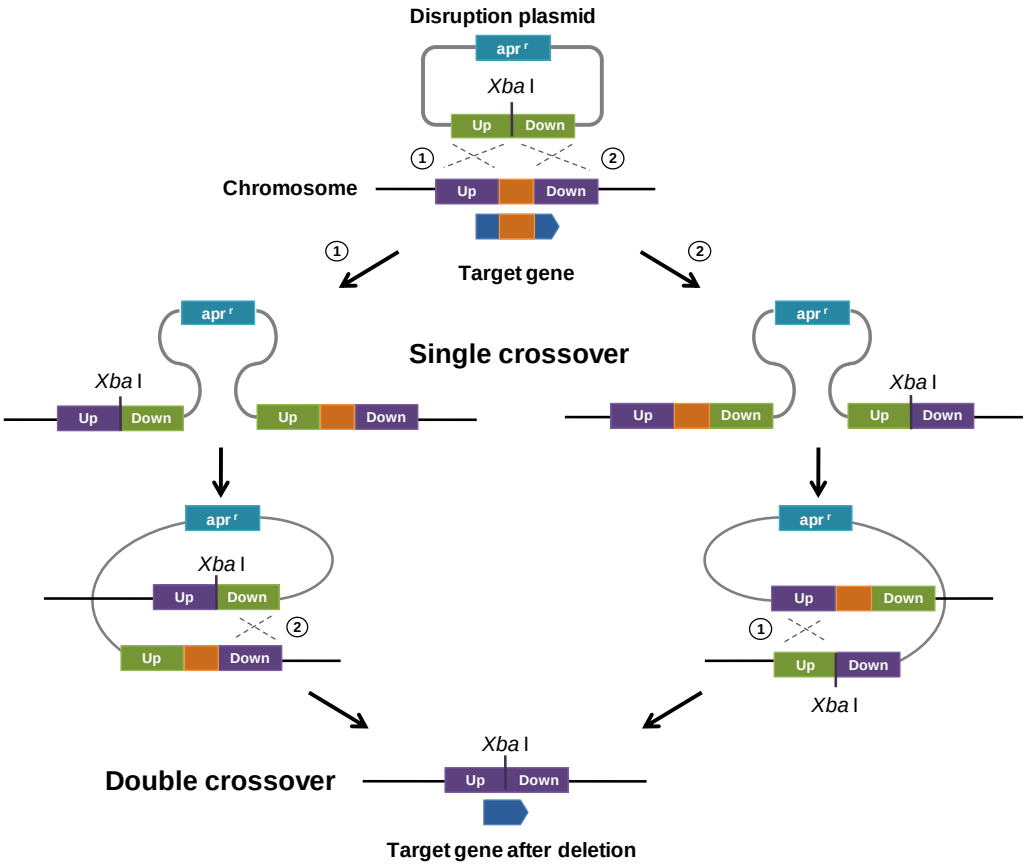


Figure 3-4. Gene inactivation using an in-frame deletion.

Inactivation of *orf5* in TW-R50-13 by single-crossover disruption

Deletion of the relatively large *orf5* gene was performed via single crossover (Figure 3-5). The disruption plasmid of *orf5* (11,442 bp) was constructed by amplifying a 2,010-bp DNA fragment near the C-terminal of the gene. The amplification of the gene was performed using the primer set listed in Table 3-2. The PCR product was then cloned into pT7blue vector and outsourced for DNA sequencing (Greiner Bio-One, Tokyo). After sequence confirmation, the plasmid obtained was digested with *Hind* III, and the DNA fragment was cloned into pUC118apr to give pUC118aprΔ*orf5*. For disruption, TW-R50-13 was transformed with pUC118aprΔ*orf5*, and the disruptants (TW-R50-13Δ*orf5*) were obtained as colonies that grew on R2YE plates containing 25 μg mL^{-1} apramycin.

The deletion mutants were confirmed by PCR amplification (Go Taq® Green Master Mix, Promega) using 2 sets of primers: P1, Del5-check-Fw (5'-GGATGCCTTGCGCGTCCGCATCGCCC-3') and Del5-check-Rv (5'-CCGCGAGGGTCTCCAGCGTAGTGCTCAG-3'); and P2, Apr-Fw (5'-CCATTTGCCTTTGCGGCAGCGGGGC-3') and Del5-check-Rv. To assess the production of compounds **3** and **3a**, TW-R50-13Δ*orf5* deletion mutants were grown in the A-94964 producing medium (supplemented with 25 μg mL^{-1} apramycin) at 27°C for 4 days before being extracted and analyzed by HRESI-MS as described in Unit 1.3 and Table 2-1, respectively.

Table 3-2. Primer sets used to construct the *orf5* disruption plasmid.

Primer	Oligonucleotide sequence (5' to 3')*	Restriction enzyme
Del5-Fw	GGGAAGCTTCGGAACTCGCTGTGCTGCAGG	*Hin*d III
Del5-Rv	GGGAAGCTTTGTCGATCATCGCGACGCCCG	*Hin*d III

*Restriction enzyme site underlined.

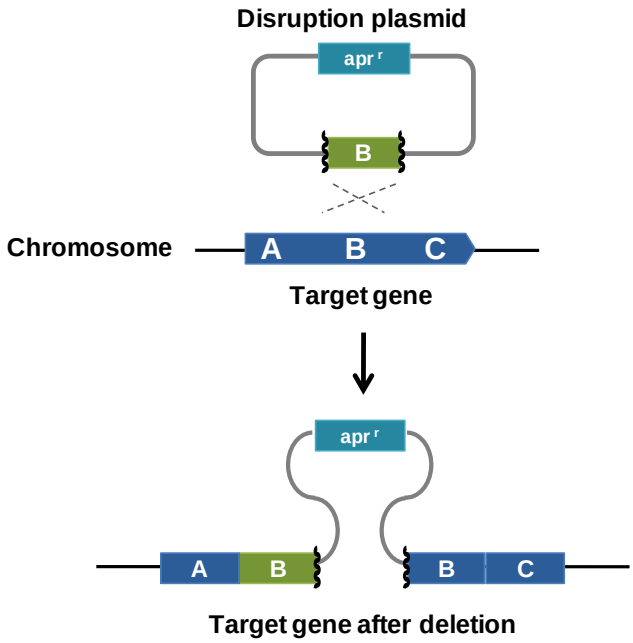


Figure 3-5. Gene inactivation via a single-crossover.

Fermentation and isolation of annimycin.

The *rif* mutant TW-R50-13 was routinely grown at 30°C on a TSB agar plate containing 3% TSB and 2% agar. For fermentation, the *rif* mutant was pre-cultured using 10 mL of a medium containing 3% TSB supplemented with 0.2 μg mL^{-1} rifampicin in 50-mL test tubes on a reciprocal shaker at 30°C for 2 days. The seed medium (2 mL) was then inoculated into a 500-mL baffled flask containing 100 mL of the A-94964 producing medium. The TW-R50-13 was cultivated in 4 L (100 mL × 40) of the producing medium at 27°C for 4 days on a rotary shaker (180 rpm). After fermentation, the culture was separated into broth and mycelial cake by centrifugation at 5,000 rpm for 10 min. The broth was passed through a DIAION HP-20 flash chromatography column (Mitsubishi Chemical, Tokyo, Japan). The column was eluted with 100% water followed by 50%, 70%, 90%, and 100% methanol in water and finally with 100% acetone. The 100% methanol fraction (60 mg) containing the target compounds was further purified using preparative HPLC with a Senshu Pak PEGASIL ODS column (20 × 250 mm, Senshu Scientific, Tokyo, Japan) and an isocratic elution of 50% acetonitrile containing 0.1% formic acid at a flow rate of 8 mL min^{-1} with monitoring at 303 nm to yield 4-*Z*-annimycin and 4-*E*-annimycin.

Results and discussion

First, to identify the biosynthetic gene cluster for both **3** and **3a**, BLAST searches on the draft sequence of SANK 60404 were performed using ALA synthase (Protein ID: AAX98209) as a query. A biosynthetic gene cluster was found containing 5 open reading frames (orfs) that show close homology to type I polyketide synthases (PKS), amide synthetase, 5-aminolevulinate synthase, and acyl-CoA ligase (Figure 3-6A and Table 3-3). Although this gene cluster seems reasonable for the production of compounds **3** and **3a**, the PKS genes consisted of only one loading module and five extension modules. These 6 modules of PKS genes may be sufficient to produce compound **3a** but not compound **3**. Additionally, the formation of the methylbenzene moiety could not be deduced by the substrates predicted on the basis of the co-linearity rule for each domain in the modules (Figure 3-6B)[69]; assuming the possibility of having an aromatic starter unit, such as *p*-aminobenzoate (PABA) or phenylalanine,[70] the number of PKS modules in this gene cluster would not be suitable for the production of **3** and **3a**.

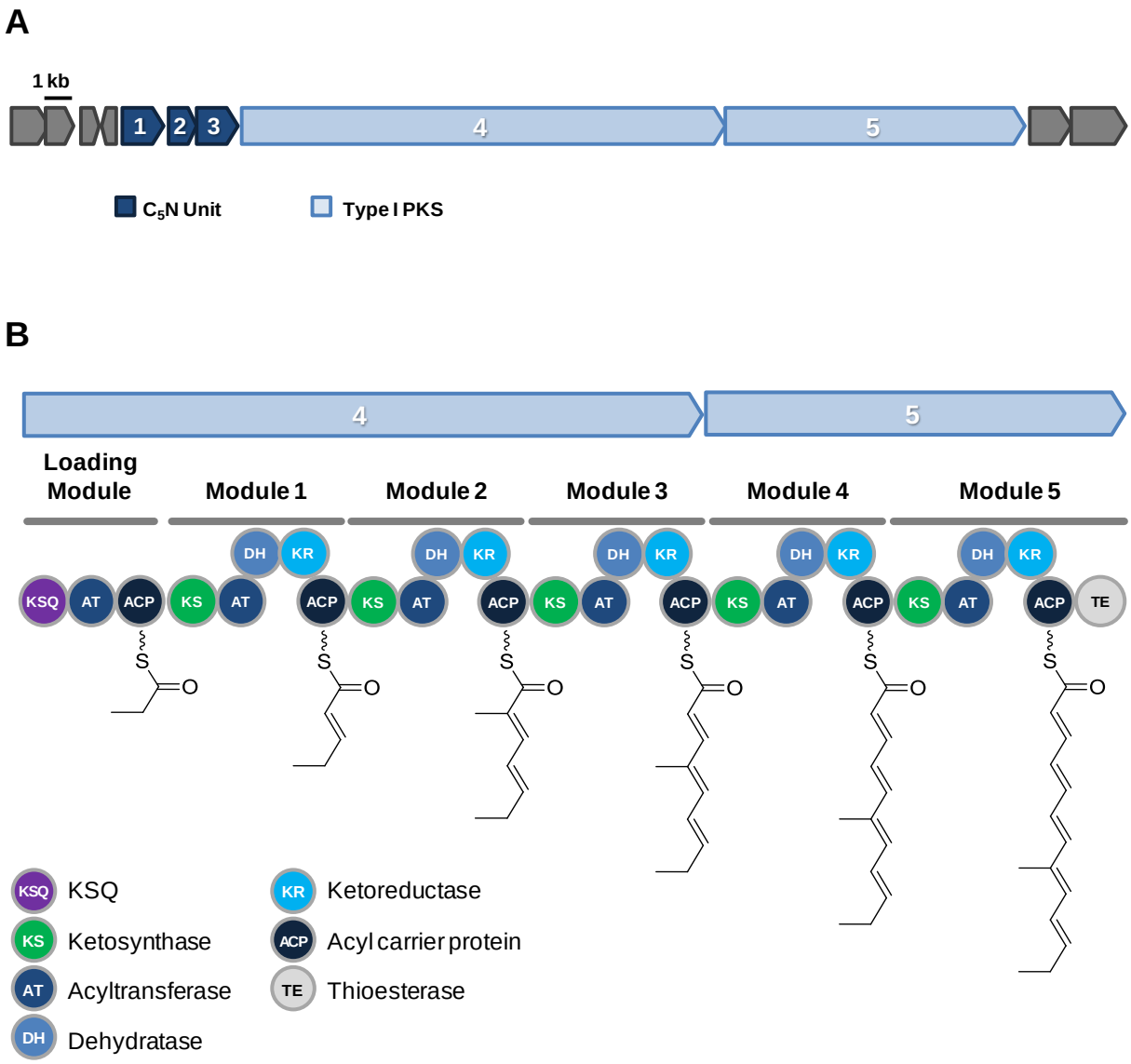


Figure 3-6. (A) Organization of the gene cluster containing the C_5N moiety biosynthetic genes. **(B)** Domain organization of the type I PKS (*orf4* and *orf5*) and possible products for each domain predicted according to the sequence alignment of the AT domains.

Table 3-3. Deduced functions of Orfs in the gene cluster containing the C_5N moiety biosynthetic genes.

Orf*	Amino acids (aa)	Proposed function	Blast hit protein [Origin]	Identity/ Similarity (%)	Protein ID
−4	440	Transporter	MFS transporter [*Streptomyces filamentosus*]	75/83	WP_010071949
−3	334	Hypothetical protein	Hypothetical protein [*Streptomyces roseochromogenus*]	80/85	WP_031227258
−2	260	Thioesterase	Type II thioesterase [*Streptomyces antibioticus*]	55/68	ACN69977
−1	193	Hypothetical protein	Hypothetical protein [*Streptomyces* sp. NRRL S-813]	33/43	WP_051845484
1	527	Amide synthetase	Ann1 [*Streptomyces calvus*]	84/91	AGY30674
2	410	5-Aminolevulinate synthase	Ann2 [*Streptomyces calvus*]	90/93	AGY30678
3	513	Acyl-CoA ligase	Ann3 [*Streptomyces calvus*]	90/95	AGY30675
4	6208	Type I PKS	Ann4 [*Streptomyces calvus*]	92/95	AGY30676
5	3813	Type I PKS	Ann5 [*Streptomyces calvus*]	90/93	AGY30677
+1	536	Transporter	Multidrug MFS transporter [*Streptomyces chattanoogensis*]	86/91	KPC66669
+2	675	DNA helicase	ATP-dependent DNA helicase [*Streptomyces chattanoogensis*]	92/96	KPC67197

*The nucleotide sequence of the gene cluster was deposited into the DDBJ/EMBL/GenBank nucleotide sequence database and assigned the accession number LC090401.

Therefore, to verify that the gene cluster is associated with the biosynthesis of both **3** and **3a**, gene inactivation of amide synthetase (*orf1*) was performed using in-frame deletion (Figure 3-7A). By deleting *orf1*, I anticipated that the synthesis of both **3** and **3a** would be abolished and that any intermediates synthesized by PKS would accumulate. As expected, the HRESI-MS analysis of the *orf1* disruptant revealed the absence of compounds **3** and **3a** (Figure

3-7B). However, no accumulation of the predicted PKS intermediates was detected in the *orf1* disruptant. The lack of the PKS intermediates may be because such intermediates are unstable to the detection conditions, or they remain intact on the acyl carrier protein without being released from the PKS system.

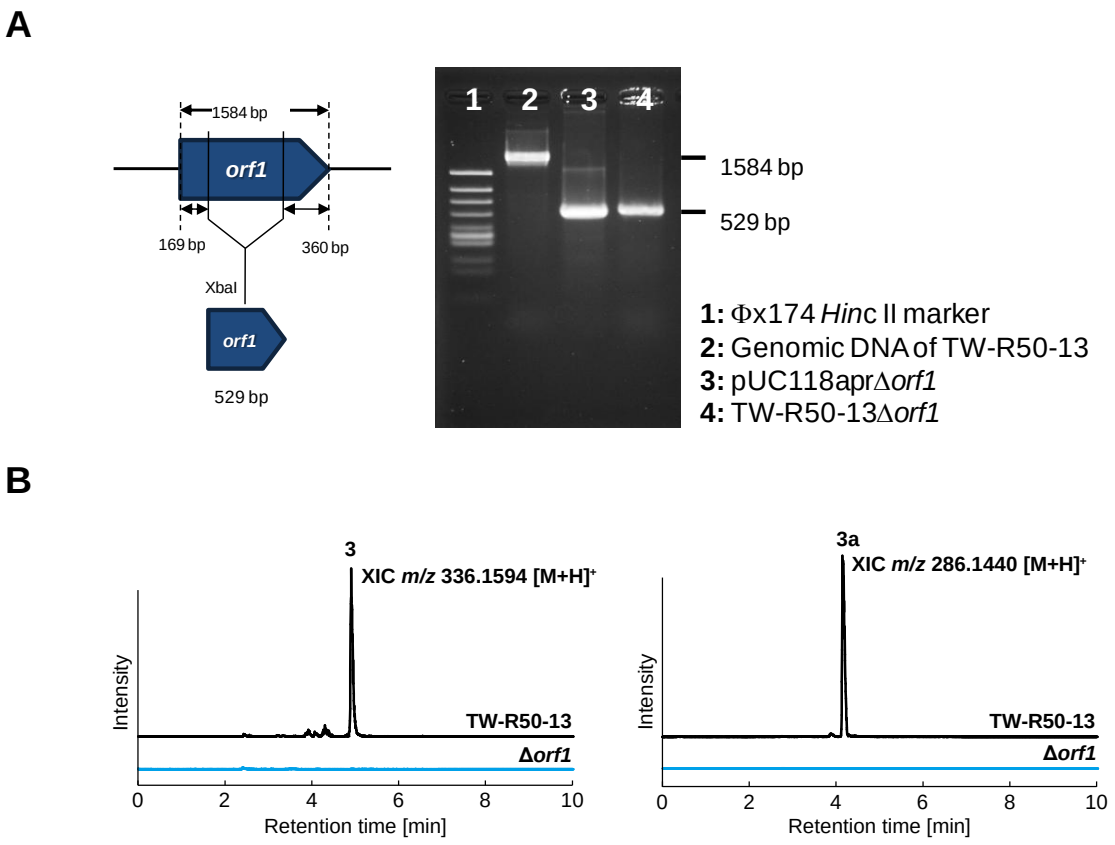


Figure 3-7. Gene inactivation of amide synthetase (*orf1*). **(A)** PCR with genomic DNA from TW-R50-13 produced a 1,584-bp DNA fragment (lane 2), whereas PCR with TW-R50-13Δ*orf1* deletion mutant gave a 529-bp DNA fragment (lane 4). Deletion plasmid pUC118aprΔ*orf1* was used as a positive control (lane 3). **(B)** Comparative HRESI-MS analysis of the culture extracts from TW-R50-13 and the *orf1* disruptant (Δ*orf1*), showing that the production of **3** and **3a** are abolished in Δ*orf1*.

During this biosynthetic study, a gene cluster (accession number: KF683117) with high homology to the one mentioned above was demonstrated to produce annimycin, which is different from **3** and **3a** (Figure 3-8).[71] In addition, annimycin was also isolated from the culture of the TW-R50-13 mutant (Table 3-4), making the involvement of the PKS genes (*orf4* and *orf5*) in the biosynthesis of **3** and **3a** unlikely.

Thus, *orf5* deletion was performed (Figure 3-9); its deletion did not abolish the production of **3** and **3a** (Figure 3-10A), but it did prevent annimycin synthesis (Figure 3-10B). Because there is no other C_5N gene cluster found in the draft genome sequence of SANK 60404, I concluded that *orf1*−*orf3* correspond to the formation of the C_5N moiety for **3**, **3a**, and annimycin but that the PKS genes (*orf4* and *orf5*) are only responsible for the biosynthesis of annimycin. The genes for the methylbenzene moiety of **3** and **3a** are presumably located at a different locus than *orf1*−*3*.

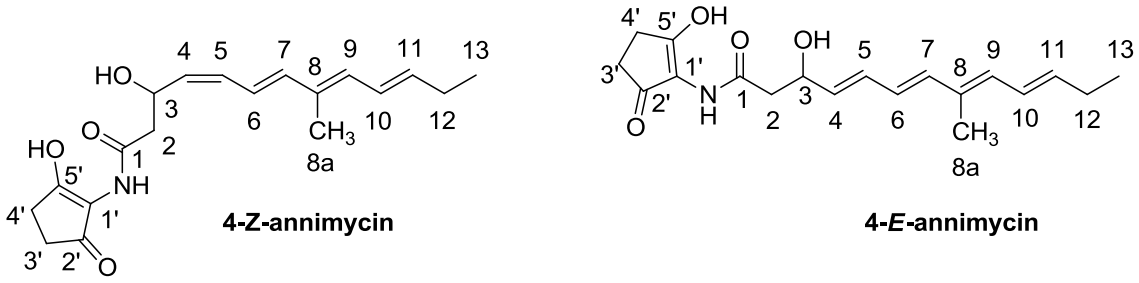


Figure 3-8. Structures of annimycins.

Table 3-4. ^{1}H NMR spectral data for annimycin.

	4-*Z*-annimycin		4-*E*-annimycin	
Position	**Found**	**Literature[71]**	**Found**	**Literature[71]**
	δ_H, mult (*J* in Hz)*	**δ_H, mult (*J* in Hz)****	**δ_H, mult (*J* in Hz)***	**δ_H, mult (*J* in Hz)****
1				
2a	2.65, dd (9, 15.6)	2.66, dd (7.8, 14.3)	2.61, m	N/A
2b	2.57, dd (3, 15.6)	2.52, dd (5.6, 16.5)	2.61, m	2.55, m
3	5.06, dt (3, 9)	5.04, m	4.61, m	4.55, m
4	5.36, t (10.2)	5.37, dd (9.9, 9.9)	5.69, dd (6.6, 15)	5.37, dd (6.4, 15.3)
5	6.13, t (10.8)	6.12, dd (10.9, 10.9)	6.34, dd (10.8, 15)	6.35, dd (10.0, 15.1)
6	6.43, dd (10.8, 15)	6.55, dd (11.2, 15.4)	6.18, dd (10.8, 15)	6.25, m
7	6.30, d (15)	6.27, d (15.1)	6.28, d (15)	N/A
8				
8a	1.86, s	1.86, s	1.84, s	N/A
9	6.09, d (11.4)	6.07, d (10.9)	6.07, d (11.4)	N/A
10	6.36, dd (11.4, 15)	6.40, dd (11.3, 15.0)	6.36, dd (10.8, 15)	N/A
11	5.82, dt (7.2, 15)	5.80, dt (7.2, 14.9)	5.81, dt (7.2, 15)	N/A
12	2.16, quin (7.2)	2.16, m	2.16, quin (7.2)	N/A
13	1.02, t (7.2)	1.02, t (7.4)	1.02, t (7.2)	N/A
NH	8.65, s		8.52, s	
1'				
2'				
3'	2.48, br	2.48, br	2.50, br	N/A
4'	2.55, br	2.48, br	2.57, br	N/A
5'				
5'-OH	13.08, s		13.01, s	

*Measured in $CDCl_3$ (600 MHz)

**Measured in CD_3OD (400 MHz)

N/A, Not available. The ^{1}H NMR spectral data in the literature[13] were collected without isolating 4-*E*-annimycin from the mixture of 4-*E*-annimycin and 4-*Z*-annimycin. In contrast, the individual isomers were isolated here, and the remaining proton signals were obtained as listed in the table.

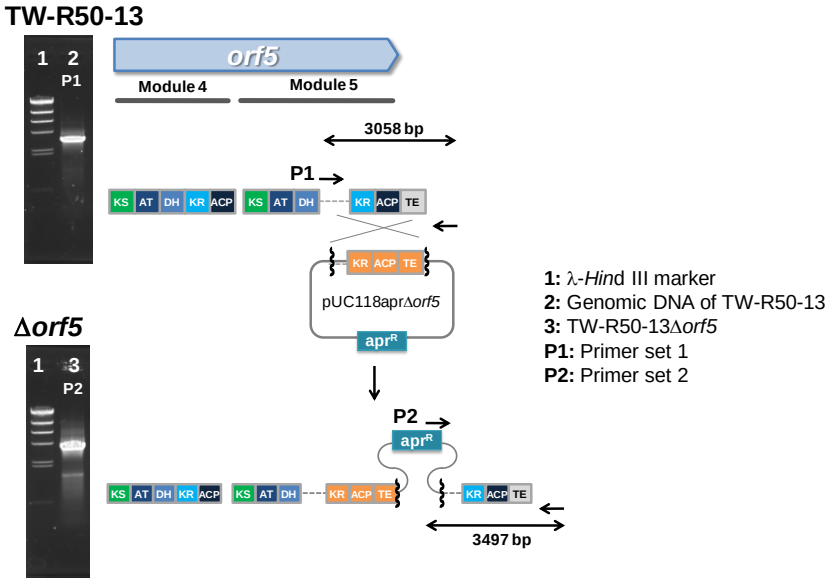


Figure 3-9. Inactivation of *orf5* in TW-R50-13 by single-crossover disruption. PCR with TW-R50-13Δ*orf5* deletion mutant gave a DNA fragment of 3,497 bp, confirming the integration of the deletion plasmid pUC118aprΔ*orf5* into the chromosome of TW-R50-13.

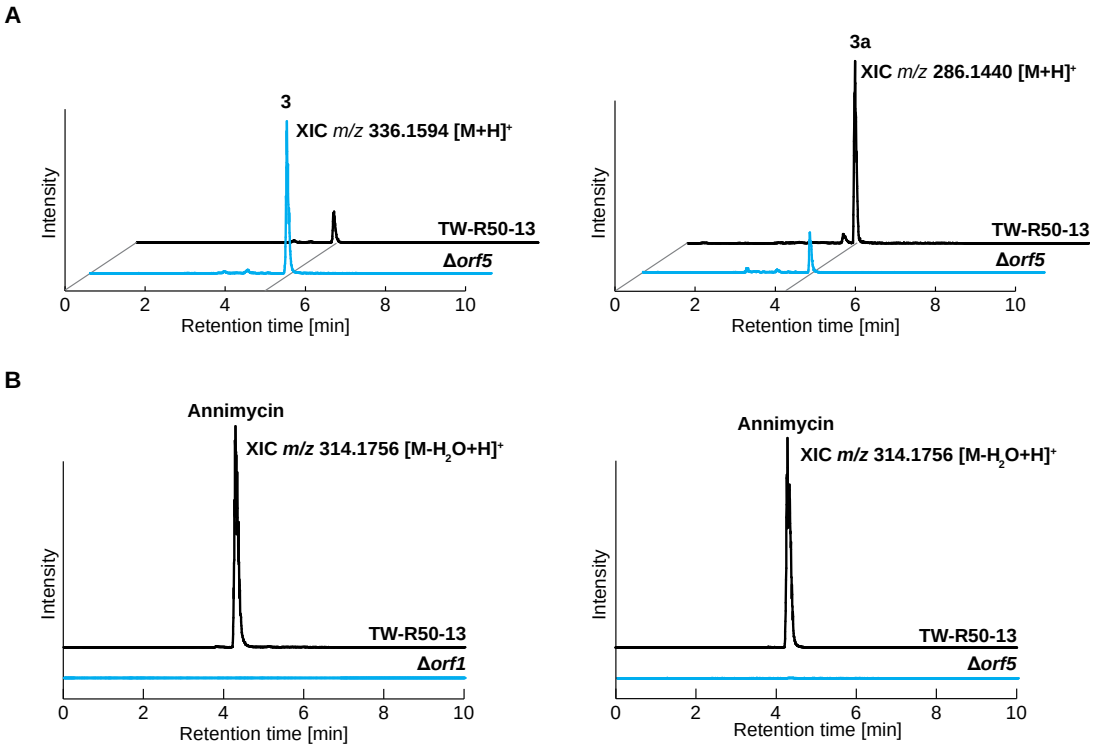


Figure 3-10. Comparative HRESI-MS analysis of the culture extracts from TW-R50-13 and the deletion mutants. **(A)** Deletion of *orf5* (Δ*orf5*) did not abolish the production of compounds **3** and **3a**. **(B)** Deletion of *orf 1* and *orf 5* prevented annimycin synthesis.

3.2 Tracer experiment with ^{13}C-labeled precursors.

Methods

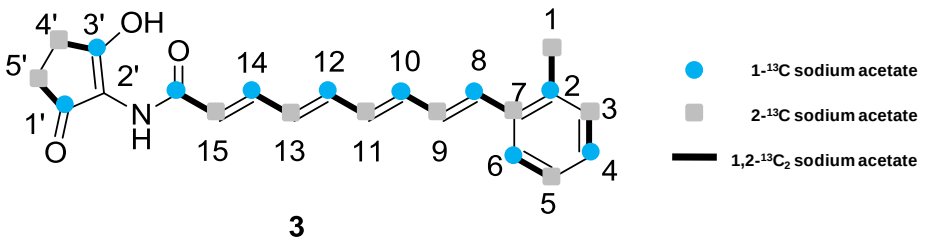
TW-R50-13Δ*orf5* was pre-cultured using 10 mL of medium containing 3% TSB supplemented with 0.2 μg mL^{-1} rifampicin and 25 μg mL^{-1} apramycin in 50-mL test tubes on a reciprocal shaker at 30°C for 2 days. The seed medium (2 mL/100 mL) was then inoculated into 500 mL (100 mL × 5) of 3% TSB medium supplemented with 25 μg mL^{-1} apramycin before being cultivated at 27°C on a rotary shaker (180 rpm). After 12 h of incubation, a ^{13}C-labeled precursor—sodium [1-^{13}C] or [1,2-$^{13}C_2$] or [2-^{13}C]acetate (Cambridge isotope laboratories, Inc., USA)—was added into each flask at 1 mg mL^{-1}, and the cultivation was allowed to continue for an additional 84 h. After fermentation, the culture was extracted as in Unit 2.2, and the labeled compound **3** was isolated directly from the ethyl acetate extract by the same preparative HPLC using an isocratic elution of 70% acetonitrile containing 0.1% formic acid with a flow rate of 8 mL min^{-1}.

Results and discussion

Although the genes involved in the formation of the C_5N moiety were identified, the biosynthetic mechanism for the methylbenzene moiety at the other ends of the compounds remains unknown. Based on the structure of the unsaturated chain attached to the methylbenzene moiety in compound **3**, I speculated that it is a product synthesized by PKS. To test this assumption, a tracer experiment was performed with TW-R50-13Δ*orf5*, which produced a greater amount of **3** than TW-R50-13 (Figure 3-10A). TW-R50-13Δ*orf5* was cultured in the presence of ^{13}C-labeled precursors (sodium [1-^{13}C], [1,2-$^{13}C_2$], and [2-^{13}C] acetate), which could be incorporated into the polyketide structure through acetyl CoA carboxylation. The distribution of the labeled carbons was determined for compound **3** only because of the low production of compound **3a** by the *orf5* deletion mutant.

In the highly conjugated section of compound **3**, feeding with [1-^{13}C]acetate contributed to the enrichment of the carbons with even numbers, whereas the [2-^{13}C]acetate label was found to be distributed among carbons with odd numbers (Figure 3-11). These labeling patterns clearly show the head-to-tail incorporation of the three acetate units into the methylbenzene moiety (C-1 to C-6) and of the five units into the unsaturated chain (C-7 to C-16) of compound **3** (Table 3-5). The distribution of the acetate units in the structure strongly suggested that it is a product of PKS. From the draft genome sequence of SANK 60404, at least 7 gene clusters were retrieved encoding different types of PKS. One of these gene clusters may be responsible for the biosynthesis of the methylbenzene moiety of compounds **3** and **3a**. Therefore, in the next unit, gene inactivation was performed in 5 of these PKS gene clusters to search for the biosynthetic gene cluster for **3** and **3a**. Additionally, it should be noted that the formation of the methylbenzene moiety is unprecedented in the PKS system, suggesting the possibility of an unusual biosynthetic mechanism.

In contrast, the C_5N moiety had the same distribution of the labeled precursors as previously reported.[72-73] [1-^{13}C] and [2-^{13}C]acetate were found on carbon C-1', C-3', C-4', and C-5' in patterns that indicated its origin from the succinate produced via the TCA cycle.[72] Regarding C-2', neither [1-^{13}C] nor [2-^{13}C]acetate was observed. This result supports the previous report indicating that C-2' and the adjacent nitrogen atom are derived from glycine.[73] Further verification of the incorporation of an intact acetate unit was achieved by the administration of [1,2-$^{13}C_2$]acetate (Figure 3-11 and Table 3-5).



3

Figure 3-11. Labeling pattern of compound **3** in tracer experiment with the ^{13}C-labeled sodium acetates.

Table 3-5. Incorporation of labeled precursors into compound **3**.

Position	δ_C	[1-^{13}C]acetate*	[2-^{13}C]acetate* (*J* in Hz)	[1,2-$^{13}C_2$]acetate (*J* in Hz)
1	19.9		19.9	43.8
2	136.1	136.1		43.8
3	130.7		130.7	56.1
4	128.2	128.2		56.1
5	126.3		126.3	56.9
6	125.2	125.2		56.9
7	135.7		135.7	57.5
8	133.2	133.2		57.5
9	129.6		129.5	57.5
10	138.6	138.6		57.5
11	132.0		131.9	57.5
12	142.0	142.1		57.5
13	129.6		129.5	56.7
14	144.5	144.5		56.7
15	119.7		119.6	67.5
16	165.9	165.8		67.5
1'	197.3	197.4	**	39.5
2'	115.2			
3'	174.2	173.9	**	45.3
4'	25.8		25.7 (33)	45.3
5'	32.2		32.2 (33)	39.5

* Attempts to determine the ^{13}C-enrichment were unsuccessful due to the weak signal of the unlabeled carbons.

**Low-intensity peak was observed

3.3 *In silico* genome analysis and gene inactivation: Identification of the methylbenzene gene cluster

Methods

In silico genome analysis

The draft genome sequence of *Streptomyces* sp. SANK 60404 used was mentioned above. *In silico* screening for gene clusters harboring PKS genes in the draft genome sequence was performed by local BLAST using the program Bioedit (http://www.mbio.ncsu.edu/bioedit/bioedit.html). Genome sequences were annotated with protein BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi), FramePlot 4.0 (2007) (http://nocardia.nih.go.jp/fp4/),[67] and 2ndFind (http://biosyn.nih.go.jp/2ndfind/). A graphical representation of the PKS and NRPS domain organization was obtained from amino acid sequences using free databases such as PKS/NRPS-domain finder ver. 1.0 (http://avermitilis.ls.kitasato-u.ac.jp/pfam-pks/).

Inactivation of the PKS gene in Contig 342 *(KS342-orf9*) in TW-R50-13 by single-crossover disruption

Deletion of the *KS342-orf9* gene was performed via single crossover in the same manner as described in Unit 3.1. The disruption plasmid of *KS342-orf9* (5,211 bp) was constructed by amplifying a 2,331-bp DNA fragment near the N-terminal of the gene. The amplification of the gene was performed using the primer set listed in Table 3-6. For disruption, TW-R50-13 was transformed with pUC118apr*Δ**KS342-orf9*, and the disruptants (TW-R50-13*Δ**KS342-orf9*) were confirmed by PCR amplification (Go Taq® Green Master Mix, Promega) using 2 sets of primers: P1, KS342-orf9-check-Fw (5'-CTGTTCGATCTCCACGGGGGAGCACTCGAC-3') and KS342-orf9-check-Rv (5'-CACGCCATCCCGATGACGGCGATCGG-3'); and P2, Apr-Fw (5'-CCATTTGCCTTTGCGGCAGCGGGGC-3') and KS342-orf9-check-Fw. To assess the production of compound **3** and **3a**, TW-R50-13*Δ**KS342-orf9* deletion mutants were grown in the A-94964 producing medium (supplemented with 25 μg mL^{-1} apramycin) at 27°C for 4 days before being extracted and analyzed by HRESI-MS as described in Unit 1.3 and Table 2-1, respectively.

Table 3-6. Primer sets used to construct the *KS342-orf9* disruption plasmid.

Primer	Oligonucleotide sequence (5' to 3')*	Restriction enzyme
KS342-orf9-Fw	GGGAAGCTTTCATCCGCCCTGGTCGCAG	*Hin*d III
KS342-orf9-Rv	GGGAAGCTTGAACCCGAGGTCCTTGAAGGAG	*Hin*d III

*Restriction enzyme site underlined.

Inactivation of the PKS gene in Contig 267 *(KS267-orf6*) in TW-R50-13 by single-crossover disruption

Deletion of the *KS267-orf6* gene was performed via single crossover in the same manner as described in Unit 3.1. The disruption plasmid of *KS267-orf6* (4,620 bp) was constructed by amplifying a 2,007-bp DNA fragment near the N-terminal of the gene. The amplification of the gene was performed using the primer set listed in Table 3-7. For disruption, TW-R50-13 was transformed with pUC118apr*Δ**KS267-orf6*, and the disruptants (TW-R50-13*Δ**KS267-orf6*) were confirmed by PCR amplification (Go Taq® Green Master Mix, Promega) using 2 sets of primers: P1, KS267-orf6-check-Fw (5'-CGCCGGGACCACGGACGACGTACCC-3') and KS267-orf6-check-Rv (5'-GGGGCGGCAGCGCGACGGTACGG-3') and P2, Apr-Fw (5'-CCATTTGCCTTTGCGGCAGCGGGGC-3') and KS267-orf6-check-Fw. To assess the production of compound **3** and **3a**, TW-R50-13*Δ**KS267-orf6* deletion mutants were grown in the A-94964 producing medium (supplemented with 25 μg mL^{-1} apramycin) at 27°C for 4 days before being extracted and analyzed by HRESI-MS as described in Unit 1.3 and Table 2-1, respectively.

Table 3-7. Primer sets used to construct the *KS267-orf6* disruption plasmid.

Primer	Oligonucleotide sequence (5' to 3')*	Restriction enzyme
KS267-orf6-Fw	GGGAAGCTTCCTCGCTGCTGGCGGTCCAC	*Hin*d III
KS267-orf6-Rv	GGGAAGCTTCGACGGTGCAGCCGTCGGC	*Hin*d III

*Restriction enzyme site underlined.

Inactivation of the PKS gene in Contig 827 *(KS827-orf5*) in TW-R50-13 by in-frame deletion

Disruption of *KS827-orf5* in TW-R50-13 was performed using in-frame deletion with slight modification on what described in Unit 3.1. The construction of the *KS827-orf5* (1,101 bp) disruption plasmid was performed by amplifying a 2,010-bp DNA fragment containing the upstream region of *KS827-orf5* and a 2,079-bp DNA fragment containing the downstream region of *KS827-orf5*. The primers used for the amplification are listed in Table 3-8.

Table 3-8. Primer sets used to construct the *KS827-orf5* disruption plasmid.

Primer	Oligonucleotide sequence (5' to 3')*	Restriction enzyme
KS827-orf5-up-Fw	GGGAAGCTTGCGAACCCTGCTGGTCCTGG	*Hin*d III
KS827-orf5-up-Rv	GGGTCTAGACTCCAGCGGCAGCGCCAG	*Xba* I
KS827-orf5-dw-Fw	GGGTCTAGACCCACAGCCGTCCACCCTC	*Xba* I
KS827-orf5-dw-Rv	GGGAAGCTTCTGGTCAGGGCGCCTGCG	*Hin*d III

*Restriction enzyme site underlined.

Each fragment was cloned into pT7blue vector to give KS827-orf5-up-pT7 and KS827-orf5-dw-pT7, and the plasmids were sent for DNA sequencing (Greiner Bio-One, Tokyo). Then, plasmids KS827-orf5-up-pT7 and KS827-orf5-dw-pT7 were digested with *Hin*d III and *Xba* I. Both fragments were simultaneously cloned into pUC118apr to give pUC118aprΔ*KS827-orf5*.

To obtain the *KS827-orf5* disruptant, TW-R-50-13 was transformed with pUC118aprΔ*KS827-orf5*. Single-crossover transformants were obtained as colonies that grew on R2YE plates containing 25 μg mL^{-1} apramycin. The single-crossover transformants obtained were confirmed by PCR, and one was used for the protoplast preparation. The prepared protoplasts were regenerated on R2YE plates. Each regenerated colony was streaked on TSB plates with or without apramycin, and an apramycin-sensitive colony was selected to obtain the *KS827-orf5* knockout mutant TW-R50-13Δ*KS827-orf5*. Successful deletion was confirmed by the PCR amplification (Go Taq® Green Master Mix, Promega) of the *KS827-orf5* gene using KS827-orf5-check-Fw

(5’-GACCGCACCCGACCGCCGACGG-3’) and KS827-orf5-check-Rv (5’-GGCGGTCGTCTACCAGCTGAGGAGGACC -3’). To assess the production of compounds **3** and **3a**, TW-R50-13Δ*KS827-orf5* deletion mutants were grown in the A-94964 producing medium (100 mL) at 27°C for 4 days before being extracted and analyzed by HRESI-MS as described in Unit 1.3 and Table 2-1, respectively.

Inactivation of the PKS gene in Contig 221 *(KS221-orf3*) in TW-R50-13 by single-crossover disruption

Deletion of the *KS221-orf3* gene was performed via single crossover as described in Unit 3.1. The disruption plasmid of *KS221-orf3* (5,892 bp) was constructed by amplifying a 2,022-bp DNA fragment near the N-terminal of the gene. The amplification of the gene was performed using the primer set listed in Table 3-9. For disruption, TW-R50-13 was transformed with pUC118aprΔ*KS221-orf3* and the disruptants (TW-R50-13Δ*KS221-orf3*) were confirmed by PCR amplification (Go Taq® Green Master Mix, Promega) using 2 sets of primers: P1, KS221-orf3-check-Fw (5’-GTGCTGCGCGGGCTGGTCCTGGTG-3’) and KS221-orf3-check-Rv (5’-GCCGGACAGGGTGGTGCCCTGGG-3’) and P2, Apr-Fw (5’-CCATTTGCCTTTGCGGCAGCGGGGC-3’) and KS221-orf3-check-Fw. To assess the production of compound **3** and **3a**, TW-R50-13Δ*KS221-orf3* deletion mutants were grown in the A-94964 producing medium (supplemented with 25 μg mL^{-1} apramycin) at 27°C for 4 days before being extracted and analyzed by HRESI-MS as described in Unit 1.3 and Table 2-1, respectively.

Table 3-9. Primer sets used to construct the *KS221-orf3* disruption plasmid.

Primer	Oligonucleotide sequence (5’ to 3’)*	Restriction enzyme
KS221-orf3-Fw	GGGAAGCTTTCGCTGCTCGCGGTCGTCGAG	*Hin*d III
KS221-orf3-Rv	GGGAAGCTTCGGTTCTTGGGGGCGGGGC	*Hin*d III

*Restriction enzyme site underlined.

Inactivation of the PKS gene of Contig 19 *(KS19-orf21*) in TW-R50-13 by in-frame deletion

Disruption of *KS19-orf21* in TW-R50-13 was performed using in-frame deletion as described in the disruption of *KS827-orf5*. The construction of the *KS19-orf21* (1,185 bp) disruption plasmid was performed by amplifying a 2,001-bp DNA fragment containing the upstream region of *KS19-orf21* and a 2,004-bp DNA fragment containing the downstream region of *KS19-orf21*. The primers used for the amplification were listed in Table 3-10.

Table 3-10. Primer sets used to construct the *KS19-orf21* disruption plasmid.

Primer	Oligonucleotide sequence (5' to 3')*	Restriction enzyme
KS19-orf21-up-Fw	GGGAAGCTTTCCGCGTCGTGCTGTCCAAG	*Hin*d III
KS19-orf21-up-Rv	GGGTCTAGAGTAGAAGGCCGGAATGGGCAC	*Xba* I
KS19-orf21-dw-Fw	GGGTCTAGACTCTTCCTGCACCACGGG	*Xba* I
KS19-orf21-dw-Rv	GGGAAGCTTACGATCGAGCCGACGGTC	*Hin*d III

*Restriction enzyme site underlined.

To obtain the *KS19-orf21* disruptant, TW-R-50-13 was transformed with pUC118aprΔ*KS19-orf21*. Successful deletion double-crossover knockout mutant (TW-R50-13Δ*KS19-orf21*) was confirmed by PCR amplification (Go Taq® Green Master Mix, Promega) of the *KS19-orf21* gene using KS19-orf21-check-Fw (5'-GTGACCGAAGTCGTCATCACCGGCCTG-3') and KS19-orf21-check-Rv (5'-GTCATCGCGGCTCACCTCGCAGCAG-3'). To assess the production of compound **3** and **3a**, TW-R50-13Δ*KS19-orf21* deletion mutants were grown in the A-94964 producing medium (100 mL) at 27°C for 4 days before being extracted and analyzed by HRESI-MS as described in Unit 1.3 and Table 2-1, respectively.

Results and discussion

As mentioned in unit 3.2, the methylbenzene moieties in **3** and **3a** are most likely the products of PKS. To identify the biosynthetic gene clusters for the moiety, local BLAST searches were performed using a ketoacyl synthase (KS) homologue as a query. Ketoacyl synthases are a group of highly conserved genes that are essential for the elongation of polyketide chains by decarboxylative Claisen condensations of malonyl-CoA-derived extender units.[74] The search returned at least 7 gene clusters harboring different types of PKS, such as type I, II, III, iterative type I, and hybrid PKS-non-ribosomal peptide synthase (PKS-NRPS), from the draft genome sequence. Among them, 5 candidate clusters that are most likely to produce **3** and **3a** from 5 different contigs were subjected to gene disruption to determine their relevance to the biosynthesis of compounds **3** and **3a** (Figure 3-12 and Table 3-11).

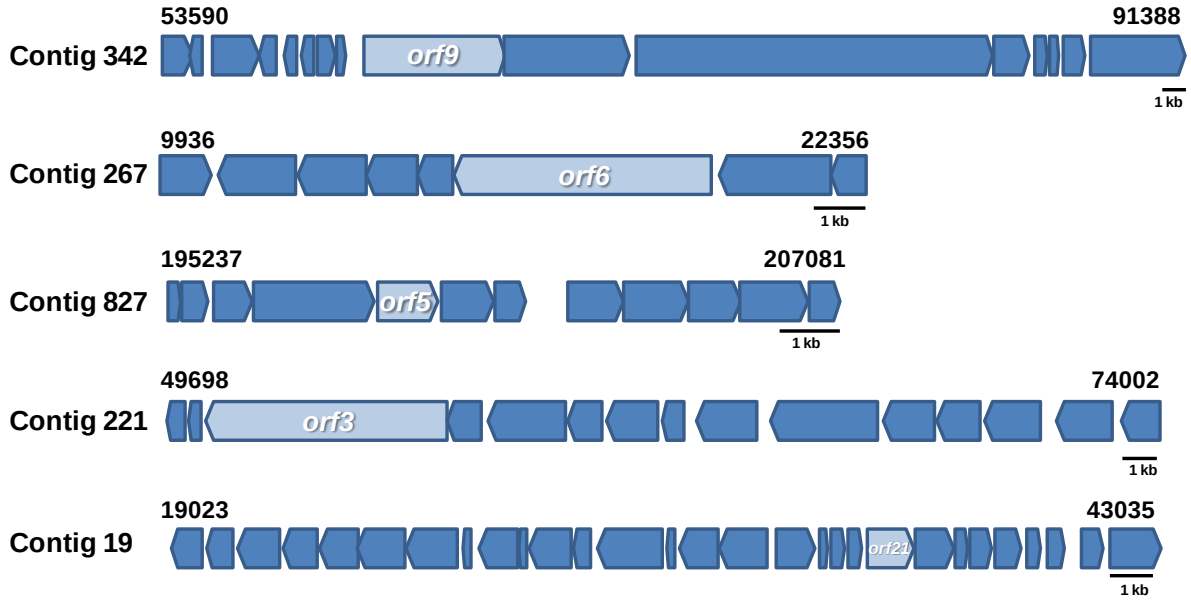


Figure 3-12. Organization of the candidate gene clusters harboring PKS. Gene disruption was performed on KS-containing genes highlighted in light blue. The numbering represents the position of the gene cluster in each contig.

Table 3-11. Annotation for each KS-containing gene selected for deletion.

Contig	Orf	Amino acid	Blast hit protein [Origin]	Identity/ Similarity (%)	Protein ID
342	9	1736	Type I polyketide synthase [*Streptomyces lavenduligriseus*]	877/1459	WP_030790738
267*	6	1539	Type I polyketide synthase [*Streptomyces griseus*]	888/1538	WP_030853235
827	5	352	Stilbene synthase [*Streptomyces aureocirculatus*]	317/364	WP_037685109
221	3	1963	PksE [*Lechevalieria aerocolonigenes*]	534/901	AAO25864
19**	21	394	3-oxoacyl-ACP synthase [*Streptomyces rubellomurinus*]	205/392	WP_045694160

*Contig267 is in the draft genome sequence of SANK 60404 assembled by Edena assembler[75] due to the incomplete assembly in Contig343 by Velvet assembler.
**Contig19 is in the new draft genome sequence of SANK 60404 acquired at Genome Research Center, NODAI Research Institute, Tokyo University of Agriculture, Tokyo, Japan.

Gene disruptions were performed either by single-crossover disruption for large genes (*KS342-orf9*, *KS267-orf6*, and *KS221-orf3*) or in-frame deletion for smaller genes (*KS827-orf5* and *KS19-orf21*) (Figure 3-13). HRESI-MS analysis of the culture extracts from the deletion mutants, unfortunately, did not reveal the abolished production of compounds **3** and **3a** (Figure 3-14). However, with careful observation of the results, I found that deletion *KS19-orf21* decreased the production of compound **3** significantly, unlike the other deletion mutants. An additional *KS19-orf21* deletion experiment on this gene cluster in Contig19 revealed that it is responsible for the biosynthesis of both compounds **3** and **3a**. Further analysis on this gene cluster will be described in the next section.

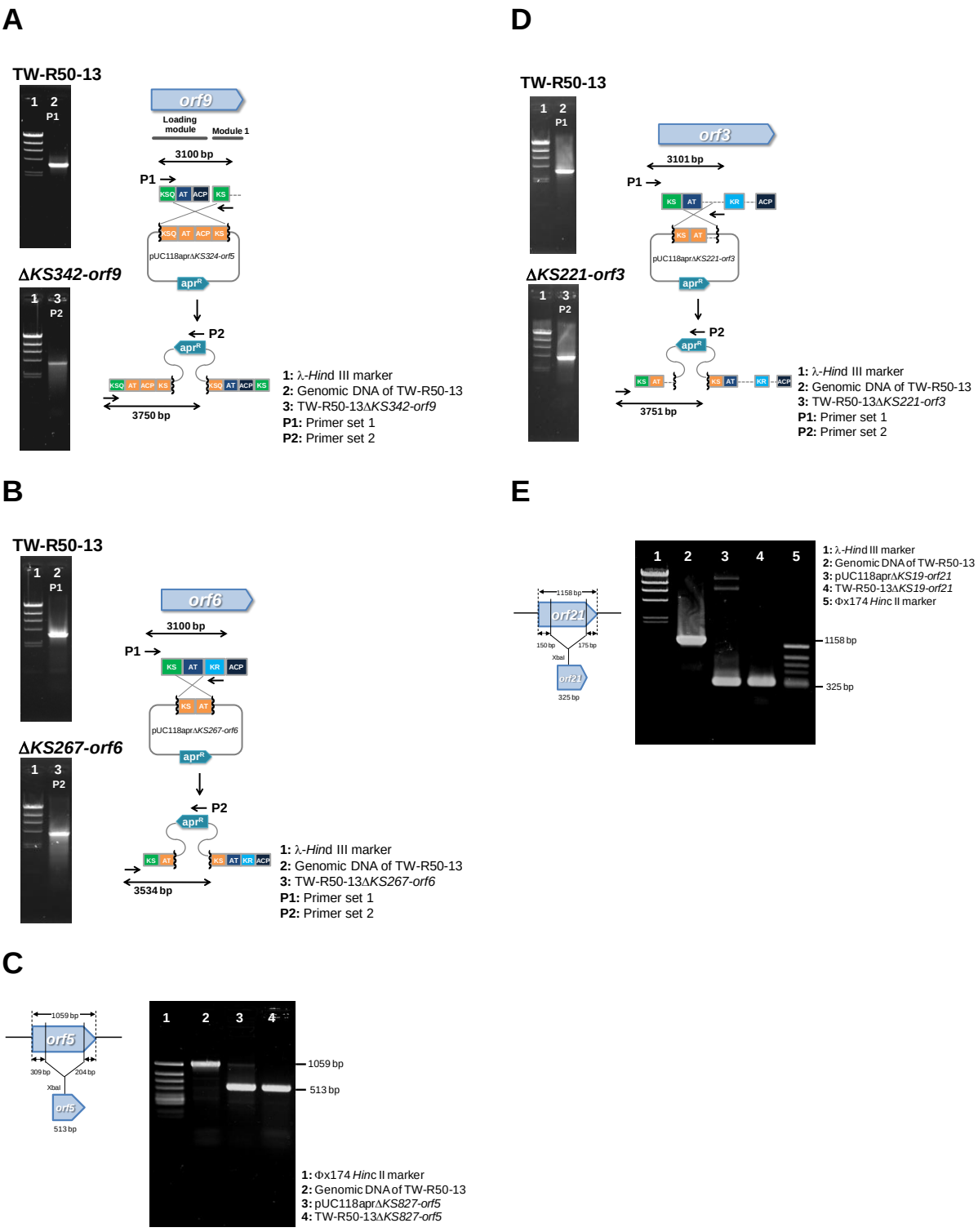


Figure 3-13. Gene inactivation in 5 candidate gene clusters via single-crossover (**A**, **B**, **D**) or in-frame deletion (**C** and **E**).

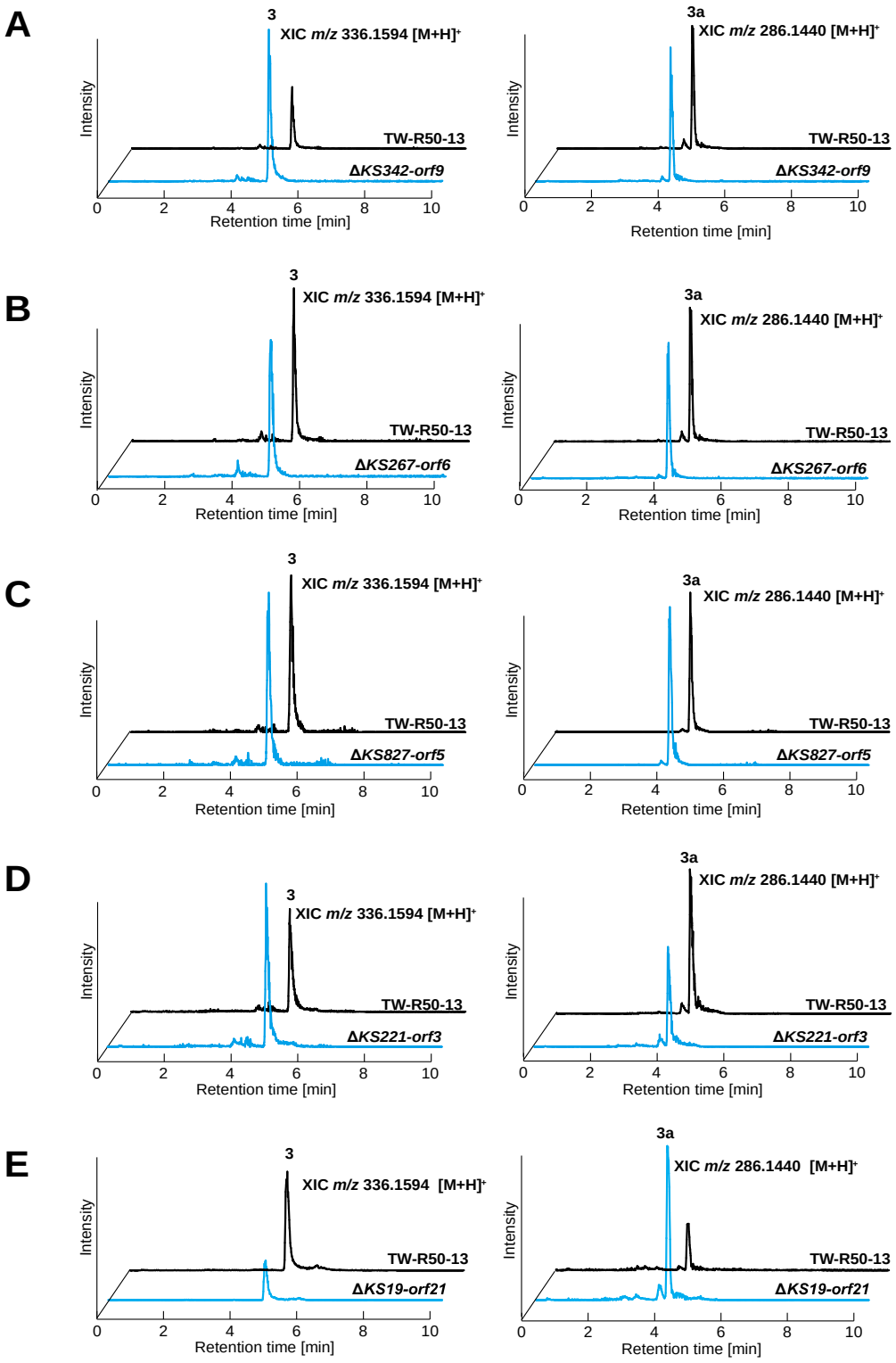


Figure 3-14. Comparative HRESI-MS analysis of the culture extracts from TW-R50-13 and the deletion mutants. **(A)** *KS342-orf9*, **(B)** *KS267-orf6*, **(C)** *KS827-orf5*, **(D)** *KS221-orf3*, **(E)** *KS19-orf21*.

3.4 Biosynthetic studies on the methylbenzene gene cluster

Methods

Inactivation of *MBG2* in Contig19 by in-frame deletion

I renamed *KS19-orf2* in Contig19 as *MBG2* here. Disruption of *MBG2* in Contig19 was performed using in-frame deletion as described in the unit 3.3. The construction of the *MBG2* (678 bp) disruption plasmid was performed by amplifying a 2,037-bp DNA fragment containing the upstream region of *MBG2* and a 2,001-bp DNA fragment containing the downstream region of *MBG2*. The primers used for the amplification were listed in Table 3-12.

Table 3-12. Primer sets used to construct the *MBG2* disruption plasmid.

Primer	Oligonucleotide sequence (5' to 3')*	Restriction enzyme
MBG2-up-Fw	GGGAAGCTTCAACGACACGATGTGGCAGC	*Hin*d III
MBG2-up-Rv	GGGTCTAGAGAACGGCAGCCAGTCGATC	*Xba* I
MBG2-dw-Fw	GGGTCTAGAAAGTTCTGGGGCGTGGACC	*Xba* I
MBG2-dw-Rv	GGGAAGCTTTCGGAAAGATCGCCACCGAG	*Hin*d III

*Restriction enzyme site underlined.

To obtain the *MBG2* disruptant, TW-R-50-13 was transformed with pUC118aprΔ*MBG2*. Successful deletion double-crossover knockout mutant (TW-R50-13Δ*MBG2*) was confirmed by PCR amplification (Go Taq® Green Master Mix, Promega) of the *MBG2* gene using MBG2-check-Fw (5'-CATGGCCCGCCGCAGGCCGC-3') and MBG2-check-Rv (5'-GTCAGCCGCAGCCGCCCGCG-3'). To assess the production of compound **3** and **3a**, TW-R50-13Δ*MBG2* deletion mutants were grown in the A-94964 producing medium (100 mL) at 27°C for 4 days before being extracted and analyzed by HRESI-MS as described in Unit 1.3 and Table 2-1, respectively.

Inactivation of *MBG4* in Contig19 by in-frame deletion

Disruption of *MBG4* (*KS19-orf4*) in Contig19 was performed using in-frame deletion as described in the unit 3.3. The construction of the *MBG4* (870 bp) disruption plasmid was performed by amplifying a 2,058-bp DNA fragment containing the upstream region of *MBG4* and a 2,001-bp DNA fragment containing the downstream region of *MBG4*. The primers used for the amplification were listed in Table 3-13.

Table 3-13. Primer sets used to construct the *MBG4* disruption plasmid.

Primer	Oligonucleotide sequence (5' to 3')*	Restriction enzyme
MBG4-up-Fw	GGGAAGCTTCACGCCTACGAGGTCGACG	*Hin*d III
MBG4-up-Rv	GGGTCTAGACAGCCCGTGCAGGAACAGC	*Xba* I
MBG4-dw-Fw	GGGTCTAGAATCACCGTGCCGTGCCTCG	*Xba* I
MBG4-dw-Rv	GGGAAGCTTCGACGAAGTCCTCAAGGCGG	*Hin*d III

*Restriction enzyme site underlined.

To obtain the *MBG4* disruptant, TW-R-50-13 was transformed with pUC118aprΔ*MBG4*. Successful deletion double-crossover knockout mutant (TW-R50-13Δ*MBG4*) was confirmed by PCR amplification (Go Taq® Green Master Mix, Promega) of the *MBG4* gene using MBG4-check-Fw (5'-GTGCCGAACTGGCACGAGTCCGCC-3') and MBG4-check-Rv (5'-GTCACGTTTCTGTGCTGAGGGACCCTGC-3'). To assess the production of compound **3** and **3a**, TW-R50-13Δ*MBG4* deletion mutants were grown in the A-94964 producing medium (100 mL) at 27°C for 4 days before being extracted and analyzed by HRESI-MS as described in Unit 1.3 and Table 2-1, respectively.

Inactivation of *MBG24* in Contig19 by in-frame deletion

Disruption of *MBG24* (*KS19-orf24*) in Contig19 was performed using in-frame deletion as described in the unit 3.3. The construction of the *MBG24* (516 bp) disruption plasmid was performed by amplifying a 2,007-bp DNA fragment containing the upstream region of *MBG24* and a 2,001-bp DNA fragment containing the downstream region of *MBG24*. The primers used for the amplification were listed in Table 3-14.

Table 3-14. Primer sets used to construct the *MBG24* disruption plasmid.

Primer	Oligonucleotide sequence (5' to 3')*	Restriction enzyme
MBG24-up-Fw	GGGAAGCTTCGGTCGTCCTCGAACGGGC	*Hin*d III
MBG24-up-Rv	GGGTCTAGACGGGTCGATCTCGGTGACCC	*Xba* I
MBG24-dw-Fw	GGGTCTAGAGGCTCGATCGTGATCACCATG	*Xba* I
MBG24-dw-Rv	GGGAAGCTTCTTCGTGTGCGCATGTGCAC	*Hin*d III

*Restriction enzyme site underlined.

To obtain the *MBG24* disruptant, TW-R-50-13 was transformed with pUC118aprΔ*MBG24*. Successful deletion double-crossover knockout mutant (TW-R50-13Δ*MBG24*) was confirmed by PCR amplification (Go Taq® Green Master Mix, Promega) of the *MBG24* gene using MBG24-check-Fw (5'-GTGATCAGCACGACGAGCCTCAAGGAGAC-3') and MBG24-check-Rv (5'-GCTCAGGCATGGGTCAACTCTCCTTCTCC-3'). To assess the production of compound **3** and **3a**, TW-R50-13Δ*MBG24* deletion mutants were grown in the A-94964 producing medium (100 mL) at 27°C for 4 days before being extracted and analyzed by HRESI-MS as described in Unit 1.3 and Table 2-1, respectively.

Inactivation of *MBG25* in Contig19 by in-frame deletion

Disruption of *MBG25* (*KS19-orf25*) in Contig19 was performed using in-frame deletion as described in the unit 3.3. The construction of the *MBG25* (651 bp) disruption plasmid was performed by amplifying a 2,013-bp DNA fragment containing the upstream region of *MBG25* and a 2,001-bp DNA fragment containing the downstream region of *MBG25*. The primers used for the amplification were listed in Table 3-15.

Table 3-15. Primer sets used to construct the *MBG25* disruption plasmid.

Primer	Oligonucleotide sequence (5' to 3')*	Restriction enzyme
MBG25-up-Fw	GGGAAGCTTGGCTACAAGTACCTGCCGAAG	*Hin*d III
MBG25-up-Rv	GGGTCTAGAGAGTTCCTCGCTCACCTGC	*Xba* I
MBG25-dw-Fw	GGGTCTAGAAAGATGGTCGAGCGCGTGCC	*Xba* I
MBG25-dw-Rv	GGGAAGCTTGTCGGCCGCGTCGATGATG	*Hin*d III

*Restriction enzyme site underlined.

To obtain the *MBG25* disruptant, TW-R-50-13 was transformed with pUC118aprΔ*MBG25.* Successful deletion double-crossover knockout mutant (TW-R50-13Δ*MBG25*) was confirmed by PCR amplification (Go Taq® Green Master Mix, Promega) of the *MBG25* gene using MBG25-check-Fw (5'-GTGGCGTTCTGCTACTCGTCGCAGGC-3') and MBG25-check-Rv (5'-CTAGACGACGATGCCACCGTCGATCTGC-3'). To assess the production of compound **3** and **3a**, TW-R50-13Δ*MBG25* deletion mutants were grown in the A-94964 producing medium (100 mL) at 27°C for 4 days before being extracted and analyzed by HRESI-MS as described in Unit 1.3 and Table 2-1, respectively.

Results and discussion

In the previous unit, the deletion of *KS19-orf21* (renamed as *MBG21* here) in Contig19 reduced the production of **3** and **3a**. Intrigued by this result, further analyses on this potential methylbenzene gene (MBG) cluster were performed. The MBG cluster exhibits a unique characteristic in that it consists of 8 stand-alone ketosynthases (*MBG3*, *MBG5*, *MBG6*, *MBG7*, *MBG15*, *MBG16*, *MBG21*, and *MBG22*) (Figure 3-15). These ketosynthases (KSs) show homology to 3-oxoacyl-ACP synthase (45% − 80% sequence identity), which catalyzes the decarboxylative condensation of malonyl-ACP and acetyl-CoA to form acetoacetyl-ACP as observed in the fatty acid biosynthesis.[76] Interestingly, sequence alignments of the KSs in this gene cluster with the reported KSs categorized them into 2 groups; one contains the Cys-His-His motif, a conserved active site for many KSs (KSI), and the other lacks this motif (KSII). *MBG6*, *MBG7*, *MBG16*, and *MBG21* were found to contain the motif (Figure 3-16), whereas the remaining KSs (*MBG3*, *MBG5*, *MBG15*, and *MBG22*) lack of it (Figure 3-17). Additionally, genes such as acyl carrier protein (ACP), ketoreductase (*MBG25*), dehydratase (*MBG23* and *MBG24*), and thioesterase (*MBG17*) are also present in the MBG cluster. All these genes are necessary for typical polyketide biosynthesis, supporting the involvement of the cluster in synthesizing the methylbenzene moiety of compounds **3** and **3a** (Table 3-16).

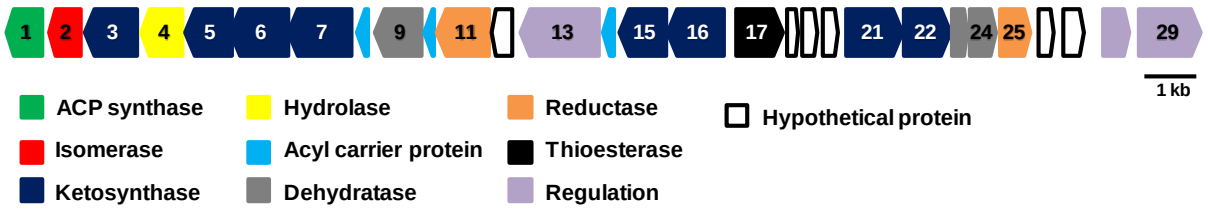


Figure 3-15. Organization of the MBG cluster in Contig19.

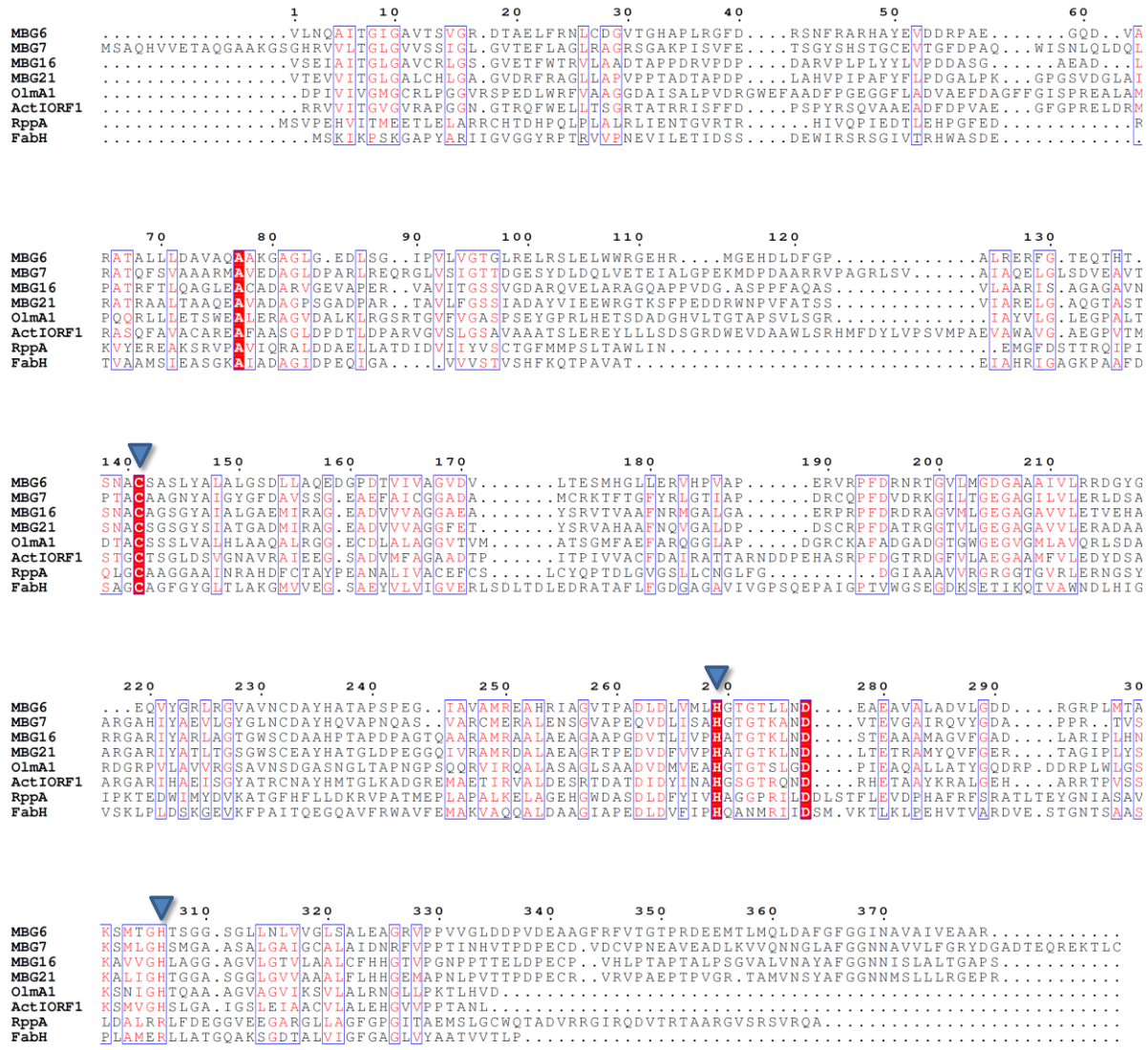


Figure 3-16. Alignment of KSI type KSs in MBG cluster with the existing KSs from type I PKS (*OlmA1*), type II PKS (*ActIOrf1*), type III PKS (*RppA*), and FabH from fatty acid biosynthesis (*FabH*). The arrows indicate the Cys-His-His motif active sites of KS.

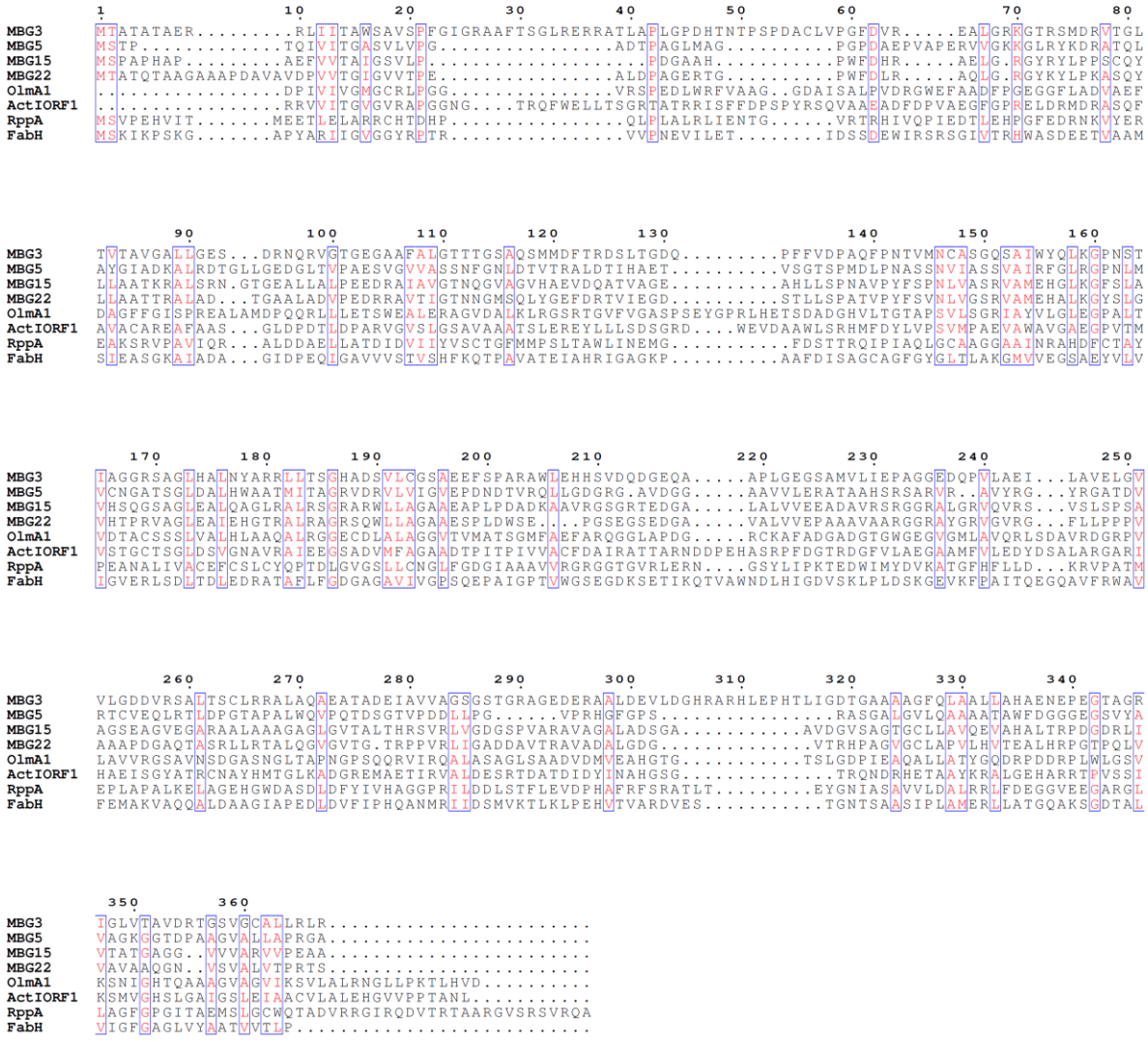


Figure 3-17. Alignment of the KSs lacking the Cys-His-His motif in the MBG cluster with the existing KSs from type I PKS (*OlmA1*), type II PKS (*ActIOrf1*), type III PKS (*RppA*), and FabH from fatty acid biosynthesis (*FabH*).

Table 3-16. Deduced functions of Orfs in the methylbenzene gene cluster.

Orf	Amino acids (aa)	Proposed function	Blast hit protein [Origin]	Identity/ Similarity (%)	Protein ID
1	213	ACP synthase	4-phosphopantetheinyl transferase, partial [*Streptomyces lavendulae*]	98/189	WP_051880150
2	225	Isomerase	Cal27 [*Streptomyces calvus*]	130/239	ALG65309
3	367	Ketosynthase	3-oxoacyl-ACP synthase [*Streptomyces* sp. NRRL WC-3742]	256/369	WP_031076408
4	289	Hydrolase	Alpha/beta hydrolase [*Streptomyces* sp. NRRL WC-3742]	144/233	WP_051838463
5	319	Ketosynthase	3-oxoacyl-ACP synthase [*Streptomyces* sp. NRRL WC-3742]	203/314	WP_031076421
6	379	Ketosynthase	3-oxoacyl-ACP synthase [*Streptomyces* sp. NRRL WC-3742]	275/379	WP_031076422
7	430	Ketosynthase	3-oxoacyl-ACP synthase [*Streptomyces* sp. NRRL WC-3742]	346/431	WP_037973653
8	84	Acyl carrier protein	Polyketide-8 synthase acyl carrier protein [*Streptomyces* sp. NRRL WC-3742]	61/84	WP_031076424
9	310	Dehydratase	Putative 3-hydroxyacyl-ACP dehydratase [*Streptomyces bingchenggensis* BCW-1]	118/288	ADI09877
10	82	Acyl carrier protein	Polyketide-8 synthase ACP [*Streptomyces avermitilis*]	49/79	WP_010985087
11	341	Reductase	Ketoreductase [*Streptomyces* sp. NRRL F-5140]	99/286	WP_030959698
12	149	Hypothetical protein	Hypothetical protein [*Streptomyces* sp. NRRL S-384]	56/143	WP_030913421
13	540	Transporter	MFS transporter [*Streptacidiphilus albus*]	315/509	WP_042436647
14	82	Acyl carrier protein	Polyketide-8 synthase acyl carrier protein [*Streptomyces europaeiscabiei*]	49/81	KND27223
15	323	Ketosynthase	3-oxoacyl-ACP synthase [*Streptomyces bingchenggensis*]	142/314	WP_014179318
16	388	Ketosynthase	3-oxoacyl-ACP synthase [*Streptomyces rubellomurinus*]	201/393	WP_045694160
17	329	Thioesterase	Putative thioesterase [*Streptomyces* sp. Acta 2897]	148/339	AEA30264
18	108	Hypothetical protein	Hypothetical protein [*Streptomyces iranensis*]	55/108	WP_044574712
19	134	Hypothetical protein	Hypothetical protein [*Streptomyces baarnensis*]	77/134	WP_030082125
20	125	Hypothetical protein	Hypothetical protein [*Streptomyces* sp. NRRL WC-3742]	74/127	WP_051838462
21	394	Ketosynthase	3-oxoacyl-ACP synthase [*Streptomyces rubellomurinus*]	205/392	WP_045694160
22	327	Ketosynthese	3-oxoacyl-ACP synthase, partial [*Streptomyces bingchenggensis*]	164/309	WP_014179318
23	128	Dehydratase	3-hydroxyacyl-ACP dehydratase [*Streptomyces violaceusniger*]	62/106	WP_014061555
24	171	Dehydratase	Beta-hydroxyacyl-ACP dehydratase [*Streptomyces rubellomurinus*]	91/157	WP_045702203
25	216	Ketoreductase	3-oxoacyl-ACP reductase [*Streptomyces globisporus*]	166/245	WP_010058215
26	141	Hypothetical protein	Hypothetical protein [*Streptomyces decoyicus*]	87/141	WP_030075195
27	166	Hypothetical protein	Hypothetical protein [*Streptomyces collinus*]	73/168	WP_020942018
28	182	Regulator	DNA-binding protein [*Streptomyces* sp. NRRL S-118]	177/182	WP_031071961
29	441	Protein kinase	Serine/threonine protein kinase [*Streptomyces* sp. NRRL S-1824]	357/480	WP_030972431

To confirm the relevance of the gene cluster in the biosynthesis of compounds **3** and **3a**, a series of in-frame deletion mutants were generated (Figure 3-18). I first deleted the dehydratase (*MBG24*) and ketoreductase (*MBG25*), which are most likely essential in the unsaturated chain formation of the polyketide structure in **3** and **3a**.

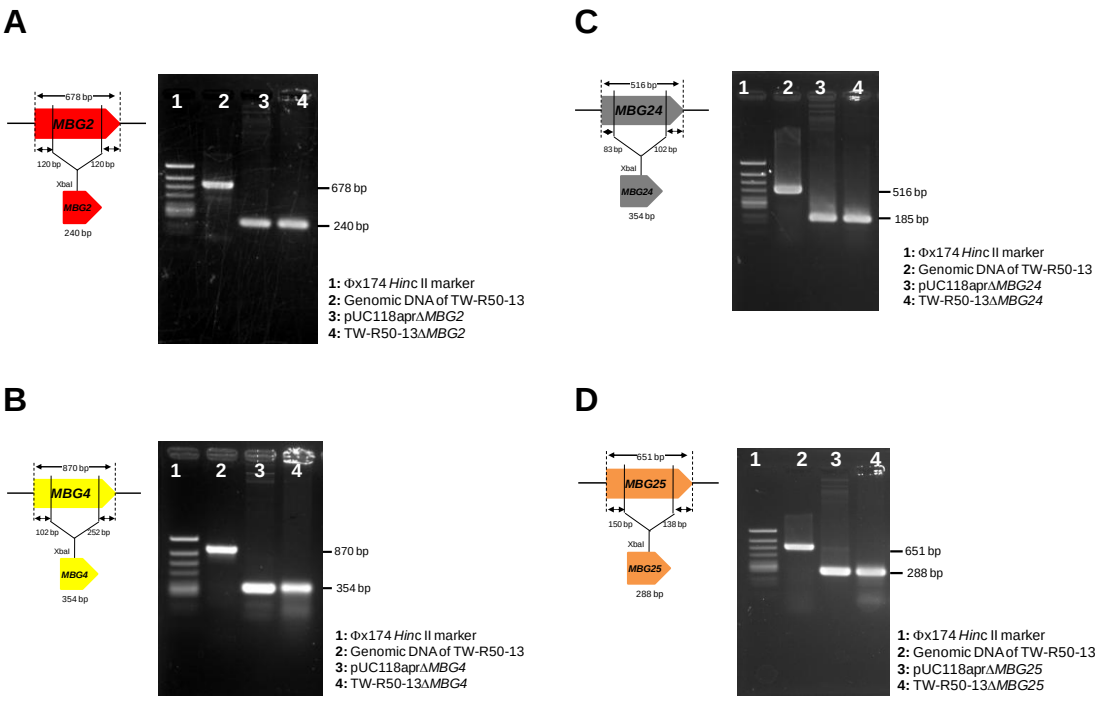


Figure 3-18. Gene inactivation of (**A**) Isomerase (*MBG2*), (**B**) Hydrolase (*MBG4*), (**C**) Dehydratase (*MBG24*), and (**D**) Ketoreductase (*MBG25*) via in-frame deletion.

As expected, the HRESI-MS analysis of the *MBG25* disruptant revealed the abolished production of both compounds **3** and **3a**, verifying that this is the gene cluster for their biosynthesis (Figure 3-19A). In the *MBG24* deletion mutant, however, not only were compounds **3** and **3a** still detected, the production of **3a** was significantly increased compared with TW-R50-13 (Figure 3-19B). The failure to abolish the production of the compounds may be due the complementation by the other dehydratase (*MBG23*) neighboring *MBG24* in the gene cluster. Furthermore, the drastic increase in the production of **3a** observed in the

MBG24 deletion mutant suggests that *MBG23* and *MBG24* may be responsible for the dehydration reaction at different chain lengths; presumably *MBG23* catalyzes the reaction primarily for short chains (< C_{12}) and *MBG24* catalyzes the reaction for longer chains (C_{12}<). In addition, the accumulation of new metabolites was also detected in the culture extract of the *MBG24* deletion mutant (Figure 3-20). The predicted molecular formula of these peaks suggested that they are most likely intermediates or shunt products from the chain elongation pathway. On the other hand, no accumulated product was detected in *MBG25* deletion mutant. This result is not unusual as the intermediate may be unstable for detection or it may remain intact on the acyl carrier protein.

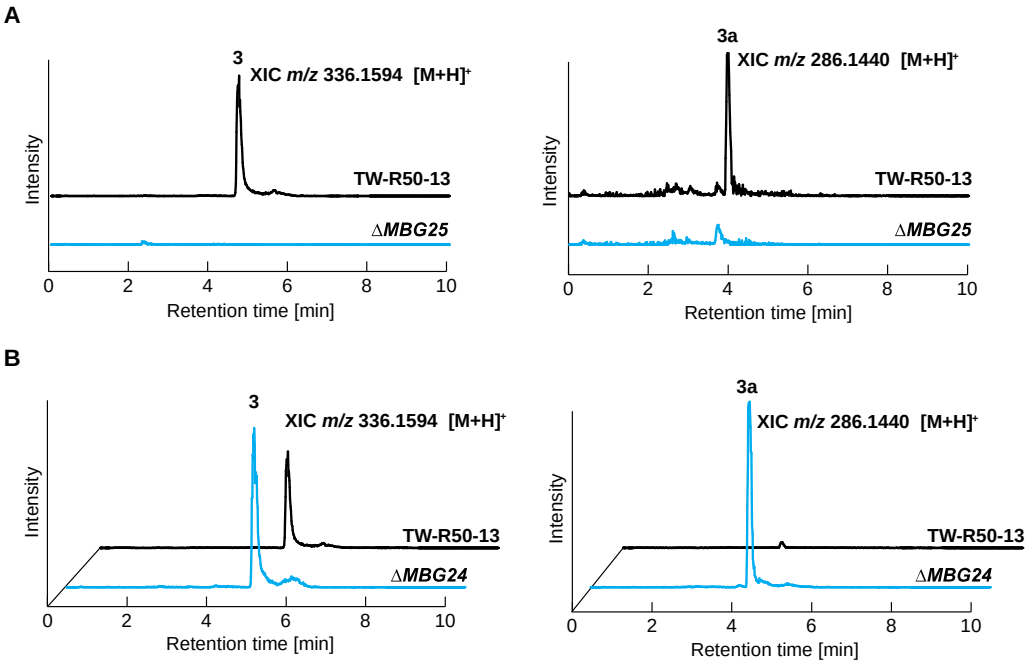


Figure 3-19. HRESI-MS analysis of the culture extracts from TW-R50-13 and the deletion mutants (**A**) *ΔMBG25* and (**B**) *ΔMBG24*.

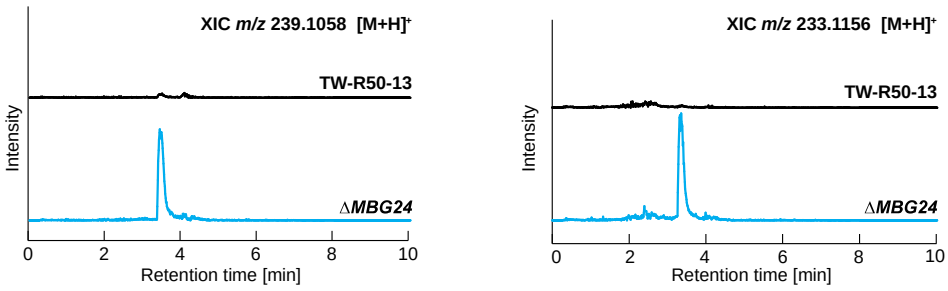


Figure 3-20. HRESI-MS analysis showing the accumulated metabolites in the *MBG24* deletion mutant. The predicted molecular formula for *m/z* 239.1058 $[M+H]^+$ and 233.1156 $[M+H]^+$ peaks are $C_{14}H_{16}O_3$ and $C_{16}H_{14}O_2$, respectively.

Then, to gain a better understanding of the biosynthetic mechanism by this MBG cluster, 2 additional genes that encode the tailoring enzymes isomerase (*MBG2*) and hydrolase (*MBG4*) were also disrupted. The deletion of *MBG2* not only abolished both compounds **3** and **3a** but, most importantly, accumulated an unknown metabolite that has the same exact mass as that of compound **3a**. From the predicted function of *MBG2*, this unknown metabolite is possibly the product prior to the cyclization that forms the methylbenzene moiety in **3a** (Figure 3-21).

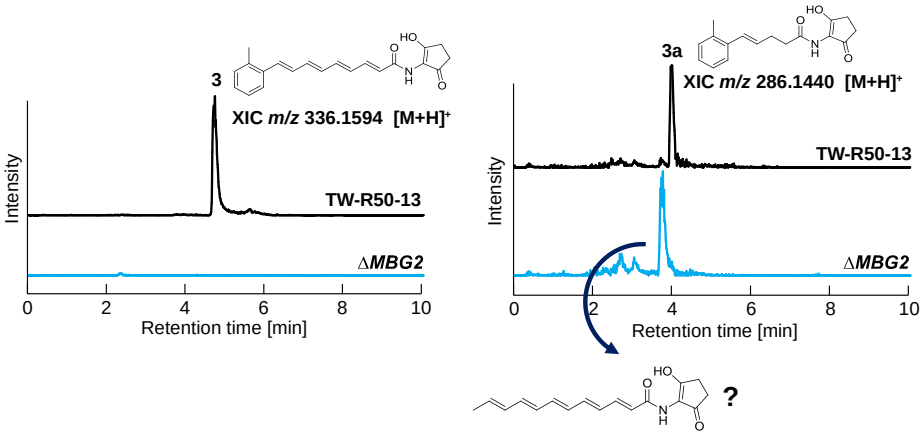


Figure 3-21. HRESI-MS analysis of the culture extracts from TW-R50-13 and *MBG2* deletion mutants showing the absence of compounds **3** and **3a** and the accumulated unknown metabolites.

In contrast, the inactivation of *MBG4* reduced but did not abolish the production of compounds **3** and **3a** (Figure 3-22). Moreover, the production of **3a** was again increased in the deletion mutant. In fact, the same phenomenon was also observed in the *MBG21* deletion described in the previous unit. Perhaps *MBG4* and *MBG21* participate in the later stages of the biosynthesis after the chain elongation reaches C_{12}. Thus, when these genes are inactivated, the chain elongation defect favors the synthesis of products with shorter chains such as compound **3a**.

Considering all the data, a proposed model for the biosynthesis of compounds **3** and **3a** will be discussed in the next section. Also in the next section, the evolutionary relation of the KSs in the MBG cluster and their distribution in other actinomycetes will be described.

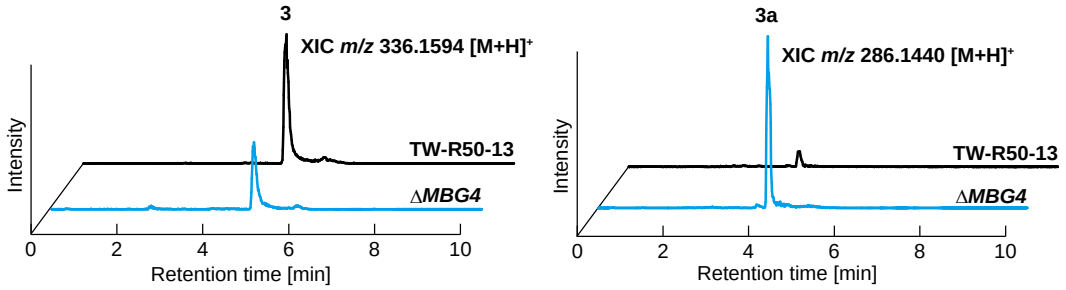


Figure 3-22. HRESI-MS analysis of the culture extracts from TW-R50-13 and *MBG4* deletion mutants, showing a lower amount of compound **3** but significantly higher production of compound **3a** in the deletion mutant.

3.5 Proposed biosynthetic pathway of the methylbenzene gene cluster and its distribution in other actinomycete strains.

Based on the collected data, the biosynthesis of compounds **3** and **3a** is proposed as shown in Figure 3-23. As mentioned previously, there are 8 KSs categorized into 2 groups (KSI and KSII) in terms of their sequence alignments in the MBG cluster. To produce the methylbenzene-containing chain of **3** and **3a**, these KSs may form a KSI/KSII complex to catalyze the chain elongation by the condensation of acetyl-ACP and malonyl-ACP as observed in the type II PKS system.[77] The building block malonyl-ACP could be generated from malonyl-CoA by the ACP-synthase homologue *MCB1* utilizing the ACP homologues *MBG8*, *MBG10*, or *MBG14*. The potential pairing of the KSs to form a complex and the number of extension units per KSI/KSII complex are currently unknown factors. A report by Pohle, S. et al. (2011) proposed that the cinnamoyl moiety in skyllamycin is putatively formed via a C_{12}-polyene produced by a set of 4 KSs (Sky17, Sky18, Sky19, and Sky22) before undergoing an aromatic cyclization reaction.[78] The sequence analyses on *Sky17*, *Sky18*, *Sky19*, and *Sky22* showed that they are homologues to *MBG7*, *MBG6*, *MBG5*, and *MBG3*, respectively, suggesting that a C_{12}-polyene could be formed by two KSI/KSII complexes. The information hints that MBG3, MBG5, MBG6, and MBG7 may be responsible for the synthesis of the first C_{12}-polyene before further extension by the remaining MBG15, MBG16, MBG21, and MBG22. This assumption matches well with the chain length of compounds **3** (C_{16}) and **3a** (C_{12}); a C_{12}-polyene could form first and then be extended twice to form a C_{16}-polyene before cyclization proceeds to form the methylbenzene moiety. Moreover, a careful analysis of *MBG15*, *MBG16*, *MBG21*, and *MBG22* revealed that they have relatively lower similarity (45% to 53% sequence identity) to 3-oxoacyl-ACP synthase in a BLAST search compared with *MBG3*, *MBG5*, *MBG6*, and *MBG7* (65% to 80% sequence identity), suggesting that they may have different chain-extension mechanisms (Table 3-16).

Additionally, the formation of the polyenes most likely involves ketoreductase (MBG25) and dehydratase (MBG23 and MBG24). From the deletion experiments, the *MBG25* deletion mutant abolished compounds **3** and **3a**, indicating that it is responsible for the synthesis of both compounds. However, the high productivity of compound **3a** detected in the *MBG24* deletion mutant suggests that MBG24 may be involved in the dehydration reactions for compound **3**, whereas MBG23 is involved for compound **3a**. After the polyenes (C_{12} and C_{16}) are generated, the chains could be hydrolyzed and released from the ACP by either MBG4 or MBG17. MBG4 resembles a putative hydrolase, whereas MBG17 possesses a hot-dog fold structure that is common to thioesterase. The deletion of *MBG4* did not abolish the production of compounds **3** and **3a**, but it triggered the accumulation of compound **3a** and significantly reduced the production of compound **3**. Again, the result suggests that MBG4 may be responsible for the release of the C_{16}-polyene from its ACP-bound state but that the deletion might be complemented by MBG17.

After the release of the polyenes from ACP, they could undergo isomerization from (2*E*, 4*E*, 6*E*)-configuration to (2*E*, 4*Z*, 6*E*)-configuration catalyzed by MBG2, bringing the carbon atoms within a reasonable distance for possibly 6π-electrocyclization to form the methylbenzene moiety.[79] However, the cyclization mechanism remains elusive because ring closure between carbon atoms 2 and 7 should be followed by dehydrogenation to produce the aromatic system, but there is no gene encoding such an enzyme in the MBG cluster. Additionally, the unknown metabolite detected in the *MBG2* deletion mutant, which is likely a C_{12}-polyene appended with a C_5N moiety, suggests that the cyclization occurs after the polyene is released from ACP. This is because the amide synthetase (*orf1*) in the C_5N unit gene cluster (described in Unit 3.1) is known to have broad substrate specificity, accepting polyketide acids of different chain lengths and condensing them to the amine of C_5N constitutes.[62] The same amide synthetase could also be involved in condensing the

methylbenzene-containing polyketide acids to the amine of the C_5N unit to form compounds **3** and **3a**.

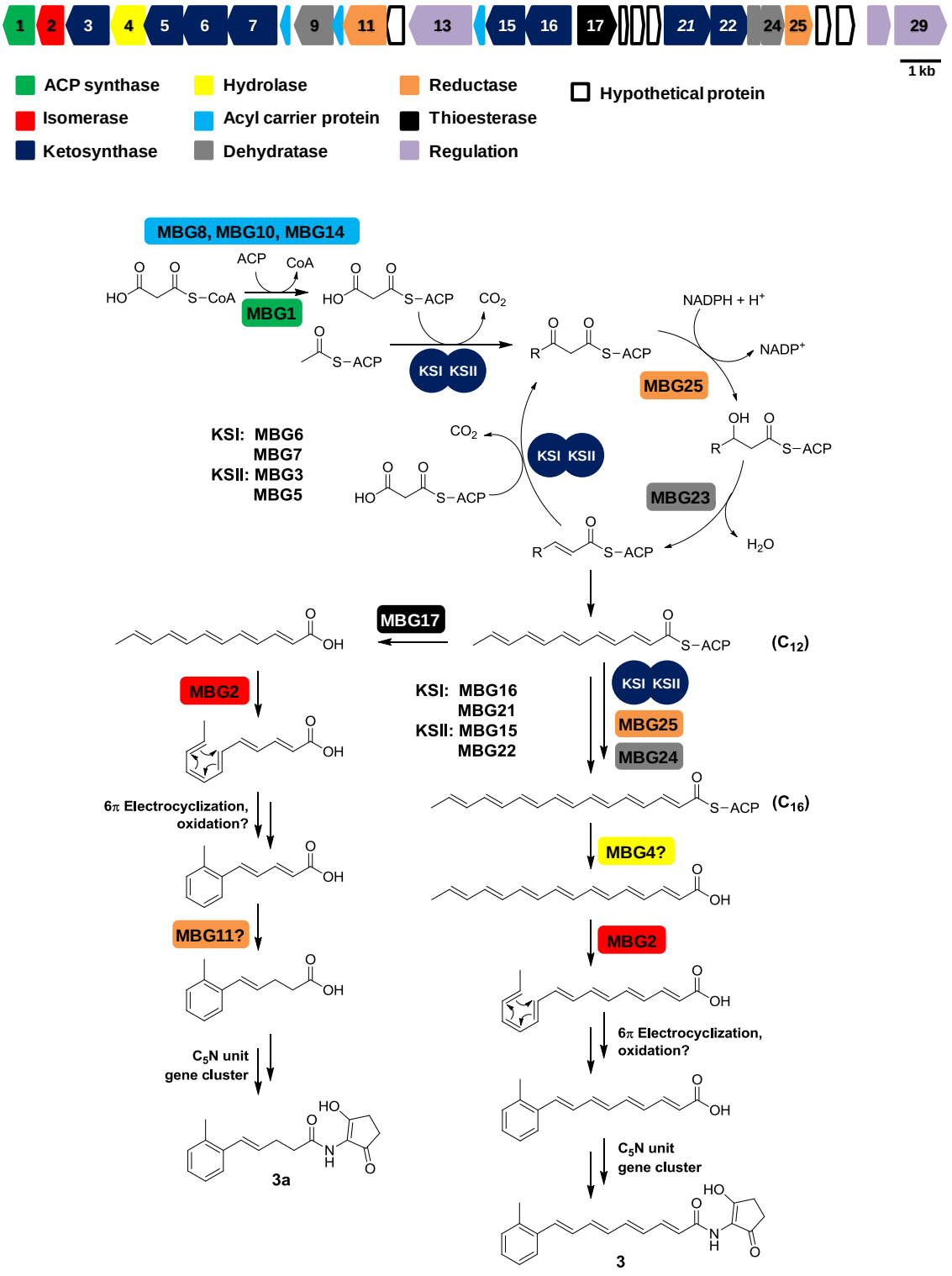


Figure 3-23. Proposed biosynthetic pathway for compounds **3** and **3a**.

Considering the unprecedented biosynthetic mechanism of compounds **3** and **3a**, a phylogenetic tree was constructed using MEGA v6.06[80] to compare KSs found in this study with the various types of existing KSs. As expected, the KSs in the MBG cluster were found in new clades distinct from the other known KSs, suggesting that they are possibly novel PKSs with unknown mechanism (Figure 3-24). Furthermore, they are classified into 2 different clades, the KSI and KSII groups. From the phylogenetic tree, the KSI group is evolutionarily the nearest to enediyne-type KS, whereas KSII has KASII as its nearest neighbor.

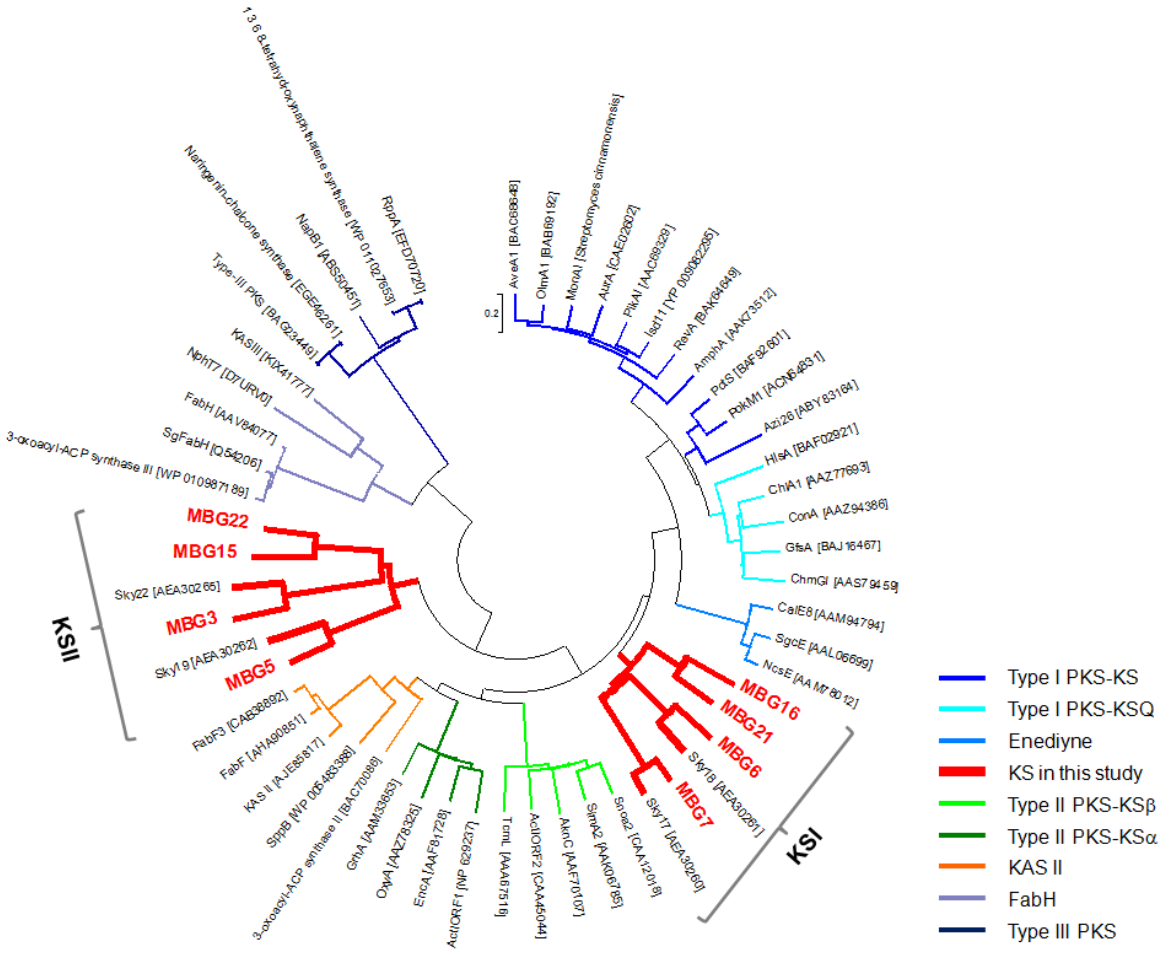


Figure 3-24. Phylogenetic tree comparing the KSs found in this study with other reported KSs.

Although the reaction mechanisms of the MBG cluster are currently unknown, there are a few reports that mention the type of KS found in this study. One of these is on the biosynthesis of skyllamycin (Figure 3-25A). A very recent report on the biosynthesis of simocyclinones (Figure 3-25B) also discussed that the tetraene chain in its structure is synthesized by 2 stand-alone KSs (smcKSI and smcKSII), which most likely belong to the same group as the KSs in the MBG cluster.[81] However, the sequences for these genes are not available yet for further confirmation.

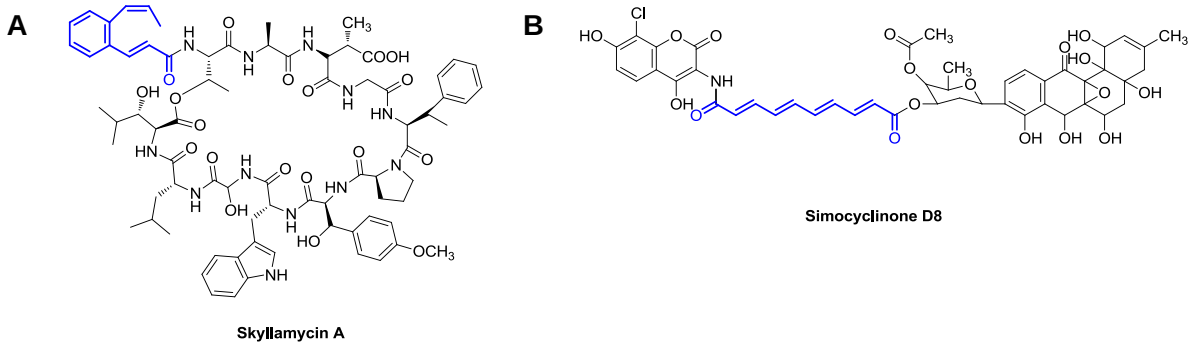


Figure 3-25. Compounds reported to utilize the same type of KSs found in this study. Structures synthesized by these KSs are highlighted in blue.

Given that there are only a number of compounds reported to utilize this type of KS in their biosynthesis, a screening for the MBG cluster in the culture collection from AIST was performed to explore the distribution of this gene cluster. The screening led to the discovery that this type of gene cluster is distributed in many strains (Figure 3-26). Generally, the clusters found in these strains consist of a cassette of gene homologues to *MBG3*, *MBG5*, *MBG6* and *MBG7* (KS), *MBG2* (isomerase), *MBG4* (hydrolase), *MBG23* and *MBG24* (dehydratase), *MBG25* (ketoreductase), and *MBG17* (thioesterase). Similar to those in the MBG cluster, the KSs found in the screening exist in pairs (KSI and KSII) varying from 2 to 6 KSs per cluster. The discovery further supported the assumption that these stand-alone genes are working together to produce polyketide chains as observed in fatty acid

biosynthesis. Furthermore, NRPS is often present in the flanking regions of MBG clusters, indicating the potential of this gene cluster as a probe in genome mining for new natural product discovery.

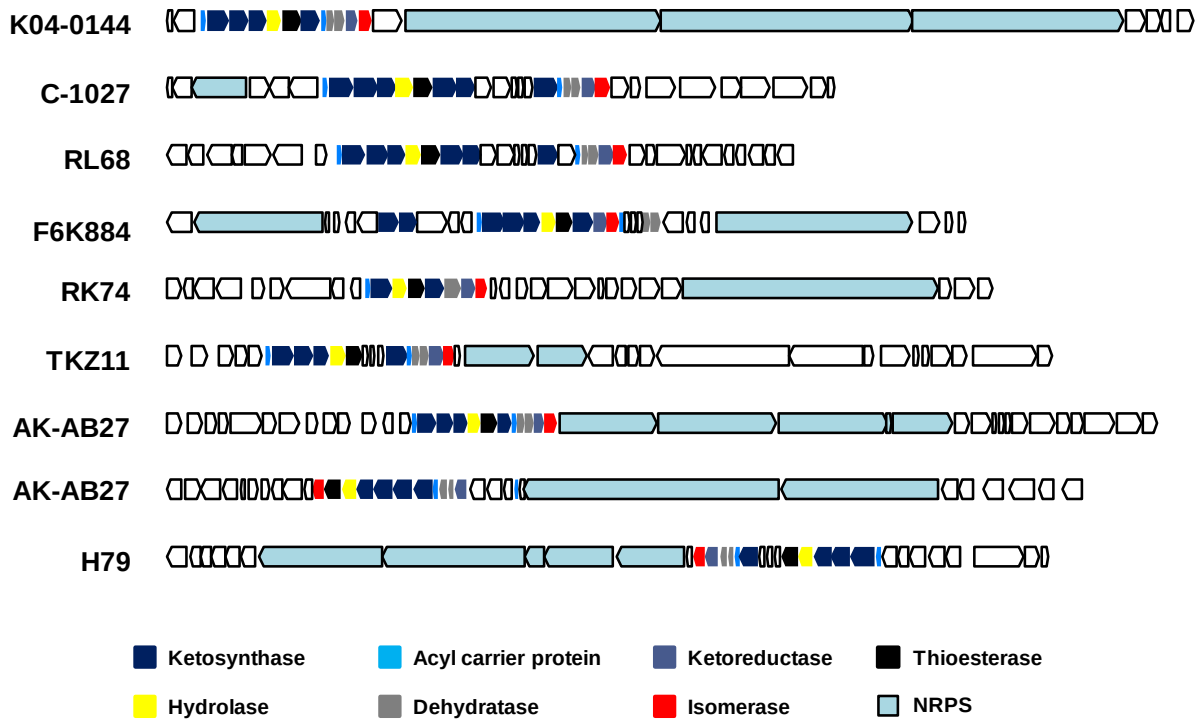


Figure 3-26. MBG clusters found in actinomycete strains in AIST cultures collection.

3.6 Conclusion

Compounds **3** and **3a** are novel compounds featuring a C_5N unit and a methylbenzene moiety. To identify the biosynthetic gene cluster for **3** and **3a**, BLAST search against the draft genome sequence of *Streptomyces* sp. SANK 60404 was performed using a 5-aminolevulinate (ALA) synthase homologue. The search revealed a biosynthetic gene cluster that may be related to the methylbenzene moiety (type I polyketide synthases) and C_5N unit (amide synthetase, ALA synthase, and acyl CoA ligase). Gene deletion experiments indicated that the gene cluster is only responsible for the C_5N moiety and that the genes for the methylbenzene moiety are located at a different locus. Then, to gain insight into the biosynthesis of the polyenoic and methylbenzene moieties of compounds **3** and **3a**, a feeding experiment using ^{13}C-labeled precursors was performed. The ^{13}C-labeling patterns demonstrated that the unsaturated moieties are most likely synthesized by PKS. Gene inactivation in 5 candidate PKS gene clusters retrieved from *in silico* screening revealed that a gene cluster in Contig19 is involved in producing the unsaturated moieties. Unlike any other known PKSs, the gene cluster in Contig19 is made up of 8 discrete KSs. A constructed phylogenetic tree showed that these KSs are categorized into 2 new clades distant from the other existing KSs, suggesting that they potentially are a new type of PKS with unprecedented mechanism. The deletion of 4 genes in this methylbenzene gene (MBG) provided better understanding of the biosynthesis of **3** and **3a**; 2 pathways may be involved for the different chain lengths in the compounds. Nevertheless, the failures to abolish compounds **3** and **3a** in some deletion mutants coupled with the lack of accumulated intermediates raises many questions about the biosynthetic mechanism. Further investigations are in progress focusing on the elucidation of the cyclization mechanism to produce the methylbenzene moiety and the polyketide chain elongation mechanism by the discrete KSs in the MBG cluster.

CHAPTER 4

CHAPTER 4: Biosynthesis of isoindolinomycin isolated from the *rif* mutant S55-50-5

Overview

Isoindolinomycin (IDM) isolated from the *rif* mutant S55-50-5 has a unique chlorinated tetracyclic structure containing an isoindolinone skeleton appended with a sugar moiety (Figure 4-1).

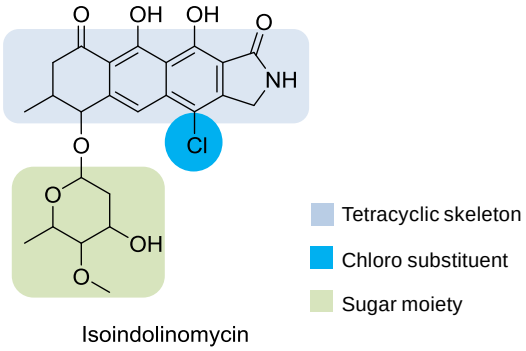


Figure 4-1. Structure of isoindolinomycin.

At a glance, the biosynthesis pathway for IDM most likely involves type II PKS for the construction of the tetracyclic skeleton and a halogenase for the chlorination of the skeleton. Type II PKS minimally consists of a KSα, a KSβ, and an ACP. These components form a complex that catalyzes the decarboxylative condensation of an acyl starter unit with malonyl-CoA extender units in an iterative manner in the chain elongation process (Figure 4-2A).[77] Then, the linear chain backbone will undergo cyclization by aldol condensation to produce a variety of fused-ring structures. In the biosynthesis of the tetracenomycin, the linear polyketide chain is cyclized by 2 different enzymes; the first enzyme catalyzes the formation of the first 3 rings, and then the polyketide is released from the ACP. The second enzyme catalyzes the cyclization of the fourth ring to produce the tetracyclic structure (Figure 4-2B).[82]

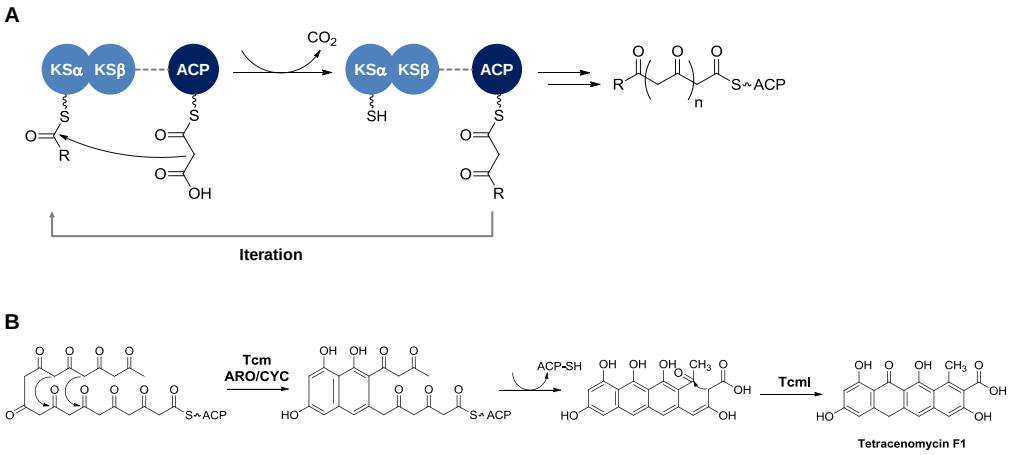


Figure 4-2. Type II PKS mechanism. (**A**) Production of the polyketide chain assembly by the minimal PKS components. (**B**) Proposed cyclization mechanism of the linear chain involving 2 different enzymes in the biosynthesis of tetracenomycin.[82]

Halogenation in natural products is not an unusual event, and in fact, the biological activity of many natural products is influenced by the presence of halogen substituents. Vancomycin exemplifies the bioactivity influenced by a halogen substituent.[83] There are a few types of halogenating enzymes. One is commonly found in bacteria and acts on aromatic moieties in a flavin-dependent manner.[84, 85] The enzymatic mechanism of flavin-dependent halogenases is shown in Figure 4-3. These enzymes use reduced flavin as the cofactor, oxygen as the oxidant, and chloride ion to catalyze halogenation.

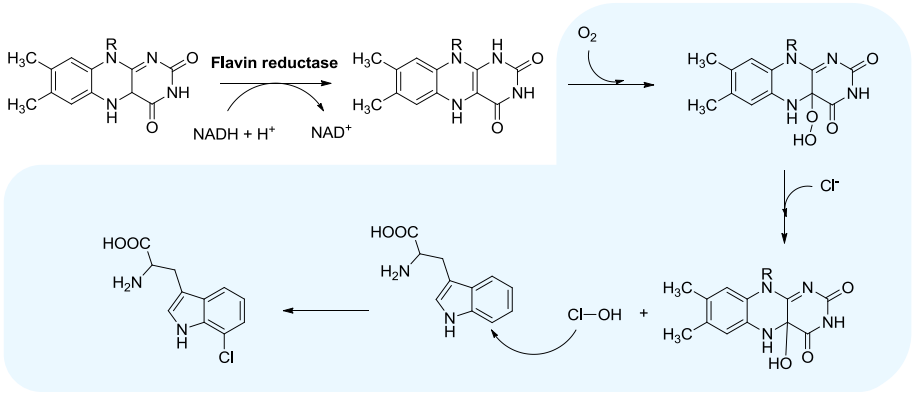


Figure 4-3. Mechanism of flavin-dependent halogenation of tryptophan. The processes occurring inside the halogenase are shaded in light blue.

In this chapter, the biosynthesis of IDM will be described from the *in silico* screening through the proposed synthesis pathway. Here, the biosynthetic studies are mainly focused on 2 aspects: 1) the attachment of the methyl moiety to the core structure and 2) the formation of the 5-membered ring in the tetracyclic structure.

4.1 *In silico* genome analysis and gene inactivation: Identification of the biosynthetic gene cluster for isoindolinomycin (IDM)

Methods

In silico genome analysis

The draft genome sequence of *Streptomyces* sp. SoC090715LN-16 was acquired at Genome Research Center, NODAI Research Institute, Tokyo University of Agriculture, Tokyo, Japan. Genome sequences were annotated with protein BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi), FramePlot 4.0 (2007) (http://nocardia.nih.go.jp/fp4/),[67] and 2ndFind (http://biosyn.nih.go.jp/2ndfind/). The specificity of the stand-alone adenylation domain was predicted using NRPSpredictor2.[86]

Inactivation of *idmB4* in S55-50-5 by in-frame deletion

The disruption of *idmB4* in S55-50-5 was performed using in-frame deletion. The construction of the *idmB4* (1,017 bp) disruption plasmid was performed by amplifying a 2,019-bp DNA fragment containing the upstream region of *idmB4* and a 2,001-bp DNA fragment containing the downstream region of *idmB4*. The primers used for the amplification are listed in Table 4-1.

Table 4-1. Primer sets used to construct the *idmB4* disruption plasmid.

Primer	Oligonucleotide sequence (5' to 3')*	Restriction enzyme
idmB4-up-Fw	GGGAAGCTTCGGATCCGACCGAACCCG	*Hin*d III
idmB4-up-Rv	GGGACTAGTGTGTCCGGTGCGCAGCAG	*Spe* I
idmB4-dw-Fw	GGGACTAGTGAGGCACGGGAGTGGCTC	*Spe* I
idmB4-dw-Rv	GGGAAGCTTTGCTGATGTTCGCGCTGCTGCG	*Hin*d III

*Restriction enzyme site underlined.

Each fragment was cloned into the pT7blue vector to produce idmB4-up-pT7 and idmB4-dw-pT7, and the plasmids were sent for DNA sequencing (Greiner Bio-One, Tokyo). Then, plasmids idmB4-up-pT7 and idmB4-dw-pT7 were digested with *Hin*d III and *Spe* I. Both fragments were simultaneously cloned into pUC118apr[68] to give pUC118aprΔ*idmB4*.

To obtain the *idmB4* disruptant, S55-50-5 was transformed with pUC118aprΔ*idmB4*. Single-crossover transformants were obtained as colonies that grew on R2YE plates containing 25 μg mL^{-1} apramycin. The single-crossover transformants obtained were confirmed by PCR, and one was used for protoplast preparation. The prepared protoplasts were regenerated on R2YE plates. Each regenerated colony was streaked on TSB plates with or without apramycin, and an apramycin-sensitive colony was selected to obtain the *idmB4* knockout mutant S55-50-5Δ*idmB4*. Successful deletion was confirmed by PCR amplification (Go Taq® Green Master Mix, Promega) of the *idmB4* gene using idmB4-check-Fw (5'-CCATGGAGTCAACGGATTCCCCGGAAGTCG-3') and idmB4-check-Rv (5'-GCTCAGTCGGCCTTGCGGGCGAC-3'). To determine the production of IDM, S55-50-5Δ*idmB4* deletion mutants were grown in GYM medium (100 mL) at 27°C for 4 days before being extracted and analyzed by HRESI-MS as described in Unit 1.3 and Table 2-1, respectively.

Inactivation of *idmB24* in S55-50-5 by in-frame deletion

Disruption of *idmB24* in S55-50-5 was performed using in-frame deletion as described above. The construction of the *idmB24* (1,047 bp) disruption plasmid was performed by amplifying a 2,001-bp DNA fragment containing the upstream region of *idmB24* and a 2,001-bp DNA fragment containing the downstream region of *idmB24*. The primers used for the amplification were listed in Table 4-2.

Table 4-2. Primer sets used to construct the *idmB24* disruption plasmid.

Primer	Oligonucleotide sequence (5' to 3')*	Restriction enzyme
idmB24-up-Fw	GGGAAGCTTCGTCGTGCGGATCTACTTCATG	*Hin*d III
idmB24-up-Rv	GGGACTAGTCTCCACCAGGACCTCGAC	*Spe* I
idmB24-dw-Fw	GGGACTAGTGGCAAGTTCCTCGACCTGGAC	*Spe* I
idmB24-dw-Rv	GGGAAGCTTTGTGGGCCGGGTCGAGATAG	*Hin*d III

*Restriction enzyme site underlined.

To obtain the *idmB24* disruptant, S55-50-5 was transformed with pUC118aprΔ*idmB24*. Successful deletion double-crossover knockout mutant (S55-50-5Δ*idmB24*) was confirmed by PCR amplification (Go Taq® Green Master Mix, Promega) of the *idmB24* gene using idmB24-check-Fw (5'-CCATGGAAGCCAACGCCGAGACCGC-3') and idmB24-check-Rv (5'-GTCAGACCGGTCGGCATTCCAGCAGG-3'). To assess the production of IDM, S55-50-5Δ*idmB24* deletion mutants were grown in the GYM medium (100 mL) at 27°C for 4 days before being extracted and analyzed by HRESI-MS as described in Unit 1.3 and Table 2-1, respectively.

Inactivation of *idmB32* in S55-50-5 by in-frame deletion

Disruption of *idmB32* in S55-50-5 was performed using in-frame deletion as described above. The construction of the *idmB32* (1,635 bp) disruption plasmid was performed by amplifying a 2,010-bp DNA fragment containing the upstream region of *idmB32* and a 2,013-bp DNA fragment containing the downstream region of *idmB32*. The primers used for the amplification were listed in Table 4-3.

Table 4-3. Primer sets used to construct the *idmB32* disruption plasmid.

Primer	Oligonucleotide sequence (5' to 3')*	Restriction enzyme
idmB32-up-Fw	GGGAAGCTTCACCCGGACATTGGCCGGG	*Hin*d III
idmB32-up-Rv	GGGTCTAGACGCCACGTCCCACACGGC	*Xba* I
idmB32-dw-Fw	GGGTCTAGACCGGACGGCGATCTCACCG	*Xba* I
idmB32-dw-Rv	GGGAAGCTTGCCCGCATCAGGGCGGTG	*Hin*d III

*Restriction enzyme site underlined.

To obtain the *idmB32* disruptant, S55-50-5 was transformed with pUC118aprΔ*idmB32*. Successful deletion double-crossover knockout mutant (S55-50-5Δ*idmB32*) was confirmed by PCR amplification (Go Taq® Green Master Mix, Promega) of the *idmB32* gene using idmB32-check-Fw (5'-CCATGCCAGATGTCGAGGTTCCCGTACTG-3') and idmB32-check-Rv (5'-CCTCTCAAAGGTGGGACAGCGAGTGCG-3'). To assess the production of IDM, S55-50-5Δ*idmB32* deletion mutants were grown in the GYM medium (100 mL) at 27°C for 4 days before being extracted and analyzed by HRESI-MS as described in Unit 1.3 and Table 2-1, respectively.

Inactivation of *idmB39* in S55-50-5 by in-frame deletion

Disruption of *idmB39* in S55-50-5 was performed using in-frame deletion as described above. The construction of the *idmB39* (1,047 bp) disruption plasmid was performed by amplifying a 2,022-bp DNA fragment containing the upstream region of *idmB39* and a 2,046-bp DNA fragment containing the downstream region of *idmB39*. The primers used for the amplification were listed in Table 4-4.

Table 4-4. Primer sets used to construct the *idmB39* disruption plasmid.

Primer	Oligonucleotide sequence (5' to 3')*	Restriction enzyme
idmB39-up-Fw	GGGAAGCTTCATCCTGGCCTCGGTCCCC	*Hin*d III
idmB39-up-Rv	GGGTCTAGACAGGTCGGCGATGCCCAGC	*Xba* I
idmB39-dw-Fw	GGGTCTAGAATCTGGACCCTCCTGGGCGG	*Xba* I
idmB39-dw-Rv	GGGAAGCTTTGCGACCGGTATGTGCGCCC	*Hin*d III

*Restriction enzyme site underlined.

To obtain the *idmB39* disruptant, S55-50-5 was transformed with pUC118aprΔ*idmB39*. Successful deletion double-crossover knockout mutant (S55-50-5Δ*idmB39*) was confirmed by PCR amplification (Go Taq® Green Master Mix, Promega) of the *idmB39* gene using idmB39-check-Fw (5'-GCATGGCTGACACGAGCACCGCCAC-3') and idmB39-check-Rv (5'-GGTCGCGGCCAGGAAGTTCCAGGAG-3'). To assess the production of IDM, S55-50-5Δ*idmB39* deletion mutants were grown in the GYM medium (100 mL) at 27°C for 4 days before being extracted and analyzed by HRESI-MS as described in Unit 1.3 and Table 2-1, respectively.

Results and discussion

To identify the biosynthetic gene cluster for IDM, a BLAST search on the draft genome sequence of *Streptomyces* sp. SoC090715LN-16 was performed using a flavin-dependent halogenase homolog (Protein ID: AFV71318) as a query. The search found two biosynthetic gene clusters, each bearing a gene with close homology to tryptophan halogenase (Figure 4-4). Analysis of the flanking regions in these gene clusters revealed that the gene cluster in Figure 4-4A contains genes encoding type II PKS, glycosyltransferase, and methyltransferase, which are highly rational for the biosynthesis of IDM.

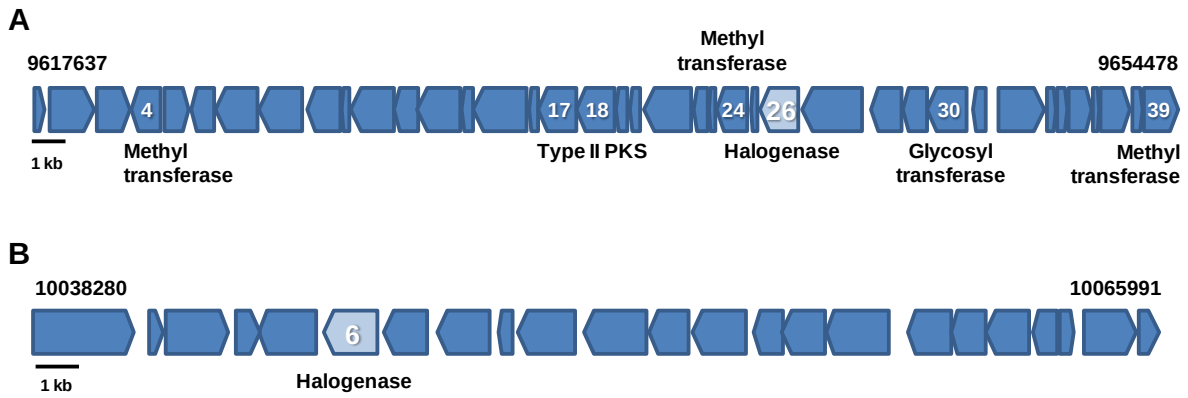


Figure 4-4. Organization of the biosynthetic gene clusters with a halogenase homolog gene. The numbers indicate the position of the gene clusters in the draft genome sequence.

To confirm the assumption, I disrupted *orf32* in the gene cluster in Figure 4-4A. The deletion of *orf32* was performed because IDM may require type II PKS to generate its tetracyclic structure, and *orf32* shows close homology to *tcmG*, a dioxygenase in the biosynthesis of tetracenomycin (Tcm) that catalyzes the triple hydroxylation in the tetracyclic Tcm A2 to form Tcm C.[87] As expected, the deletion abolished the production of IDM, verifying that the gene cluster is responsible for IDM biosynthesis (Figure 4-5).

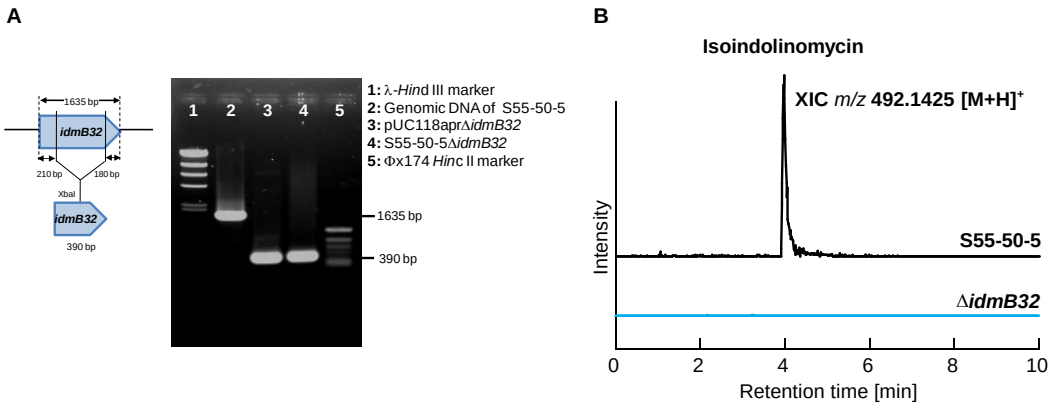


Figure 4-5. Deletion of *orf32* in the gene cluster A. (**A**) Deletion confirmation by PCR amplification. (**B**) Comparative HRESI-MS analysis of the culture extracts from S55-50-5 and the *orf32* (renamed to *idmB32*) deletion mutant showing the loss of IDM.

Because the gene cluster for IDM was identified, I attempted to predict its biosynthesis based on the proposed function (Table 4-5). From the predicted pathway (Figure 4-6), I was interested in the attachment of the methyl moiety to the tetracyclic structure because, to the best of my knowledge, this type of methylation at the skeleton is unprecedented. To investigate this mechanism, the deletion of three methyltransferases (*idmB4*, *idmB24*, and *idmB39*) in the gene cluster were performed (Figure 4-7).

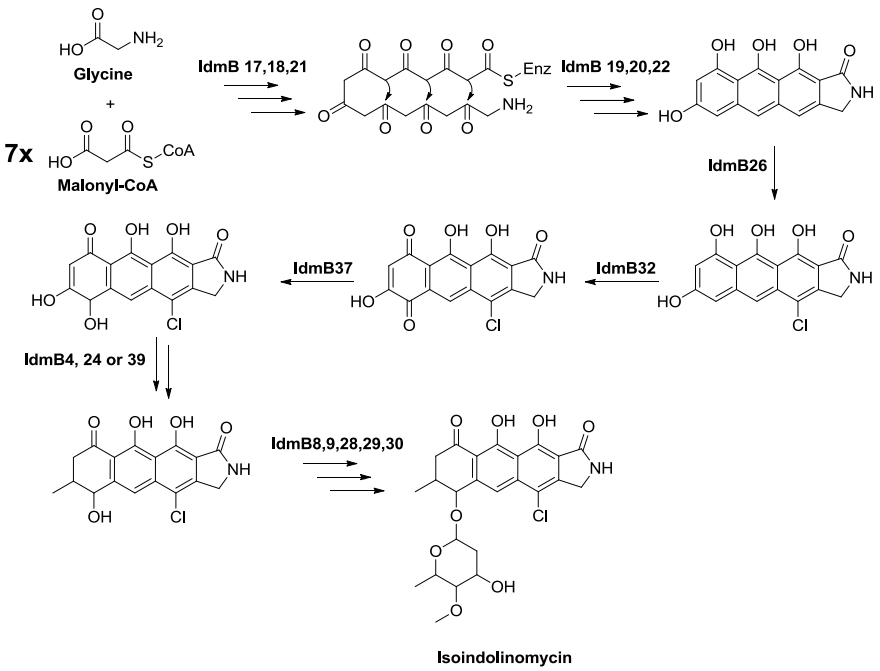


Figure 4-6. Initial proposed biosynthetic pathway of IDM.

Table 4-5. Deduced functions of Orfs in the biosynthesis gene cluster of IDM.

Orf	Amino acids (aa)	Proposed Function	Blast hit protein [Origin]	Identity/ Similarity (%)	Protein ID
1	110	Hypothetical protein	Hypothetical protein [*Streptomyces* sp. NRRL S-813]	28/65	WP_030186183
2	474	Transporter	Drug resistance transporter, EmrB/QacA subfamily [Streptomyces iranensis]	249/503	CDR06316
3	359	Aminotransferase	Aminotransferase [Streptomyces lividans 1326]	170/360	EOY48903
4	338	Methyltransferase	Lct35 [*Streptomyces rishiriensis*]	201/322	ABX71118
5	260	Dehydrogenase	Dehydrogenase [*Streptomyces fulvoviolaceus*]	134/251	WP_030619294
6	243	Regulatory protein	Regulatory protein TetR [*Streptomyces iranensis*]	120/227	CDR06320
7	452	Oxidase	FAD linked oxidase domain-containing protein [*Amycolatopsis decaplanina*]	228/436	WP_007031970
8	457	Dehydratase	NDP-hexose 2,3-dehydratase [*Streptomyces* sp. NRRL S-1813]	248/444	WP_030984725
9	357	dTDP-1-glucose synthase	Putative dTDP-1-glucose synthase [*Streptomyces vietnamensis*]	199/352	ADO32770
10	101	Monooxygenase	Monooxygenase [*Streptacidiphilus anmyonensis*]	40/74	WP_042428701
11	451	Na: H Exchanger	Sodium:proton exchanger [*Streptomyces* sp. NRRL WC-3773]	255/399	WP_031005911
12	253	Thioesterase	Lct34 [*Streptomyces rishiriensis*]	109/241	ABX71117
13	449	Pyruvate carboxylase	Pyruvate carboxylase subunit A [*Streptomyces iranensis*]	343/447	WP_044581358
14	153	Biotin carboxyl carrier protein	Acetyl-CoA carboxylase biotin carboxyl carrier protein subunit [*Streptomyces iranensis*]	75/149	WP_052701348
15	571	Acetyl-CoA carboxylase	Acetyl-CoA carboxyl transferase [*Streptomyces glaucescens*]	332/557	WP_052413654
16	106	Monooxygenase	Lct33 [*Streptomyces rishiriensis*]	58/101	ABX71116
17	408	Beta-ketoacyl synthase	Lct32 [*Streptomyces rishiriensis*]	277/400	ABX71115
18	419	Beta-ketoacyl synthase	Lct31 [*Streptomyces rishiriensis*]	336/416	ABX71114
19	130	Polyketide cyclase	Lct30 [*Streptomyces rishiriensis*]	84/129	ABX71113
20	116	Polyketide cyclase	Lct29 [*Streptomyces rishiriensis*]	70/109	ABX71112
21	510	Adenylation	Lct28 [*Streptomyces rishiriensis*]	310/514	ABX71111
22	156	Polyketide cyclase	Lct27 [*Streptomyces rishiriensis*]	115/152	ABX71110
23	83	Acyl carrier protein	Lct26 [*Streptomyces rishiriensis*]	37/80	ABX71109
24	348	Methyltransferase	Lct25 [*Streptomyces rishiriensis*]	210/349	ABX71108
25	84	Acyl carrier protein	Lct24 [*Streptomyces rishiriensis*]	48/84	ABX71107
26	409	Tryptophan halogenase	FAD-binding monooxygenase [*Streptomyces iranensis*]	292/409	WP_044569677
27	638	Regulatory protein	Lct23 [*Streptomyces rishiriensis*]	278/551	ABX71106
28	334	Ketoreductase	dTDP-4-keto-6-deoxyhexose 2,3-reductase [*Streptomyces halstedii*]	224/326	BAD08364
29	265	Ketoreductase	NDP-hexose 4-ketoreductase [*Streptomyces chattanoogensis*]	132/248	KPC66697
30	430	Glycosyl transferase	Glycosyl transferase [*Amycolatopsis mediterranei*]	166/404	WP_013227238
31	139	Hypothetical protein	Hypothetical protein [*Propionibacterium* sp. KPL1844]	78/134	WP_023034785
32	491	Dioxygenase	Lct41 [*Streptomyces rishiriensis*]	282/468	ABX71124
33	107	ABM	Lct33 [*Streptomyces rishiriensis*]	51/104	ABX71116
34	106	ABM	Lct42 [*Streptomyces rishiriensis*]	48/99	ABX71125
35	247	Abhydrolase	Carboxymethylenebutenolidase [*Streptomyces* sp. AcH 505]	174/247	WP_041988759
36	93	Hypothetical protein	Hypothetical protein [*Streptomyces* sp. NRRL F-5126]	56/85	WP_030904778
37	313	Quinone reductase	Lct43 [*Streptomyces rishiriensis*]	174/308	ABX71126
38	102	Hypothetical protein	Hypothetical protein [*Solirubrobacter soli*]	38/74	WP_028064479
39	348	Methyltransferase	Methyltransferase [*Streptomyces griseus*]	123/335	WP_030766712

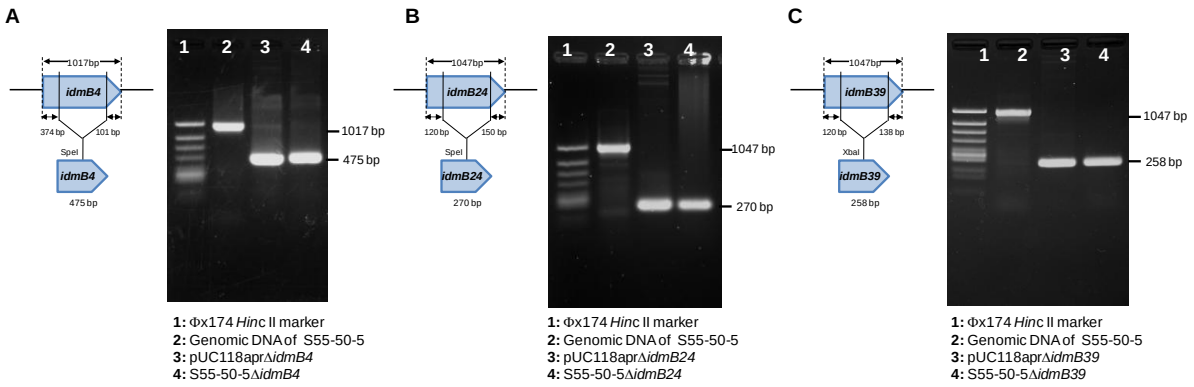


Figure 4-7. Gene inactivation of the methyltransferases in the biosynthetic gene cluster of isoindolinomycin. (**A**) *idmB4*, (**B**) *idmB24*, (**C**) *idmB39*.

The deletion of *idmB39* revealed that it may be responsible for the *O*-methylation in the sugar moiety. This deduction is based on the observation that IDM production was abolished, and an unknown metabolite, most likely an IDM lacking a methyl group in its sugar moiety, was detected (Figure 4-8A and B). Moreover, the detection of IDM aglycone in the deletion mutant further confirmed that IdmB39 is not responsible for the transfer of a methyl moiety to the tetracyclic structure (Figure 4-8C). IDM aglycone has been detected in the culture extract of S55-50-5, but due to its low yield, the structure can only be predicted from LC-MS/MS spectra (Figure 4-9).

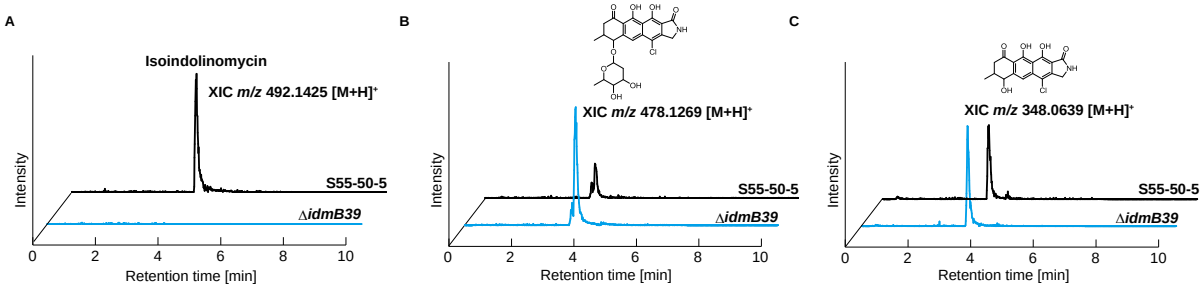


Figure 4-8. Comparative HRESI-MS analysis of the culture extracts from S55-50-5 and the *idmB39* deletion mutant showing (**A**) the absence of IDM (**B**) the accumulation of a possible IDM lacking a methyl group in its sugar moiety and (**C**) the detection of IDM aglycone.

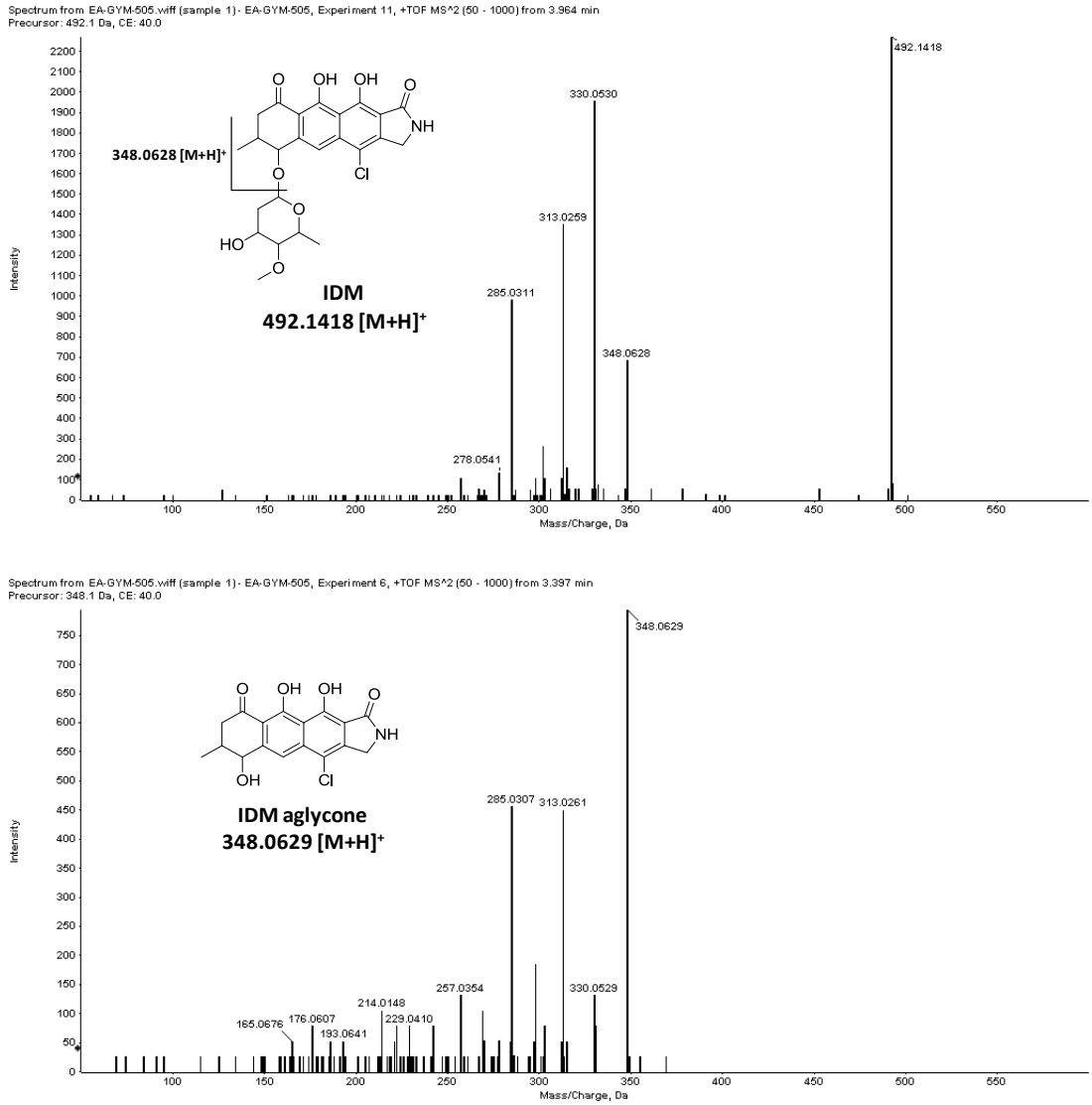


Figure 4-9. LC-MSMS spectra of IDM and IDM aglycone.

In contrast, the deletion of *idmB4* abolished the production of both IDM and its aglycone, indicating that IdmB4 may be the enzyme catalyzing the transfer of methyl substituents to the core structure (Figure 4-10). Here, I was expecting the accumulation of an intermediate that allows the elucidation of the reaction, but no intermediate was detected. This result is intriguing because it suggests a possible unprecedented methylation by IdmB4, which may occur during the generation of the polyketide chain. Further experiments and gene complementation are required to confirm the assumption.

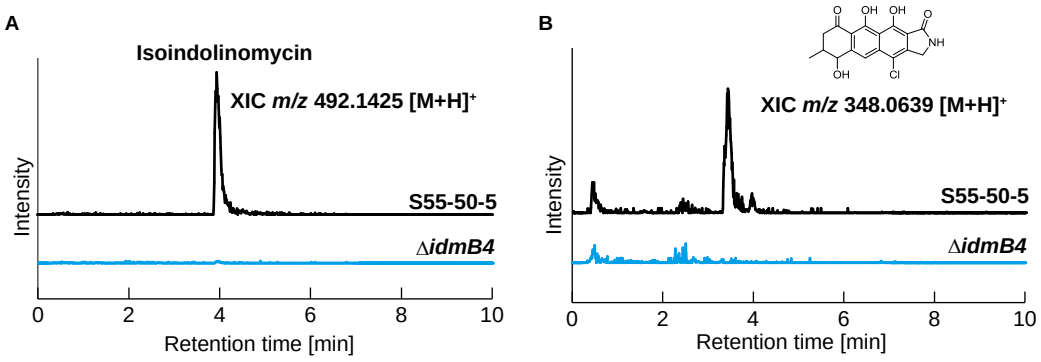


Figure 4-10. Comparative HRESI-MS analysis of the culture extracts from S55-50-5 and the *idmB4* deletion mutant showing the loss of IDM and its aglycone.

In contrast to *idmB4* and *idmB39*, the deletion of *idmB24* did not terminate the production of IDM (Figure 4-11A). Although the result suggests that *idmB24* may not be related to the biosynthesis of IDM, an unknown metabolite accumulated in the deletion mutant (Figure 4-11B). The isolation and structural elucidation of this metabolite may provide information on the function of this enzyme.

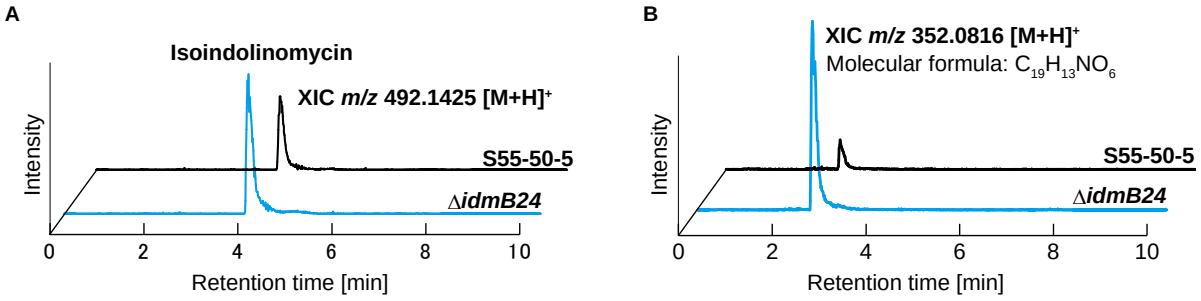


Figure 4-11. Comparative HRESI-MS analysis of the culture extracts from S55-50-5 and the *idmB24* deletion mutant showing (**A**) the detection of IDM and (**B**) the accumulation of an unknown metabolite.

4.2 Substrate determination of the stand-alone adenylation domain (IdmB21).

Methods

Preparation of recombinant IdmB21 protein

The primers listed in Table 4-6 were used to amplify the *idmB21* gene from S55-50-5 genomic DNA. The PCR-amplified 1,551 bp DNA fragment was cloned into the pT7blue vector to produce IdmB21-pT7, and the plasmids were sent for DNA sequencing (Greiner Bio-One, Tokyo). IdmB21-pT7 was then digested with *Nco* I and *Hind* III and cloned into the same sites of pHis8 vector to give IdmB21-pHis8. IdmB21-pHis8 was then used to transform *E. coli* BL21 (DE3). *E. coli* BL21 (DE3) transformed with IdmB21-pHi8 were first pre-cultured in LB medium supplemented with 50 μg mL^{-1} kanamycin at 37°C. The pre-culture was then inoculated into TB medium, which is also supplemented with 50 μg mL^{-1} kanamycin, and grown at 37°C for 90 min. For induction, the culture was cooled on ice for 10 min; then, isopropyl β-thiogalactopyranoside (final 0.1 mM) was added, and the culture was cultivated for an additional 12 h at 18°C.

After cultivation, the cells were harvested by centrifugation at 3,900×*g* for 10 min. Then, the cells were washed and suspended with wash buffer (50 mM Tris-HCl pH 8.0, 0.5 M NaCl, 20 mM imidazole pH 8.0, 20% Glycerol). The cell suspension was homogenized by sonication, and the lysate was centrifuged at 35,000×*g* for 20 min. IdmB21 was purified by applying the supernatant to a column loaded with Ni-NTA Superflow resin (Qiagen, Tokyo, Japan). The column was washed with wash buffer, and IdmB21 was eluted with elution buffer (50 mM Tris-HCl pH 8.0, 0.5 M NaCl, 250 mM imidazole pH 8.0, 20% Glycerol). The eluent containing the IdmB21 was concentrated with Vivaspin 10,000 MWCO (Millipore).

Table 4-6. Primer sets used to construct the *idmB21* expression plasmid.

Primer	Oligonucleotide sequence (5' to 3')*	Restriction enzyme
IdmB21-pHis8-Fw	GGGCCATGGCTCTGAACGTACGCTGTAC	*Nco* I
IdmB21-pHis8-Rv	GGGAAGCTTTCAGTCGGGGTCGAGCC	*Hin*d III

*Restriction enzyme site underlined.

IdmB21 *in vitro* assay

The IdmB21 assay was performed in 100 μL reaction mixture (Table 4-7) and incubated at 30°C for 30 min. After incubation, the reaction mixture was centrifuged at 15,000 rpm for 1 min before proceeding to the malachite green phosphate assay.

Table 4-7. Reaction mixture of IdmB21 assay.

200 mM	Tris-HCl pH 8
5 mM	$MgCl_2$
0.1 mM	ATP
1 mM	Amino acid*
0.02 U	Pyrophosphatase
0.5 mM	DL-Dithiothreitol (DTT)
285 μg mL^{-1}	IdmB21
	MilliQ
Total	100 μl

*20 L-amino acids were tested, each in a different reaction mixture

Malachite green phosphate assay

The malachite green phosphate assay was performed according to the protocol in the assay kit (Bioassay Systems, USA). Eighty microliters of the reaction product (from IdmB21 assay above) was added to the 20 μL of working reagent before incubating at room temperature for 30 min. Then, 900 μL of milliQ water was added to the mixture, and the absorbance was measured at 620 nm.

Results and discussion

Another point of interest in the biosynthesis of IDM is the formation of the 5-membered ring in the core structure. Based on the structure, an amino acid, presumably glycine, is predicted to act as the starter unit in the type II PKS system. To do this, a glycine must be adenylated before loading onto an ACP and being incorporated into the polyketide chain backbone (Figure 4-12). A recent report on streptothricin biosynthesis demonstrated that 2 stand-alone adenylation (A) domains each catalyze a reaction; one adenylates a starter unit L-β-lysine for PCP loading, and the other adenylates L-β-lysine for the elongation of the L-β-lysine peptide chain.[88] Therefore, the adenylation reaction in IDM biosynthesis is most likely catalyzed by the stand-alone A domain (IdmB21) found in the gene cluster. Sequence analysis of *idmB21* revealed that the signature code for this gene is DLFFLCLLSK, and it is predicted by a web software NRPSpredictor2 to recognize glycine or alanine.[86] To determine the substrate for IdmB21, the malachite green phosphate assay was performed. Here, the recombinant IdmB21 protein was purified (Figure 4-13A) and used in the in vitro assay with 20 amino acids. The reaction products were then tested with the malachite green phosphate kit to measure phosphate liberated from the degradation of ATP during the adenylation reaction (Figure 4-12).

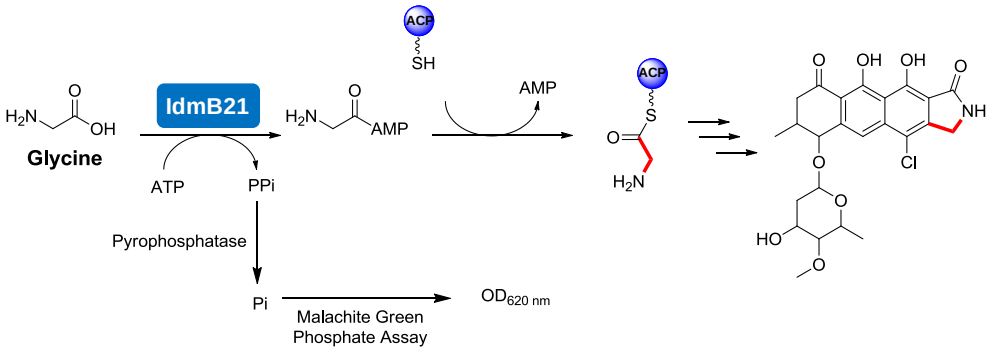


Figure 4-12. Proposed function of IdmB21 and the detection mechanism of the malachite green phosphate assay.

As a result, IdmB21 displayed strict specificity toward glycine only (Figure 4-13B), confirming the involvement of glycine as the starter unit of the polyketide chain backbone of IDM. After the completion of the polyketide chain elongation, the 5-membered ring formation, which may be catalyzed by IdmB19, IdmB20, or IdmB22, may occur. Considering all the data obtained from *in silico* analysis, gene deletions, and *in vitro* assay, the initial proposed biosynthetic pathway for the unprecedented skeleton of IDM is revised and will be discussed in the next section.

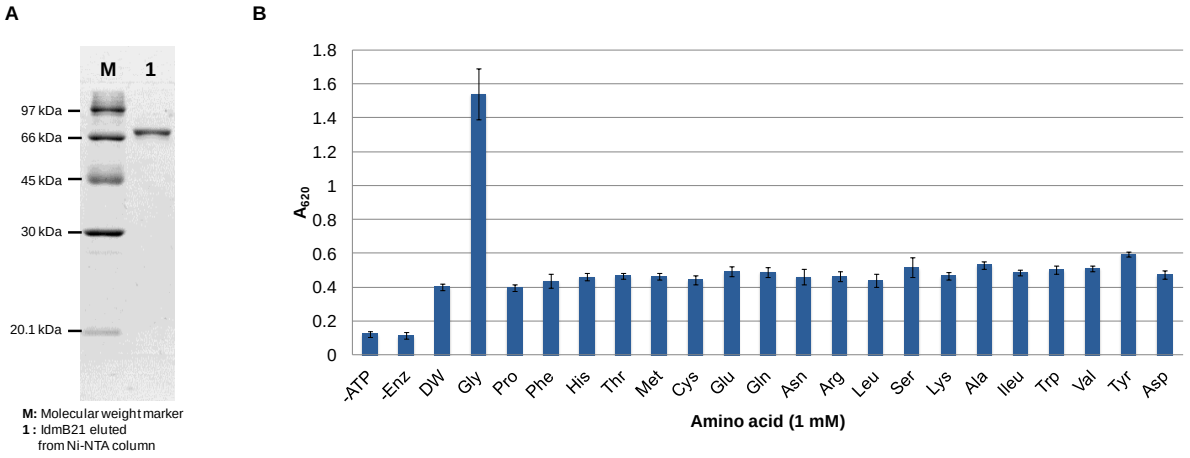


Figure 4-13. (**A**) SDS-PAGE analysis of purified IdmB21. (**B**) Absorbance detected at 620 nm for each amino acid used in the IdmB21 assay. Reactions without ATP (-ATP), enzyme (-Enz), and milliQ water in place of amino acid (DW) serve as negative controls.

4.3 Proposed biosynthetic pathway of isoindolinomycin

Analysis of the genes in the IDM gene cluster revealed that many are homologous to those in the biosynthesis gene cluster of lactonamycin, a compound with the closest similarity to IDM (Table 4-5). I speculate that the homologous genes identified mainly encode proteins that participate in the formation of the polyketide skeleton, such as type II PKS, polyketide cyclase, and tailoring enzymes. This speculation is based on the similarity in the core skeletons of IDM and lactonamycin, particularly the 5-membered ring (A-ring). The precursor incorporation experiments by Zhang et al. (2008) revealed that a glycine or its derivative is involved in the biosynthesis of lactonamycin.[89] However, no evidence was provided with the enzyme and the actual substrate to support the observation. Here, the in vitro and malachite green phosphate assay of IdmB21 demonstrated that the biosynthesis of IDM starts with the adenylation of glycine before the condensation and extension of 7 acetate units by IdmB17 and IdmB18 (Figure 4-14). After the polyketide chain is assembled, cyclizations occur and are most likely catalyzed by IdmB19, IdmB20, and IdmB22. The cyclization mechanism and the timing for the release of the polyketide chain from its ACP are currently unknown. Another reaction with uncertain timing is the halogenation of IDM. Thus, a gene deletion of the halogenase (*idmB26*) is in progress to elucidate the chlorination reaction.

Next, I was intrigued by the attachment of the methyl substituent to the tetracyclic structure. This is because, after the tetracyclic is being formed, methylation at the D-ring may be unprecedented. Here, I hypothesize that the D-ring could lose its aromaticity probably by the activities of IdmB32 and IdmB37 before the transfer of the methyl group proceeds. The disruption of *idmB32* abolished the production of IDM; however, no accumulation of intermediate was detected, giving no information on its reaction. Then, to identify the gene

responsible for the methyl substituent, gene inactivation of all methyltransferases (*idmB4*, *idmB24*, and *idmB39*) found in the gene cluster was performed. As a result, *idmB4* was the most relevant for the methyl group at the D-ring, whereas IdmB39 may catalyze the *O*-methyl transfer at the sugar moiety. The deletion of *idmB4* did not accumulate any detectable intermediates, suggesting that the methyl transfer may occur during the elongation of the polyketide chain. This is because the polyketide intermediate accumulated during the chain elongation step is often unstable.[77] On the other hand, the accumulation of the IDM lacking a methyl group in its oliose moiety in the *idmB39* deletion mutant implied that the methyl transfer occurs after the glycosidic bond is formed.

The biosynthesis of an oliose sugar has been proposed in the biosynthetic study of chromomycin.[57] Based on the report, the biosynthesis of the oliose sugar moiety in IDM is proposed as shown in Figure 4-14. However, no NDP-4,6-dehydratase homolog was identified in the cluster, which is unusual and suggests the possibility of another sugar synthesis pathway. After the sugar biosynthesis, a glycosidic bond between the sugar moiety and IDM aglycone could be formed by the glycosyltranferase (IdmB30), which would be followed by the *O*-methyl transfer by IdmB39 to produce IDM.

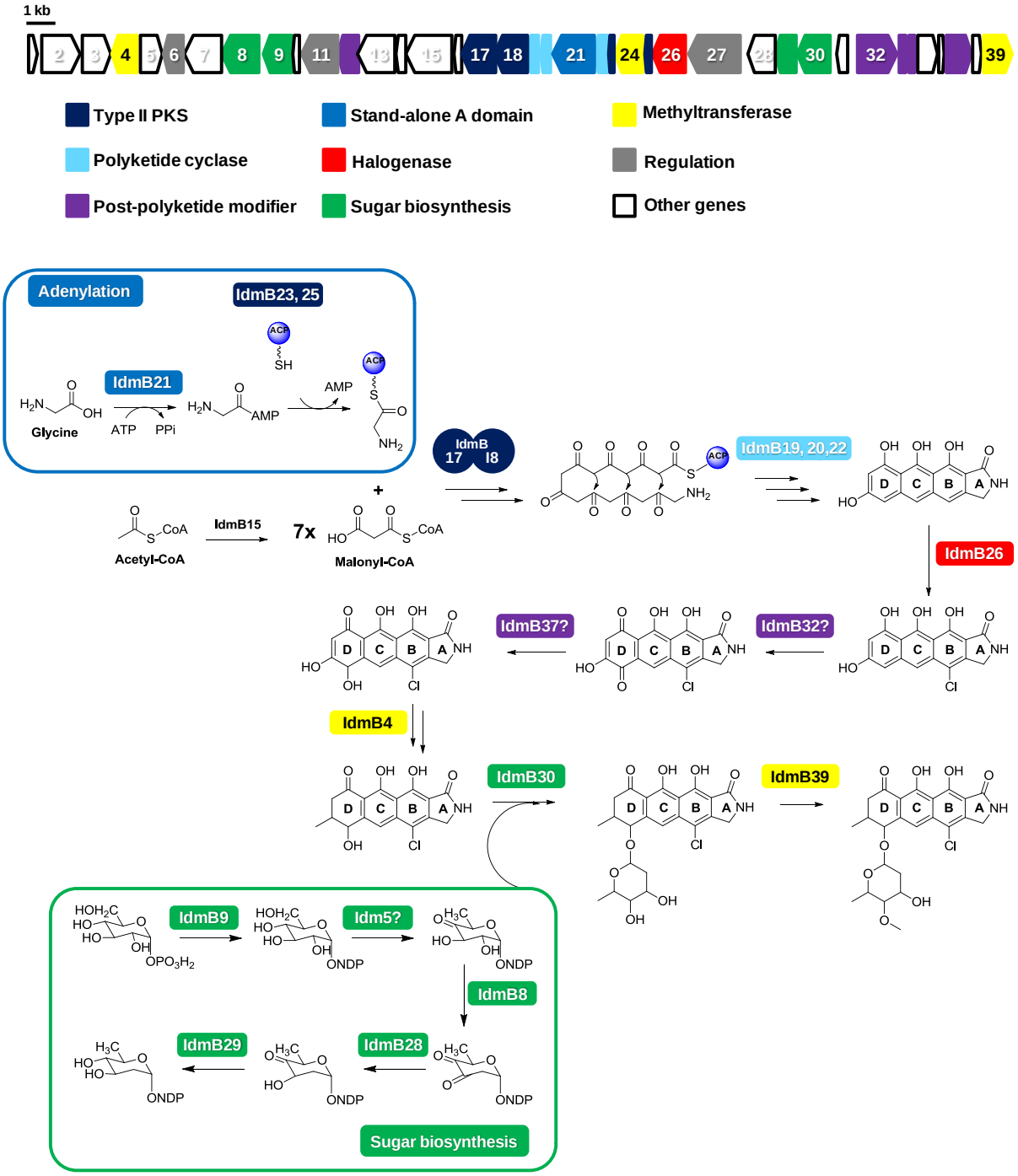


Figure 4-14. Proposed biosynthetic pathway for isoindolinomycin.

4.4 Conclusion

IDM has a novel tetracyclic skeleton with a chloro substituent and an oliose sugar moiety. BLAST searches on the draft genome sequence of *Streptomyces* sp. SoC090715LN-16 retrieved two biosynthetic gene clusters; each has a gene with close homology to tryptophan halogenase. Bioinformatics analysis and gene disruption experiments demonstrated that one of these gene clusters, which contains type II PKS, glycosyltransferase, and methyltransferases, is responsible for IDM biosynthesis.

According to the IDM gene cluster identified, the polyketide chain synthesis by the type II PKS is proposed to be initiated by a glycine as the starter unit. The in vitro assay on the stand-alone A domain (IdmB21) supported this assumption. The involvement of an alternate starter unit in a type II PKS system is rare, although there are some cases reported such as the starter unit propionate for daunorubicin biosynthesis in *Streptomyces peucetius*,[90] and benzoate functions in the synthesis of enterocin in *Streptomyces maritimus*.[91]

Although gene deletion experiments have identified the methyltransferases for each methyl moiety in IDM, there are several unusual features in the gene cluster that still remain to be characterized. For example, the presence of a thioesterase (IdmB12) is rarely associated with type II PKS systems. A study by Yi Tang et al. on an acyltransferase homologue, ZhuC from the initiation module of the R1128 PKS, showed that ZhuC has a thioesterase activity that edits acetyl primer units.[92] The group demonstrated that ZhuC has an editing role to remove the usual acetate starter units that compete with the alternate alkyl acyl primer units. Therefore, it is conceivable that IdmB12 may have a similar function as ZhuC that prevents the initiation of IDM synthesis with acetate. Additional studies are necessary to elucidate the mechanism of glycine incorporation and the full biosynthetic pathway of IDM.

CONCLUSIONS

CONCLUSIONS

Over the last 5 decades, drug discovery and development have relied on natural resources. Actinomycetes have played active roles as a rich source of natural products, supplying a wide variety of bioactive secondary metabolites. Recently, however, efforts have shifted from the conventional screening of bioactive compounds to the generation of synthetic structural analogs. The major setback of natural products screening is the decreasing hit rate of novel bioactive compounds with undiscovered chemical skeletons. On the contrary, genome sequencing reveals that an actinomycete strain is capable of synthesizing more than 20 secondary metabolites. However, most of these metabolites are either not being produced or are produced in an undetectable yield due to the silenced or poorly expressed biosynthetic gene clusters. In this study, to activate the potential gene clusters in actinomycetes, 20 actinomycete strains were used to screen for rifampicin-resistant mutant. Strains that acquired rifampicin resistance were reported to have mutation(s) in their *rpoB* gene (encoding RNA polymerase β-subunit), which is able to activate the production of bioactive compounds as demonstrated in the discovery of a cryptic new type of antibiotic called piperidamycin.

This study started with the determination of the minimum inhibitory concentration (MIC) of rifampicin in each actinomycete strain. MIC values obtained were used to guide the choice of rifampicin concentration for the induction of resistance. A total of 278 spontaneous rifampicin-resistant (*rif)* mutants generated were isolated. Among them, 200 *rif* mutants were selected and cultivated along with their native strains to perform comparative metabolic screening. The screening resulted in the detection of 4 potential overproduced metabolites (**1**, **3**, **6**, and **7**) in the culture extract of the *rif* mutants WK2057-10-39, TW-R50-13, S45-50-3, and S55-50-5, respectively.

The purification of metabolite **1** in WK2057-10-39 gave 4 known phenoxazinones, which include questiomycin A, 2-ethanolamino-3H-phenoxazin-3-one,

2-acetylamino-3H-phenoxazin-3-one, and 2-methylamino-3H-phenoxazin-3-one. In contrast, the isolation of metabolite **3** in TW-R50-13 yielded 2 novel compounds **3** and **3a** (Figure 1), which were elucidated to have an identical 2-amino-3-hydroxycyclopent-2-enone (C_5N) and methylbenzene moieties. Metabolite **6** in S45-50-3 was identified as cinerubin A, whereas metabolite **7** was elucidated as an unprecedented compound with a novel skeleton featuring an isoindolinone structure (Figure 1). This novel compound **7** was found to exhibit potent cytotoxicity against all cell lines tested and antimicrobial activity toward *Staphylococcus aureus*. In contrast, no biological activity was observed for the new compounds **3** and **3a** in the same assays tested. The bioactivity displayed by compound **7** earned its right to be named isoindolinomycin. Additionally, mutation analysis on the *rpoB* gene of the *rif* mutants (TW-R50-13 and S55-50-5) producing the novel compounds revealed that they both have a point mutation at position 437, which is a mutation point known for activating secondary metabolite production.

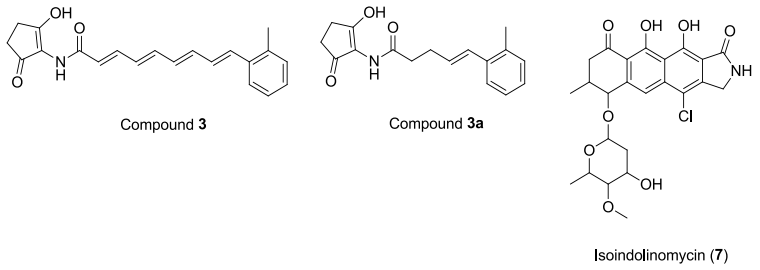


Figure 1. Novel compounds isolated in this study.

The identification of the biosynthetic gene cluster for compounds **3** and **3a** was achieved by *in silico* screening and gene inactivation. Initial investigations on a gene cluster in Contig12 resulted in the identification of the gene cluster for the C_5N moiety only, and the biosynthetic mechanism of the methylbenzene moiety remained elusive. The methylbenzene moiety was then determined to be most likely the product of PKS activity through the incorporation of ^{13}C-labeled precursors. Therefore, to search for the biosynthetic gene cluster

for the methylbenzene moiety, *in silico* screening for PKS-containing gene clusters were performed. A series of gene inactivations on 5 PKS-containing gene clusters led to the discovery of a gene cluster composed of 8 discrete KSs in Contig19 (Figure 2). Additional gene deletion experiments not only verified the MBG gene cluster but also provided insight into the biosynthesis of compounds **3** and **3a**. Unlike any other known PKSs, the MBG cluster contains most of the typical PKS components (KS, ACP, KR, DH, and TE) as stand-alone genes. A constructed phylogenetic tree showed that the KSs in this cluster are distant from the other existing KSs, suggesting the potential of this MBG cluster as a new type of PKS with an unprecedented mechanism.

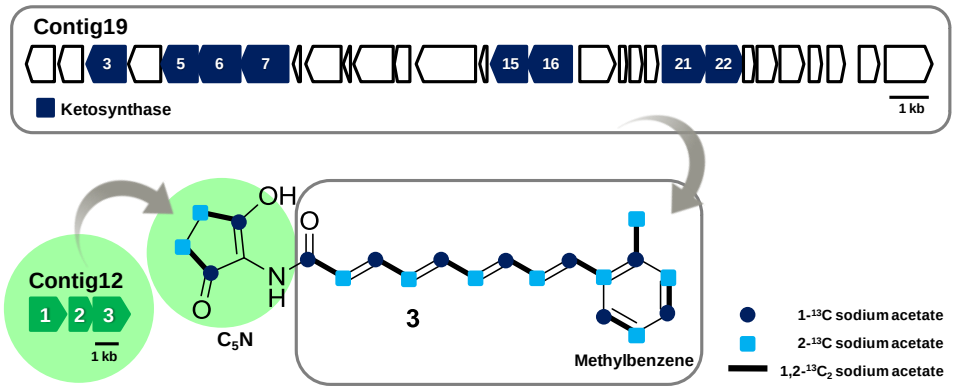


Figure 2. ^{13}C-labeling pattern and biosynthetic gene clusters identified to synthesize compounds **3** and **3a**.

Similarly, the biosynthetic gene cluster of IDM was identified through BLAST searches against the draft genome sequence of S55 followed by gene inactivation (*idmB32*) for confirmation. Based on the identified gene cluster, a biosynthesis pathway for IDM was proposed (Figure 3). In this proposed pathway, I found the attachment of the methyl moiety to the D-ring of IDM intriguing. To identify the gene responsible for the reaction, the gene disruption of 3 methyltransferases (*idmB4*, *idmB24*, and *idmB39*) in the gene cluster was performed. As a result, IdmB4 and IdmB39 were found to catalyze methyl transfer at the

D-ring and the sugar moiety, respectively. Next, for the formation of the unique 5-membered A-ring, IdmB21 was found to have strict specificity toward glycine, which could be adenylated before being incorporated into the elongating polyketide chain by the type II PKS. Collectively considering the data from *in silico* analysis, gene deletions, and in vitro assay, the biosynthesis pathway for the unprecedented skeleton of IDM was proposed.

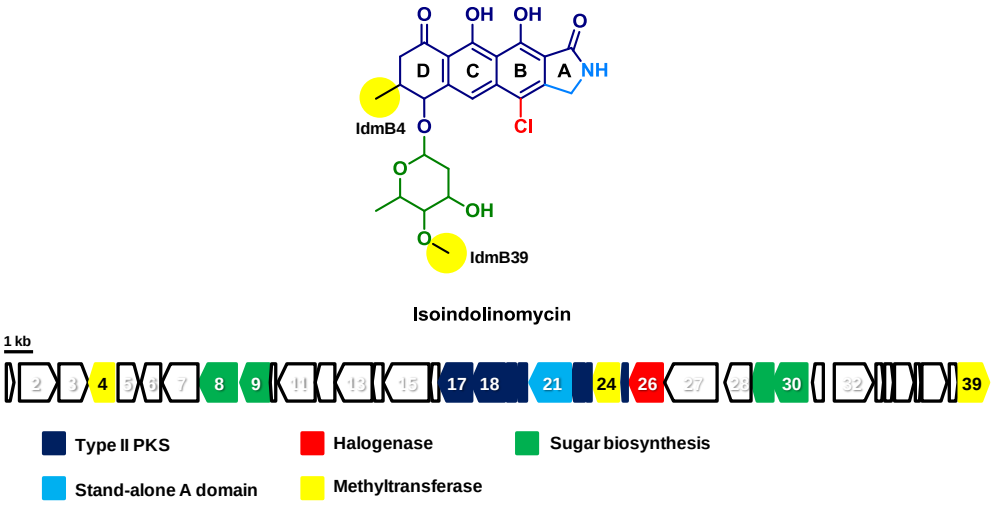


Figure 3. Organization of the biosynthetic gene cluster for IDM. The proposed functions of some genes in the cluster are highlighted with the same color in the structure.

The aim of this study is to explore the "hidden" potential of actinomycetes via the activation of cryptic gene clusters. The goal was achieved when 3 novel compounds were isolated from the *rif* mutants. These unprecedented compounds possess a novel skeleton with potent bioactivity (IDM), and their chemical structures have steered the research to the discovery of possible new biosynthetic mechanisms. Although the full biosynthetic pathways of the novel compounds are yet to be elucidated, this report demonstrated the potential of actinomycetes as a source for natural products discovery.

Naturally, there is no single method that can unlock all cryptic pathways in actinomycetes. Advances in genome sequencing coupled with software tools for genome

annotation have allowed us to predict and identify potential novel natural products. However, in many cases, *in silico* screening works only for highly established natural products and is unable to recognize new types of biosynthetic assemblies. Moreover, the complex regulation of secondary metabolite biosynthesis constantly encourages us that there is more to be studied before we can gain the full potential of actinomycetes. The discovery of new synthetically challenging bioactive substances is expected from these microorganisms, and with them, we can contribute to better disease therapies in the future.

EXPERIMENTAL

EXPERIMENTAL

Media

TSB medium

Tryptone Soya Broth	3.0%
(agar)	1.5%

K medium

Starch	2.5%
Soybean Meal	1.5%
Dry Yeast	0.2%
CaCO3	0.4%
pH	adjust to 6.2

Phenazinomycin producing medium

Glucose	2.0%
Peptone	0.5%
Meat extract	0.5%
Dry yeast	0.3%
NaCl	0.5%
CaCO3	0.3%
pH	adjust to 7.0

GYM (per liter)

Glucose	4 g
Yeast Extract	4 g
Malt Extract	10 g
NZ-amine, type A	1 g
NaCl	2 g
pH	adjust to 7.2

A1 medium

Starch	1.0%
Yeast extract	0.4%
Peptone	0.2%
pH	adjust to 7.0

PC1 medium

Starch	1.0%
Molasses	1.0%
Meat Extract	1.0%
Polypeptone	1.0%
pH	adjust to 7.2

A-94964 producing medium

Glucose	1.0%
Soluble Starch	4.0%
Polypeptone	1.0%
Dry Yeast Extract	0.45%
Corn Steep Liquid	0.50%
Trace Element*	0.10%

Minimal liquid medium (NMMP)(per liter)

$(NH_4)SO_4$	2 g
Casamino acids	5 g
$MgSO_4.7H_2O$	0.6 g
PEG6000	50 g
Trace Element*	1 mL
H_2O	800 mL

At time of use, add (autoclave separately):

NaH_2PO_4/K_2HPO_4 buffer (0.1M, pH6.8)	15mL/80 mL media
Glucose (20%)	2.5mL/80 mL media

R2YE medium (per 100 mL)

Sucrose	10.3 g
K_2SO_4	25 mg
Glucose	1.0 g
$MgCl_2.6H_2O$	1.0 g
Casamino acid	10 mg
Yeast extract	500 mg
TES	573 mg
Trace element*	200 μL
H_2O	85 mL

At time of use, add (autoclave separately):

0.5% KH_2PO_4	1 mL
5M $CaCl_2.2H_2O$	400 μL
20% Proline	1.5 mL
1N NaOH	700 mL

LB medium

Bacto Tryptone	1.0%
Bacto Yeast Extract	0.5%
NaCl	1.0%
(Bacto Agar)	1.5%

TB medium (per liter)

Bacto tryptone	12 g
Bacto yeast extract	24 g
Glycerol	4 mL
H_2O	fill up to 900 mL

At time of use, add (autoclave separately):

KH_2PO_4	2.31 g
K_2HPO_4	12.54 g
H_2O	fill up to 100 mL

Solutions

TE

0.5 M Tris-HCl (pH8.0)	200 μL
1 M EDTA (pH8.0)	1 mL
H_2O	98.8 mL

P1

Tris-HCl (pH 8.0)	6.06 g
EDTA	3.72 g
H_2O	800 mL

P2

NaOH	0.2 N
SDS	1 %

P3

5 N Potassium acetate	60 mL
pH	adjust to 5.5 using glacial acetic acid

P buffer (per 100 mL)

Sucrose	10.3 g
K_2SO_4	25 mg
$MgCl_2.6H_2O$	0.202 g
Trace element*	200 μL
H_2O	80 mL

At time of use, add (autoclave separately):

5.73% TES buffer (pH 7.2)	10 mL
3.68% $CaCl_2.2H_2O$	10 mL
0.5% KH_2PO_4	1 mL

*Trace Element

$ZnCl_2$	40 mg
$FeCl_2.6H_2O$	20 mg
$CuCl_2.2H_2O$	20 mg
$MnCl_2.4H_2O$	20 mg
$NaB_4O_7.10H_2O$	20 mg
$(NH_4)6Mo_7O_{24}.4H_2O$	10 mg
H_2O	fill up to 1L

General experimental methods

Genomic DNA preparation

The actinomycetes strain was cultured in 10 mL of TSB medium for 2-3 days with shaking at 30°C. Cells from 500 μL of culture were harvested by centrifugation (15,000 rpm, 1 min) and resuspended in the same volume of buffer (25 mM Tris-HCl pH 8.0, 25 mM EDTA pH 8.0, 0.3 M sucrose and 2 mg mL^{-1} lysozyme). The mixture was then gently mixed with a rotary mixer for 30 min, 500 μL of 1% SDS was added, and the mixture was incubated for 1 min. After 250 μL phenol-chloroform was added, the sample was gently inverted and centrifuged at 15,000 rpm for 5 min. The supernatant was then transferred into a clean tube and subsequently precipitated with 1/10 volume of 3 M sodium acetate and the same volume of isopropanol. After centrifugation (15,000 rpm, 10 min), the supernatant was removed, and the DNA was washed with 70% (v/v) ethanol. The solvent was then removed by evaporation, and the DNA was dissolved in 100 μL of TE buffer containing RNase. Finally the sample was incubated in a water bath at 37°C for 20 min before the quantity and quality of the DNA was determined by agarose gel electrophoresis.

Plasmid DNA preparation by alkaline lysis

Cells from 1.5 mL of culture were harvested by centrifugation (21,500×*g*, 1 min) and resuspended in 300 μL of P1. The mixture was then lysed with 300 μL P2 and incubated for 5 min on ice. A volume of 300 μL P3 was added, and the mixture was incubated 5-15 min on ice. The resulting mixture was then centrifuged at 21,500×*g* for 10 min. Next, 700 μL of the supernatant was transferred to a new tube, and the same volume of phenol-chloroform was added. The mixture was then inverted and centrifuged at 15,000 rpm for 5 min. The supernatant (500 μL) was again transferred to a new tube, and 0.6 volume of isopropanol was added. The resulting mixture was incubated on ice for 10 min followed by centrifugation at 15,000 rpm for 10 min. The supernatant was then removed, and the remaining DNA was washed with 70%

(v/v) ethanol. The solvent was removed by evaporation, and the DNA was dissolved in 50 μL of TE buffer containing RNase.

Protoplast preparation

The *rif* mutant was pre-cultured using 10 mL of a medium containing 3% TSB supplemented with rifampicin (at MIC) in 50-mL test tubes on a reciprocal shaker at 30°C for 2 days. The seed medium (2 mL) was then inoculated into 10 mL of R2YE medium (supplemented with glycine) in 50-mL test tubes and cultivated on a reciprocal shaker at 30°C for 3 days. After cultivation, the cells were harvested by centrifugation at 3,600×*g* for 10 min before washing with 10.3% sucrose (autoclaved) and centrifuging again. After centrifugation, the cell pellet was resuspended with P buffer, and filter-sterilized lysozyme (final concentration 2 mg mL^{-1}) was added to the mixture. The mixture was then incubated at 30°C for 30 min to 1 h for the enzymatic reaction of lysozyme to take place. After confirmation under a microscope, the protoplasts were filtered, centrifuged at 3,600×*g* for 10 min, and washed with P buffer twice. Finally, the protoplasts were resuspended with P buffer before storing at –80°C.

Transformation of *Streptomyces*

Frozen protoplasts were thawed, and 5 μL of plasmid DNA was added. Then, 500 μL of PEG 1450 (25%) was added immediately, and the mixture was incubated at room temperature for 30 sec. Subsequently, 700 μL of P buffer was added and plated on a R2YE agar plate. The plate was incubated for 14 to 16 h at 30°C before overlaying with a selection marker and incubating for an additional 56 h. Successful transformants are colonies that grow after the addition of the selection marker.

Instruments

UHPLC Column	Shiseido Capcell Pak C18 2.0 mm × 50 mm
UHPLC Pump	JASCO X-LC 3185PU
UHPLC PDA	JASCO X-LC 3110MD
UHPLC Autosampler	JASCO X-LC 3159AS
UHPLC Degasser	JASCO X-LC 3080DG
UHPLC Mixing unit	JASCO X-LC 3180MX
UHPLC Valve unit	JASCO X-LC 3180HV
UHPLC Column oven	JASCO X-LC 3067CO

HPLC Column	Senshu Pak PENGASIL ODS 4.6ϕ × 250 mm
HPLC Pump	JASCO PU-2089 Plus
HPLC PDA	JASCO MD-2010 Plus
HPLC Autosampler	JASCO AS-2057 Plus

Prep. HPLC Column	Senshu Pak PENGASIL ODS 20ϕ × 250 mm
Prep. HPLC UV/VIS Detector	SSC-5410, Senshu scientific co. ltd
Prep. HPLC Pump	SSC-3462, Senshu scientific co. ltd

LCMSMS

LC Column	Shiseido Capcell Pak C18 2.0 mm × 50 mm
Pump	Shimadzu LC-30AC
PDA	Shimadzu SPD-M20A
Autosampler	Shimadzu SIL-30AC
Column oven	Shimadzu CTO-30A
Degasser	Shimadzu DGU-20A_5
MS	AB Sciex Triple TOF^{TM} 5600

NMR	JEOL Superconducting Magnet
	600 MHz (^{1}H NMR), 150 MHz (^{13}C NMR)

APPENDIX:

NMR spectra of novel compounds

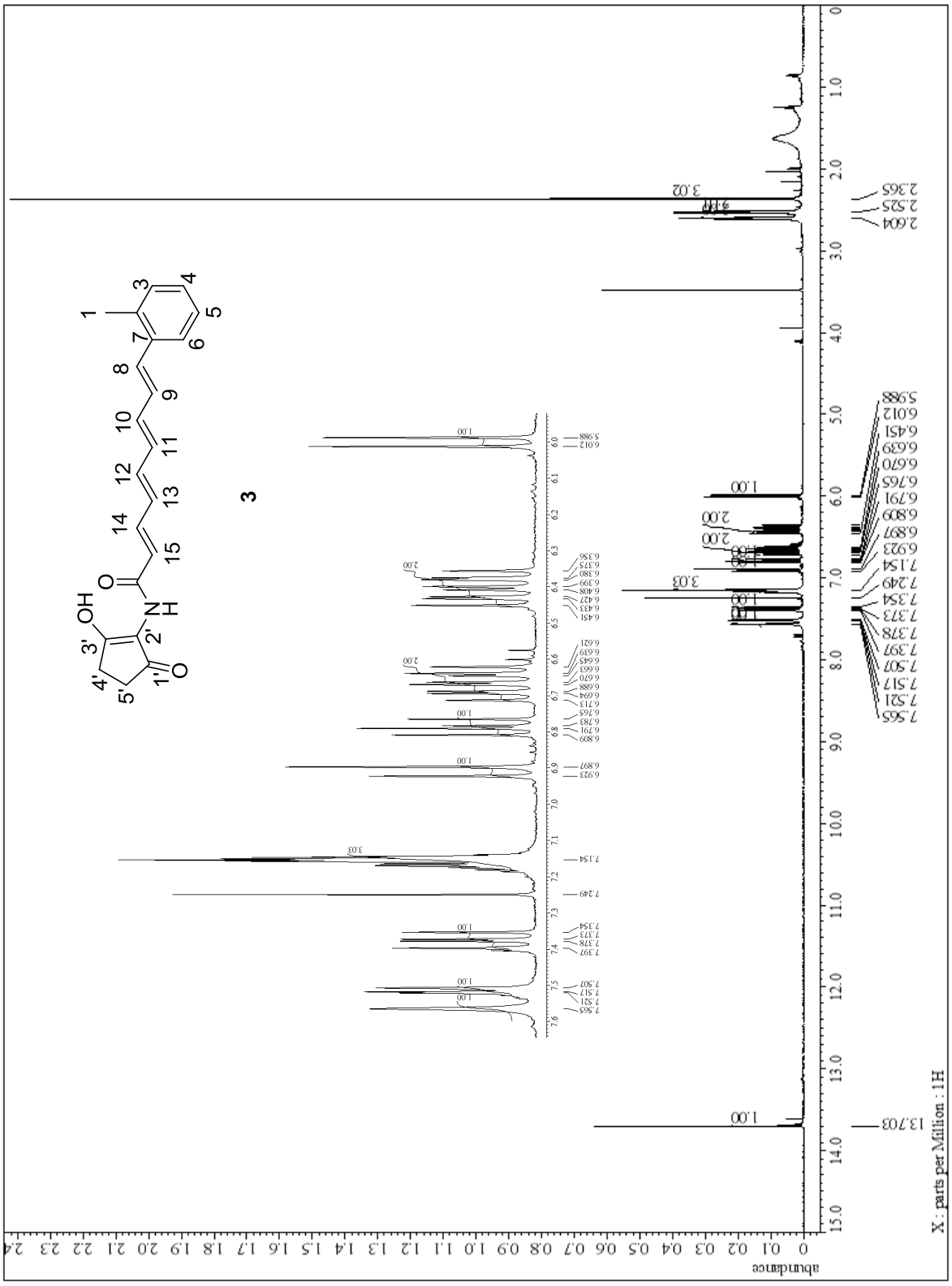


Figure S1. ^{1}H NMR (600 MHz, $CDCl_3$) spectrum of **3**.

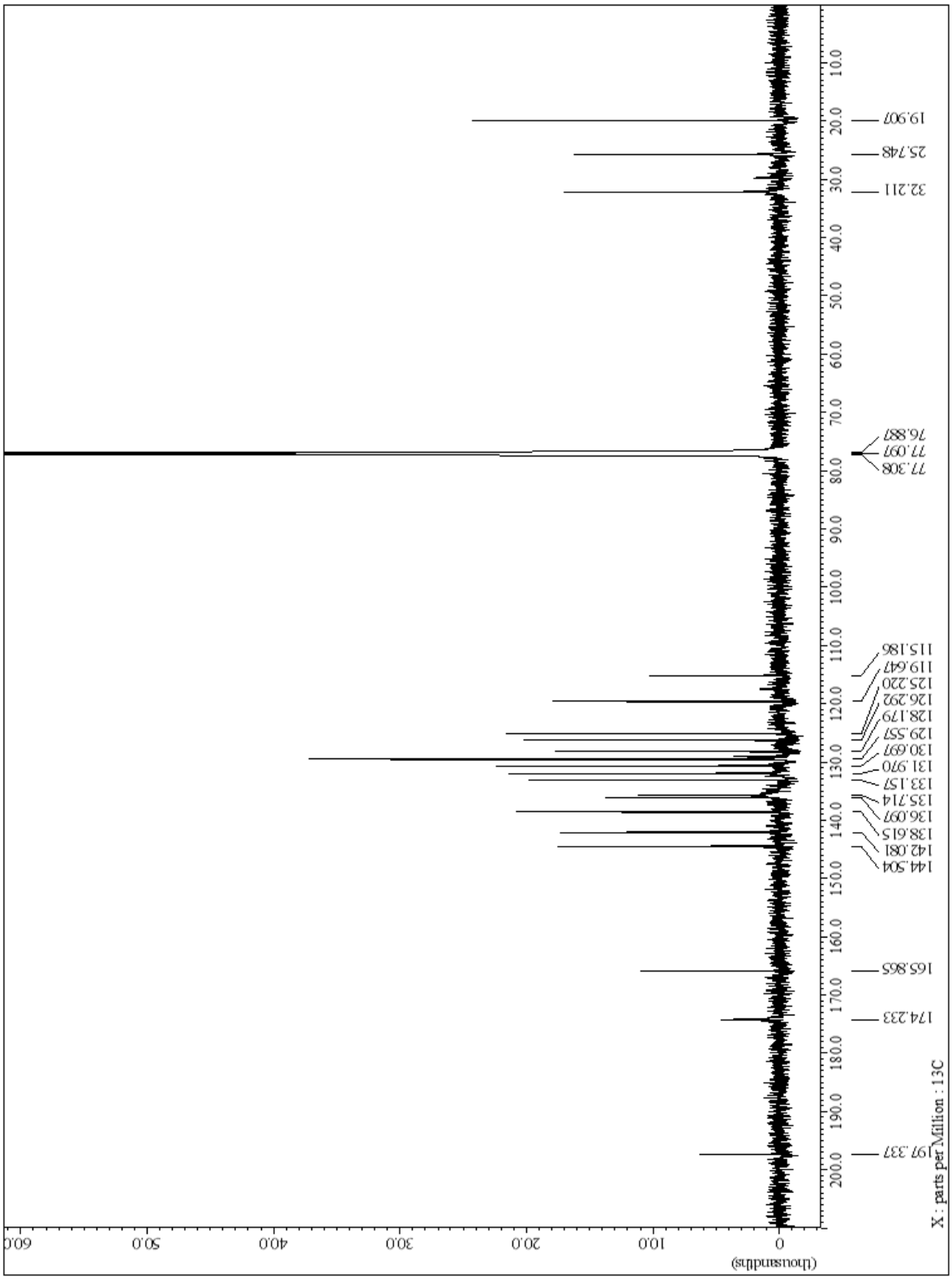


Figure S2. ^{13}C NMR (150 MHz, $CDCl_3$) spectrum of **3**.

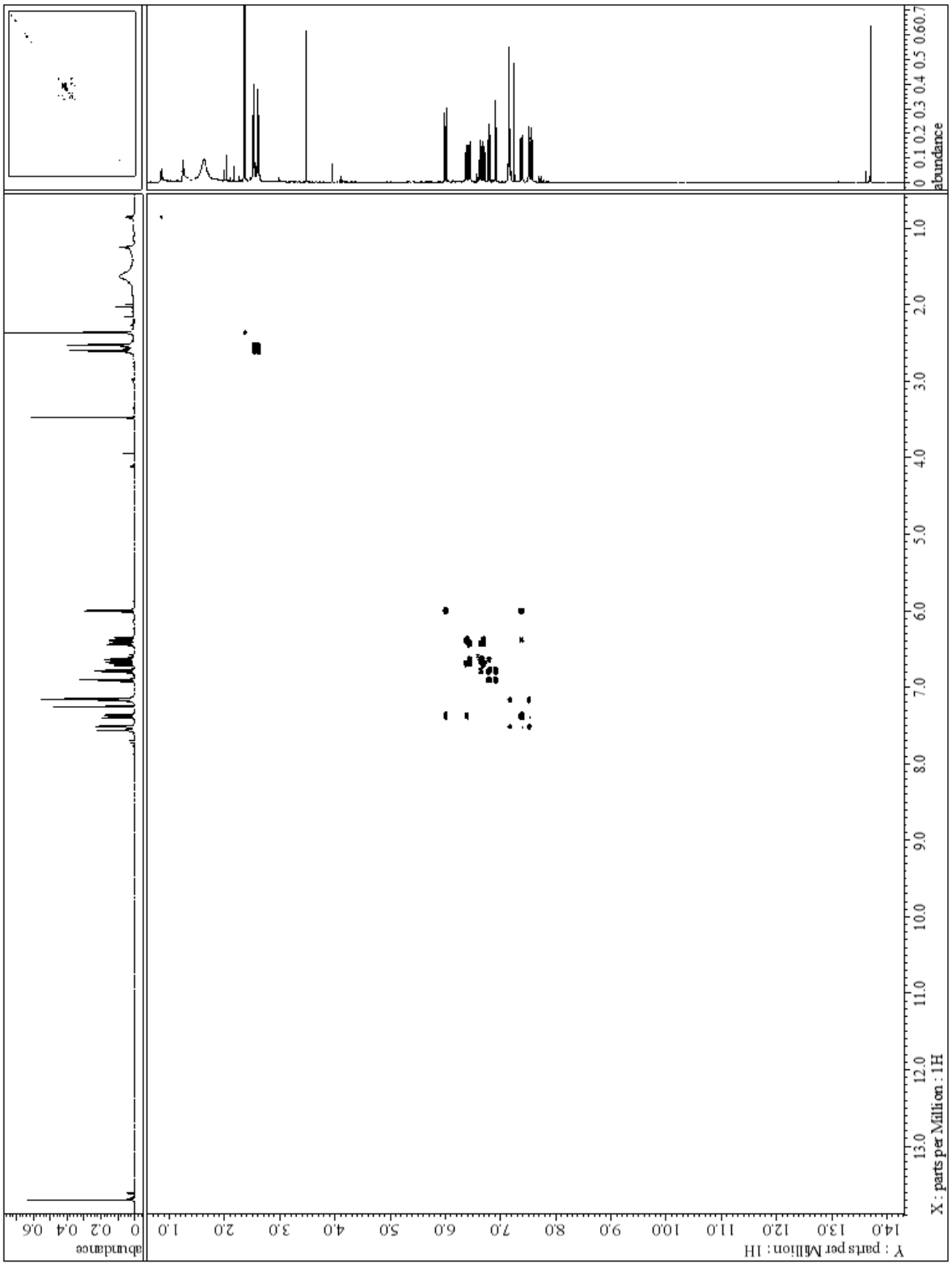


Figure S3. ^{1}H-^{1}H COSY ($CDCl_3$) spectrum of **3**.

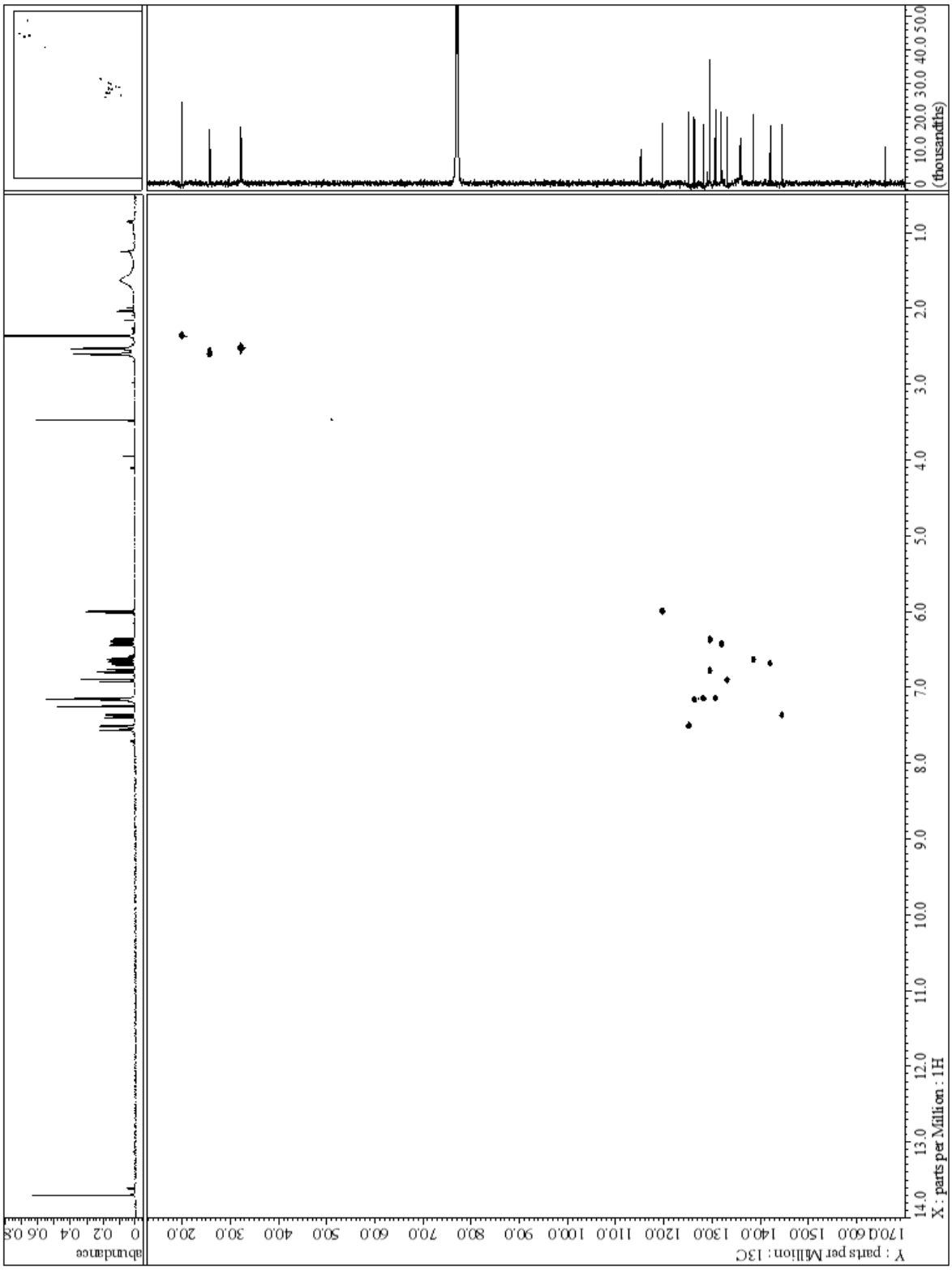


Figure S4. HSQC ($CDCl_3$) spectrum of **3**.

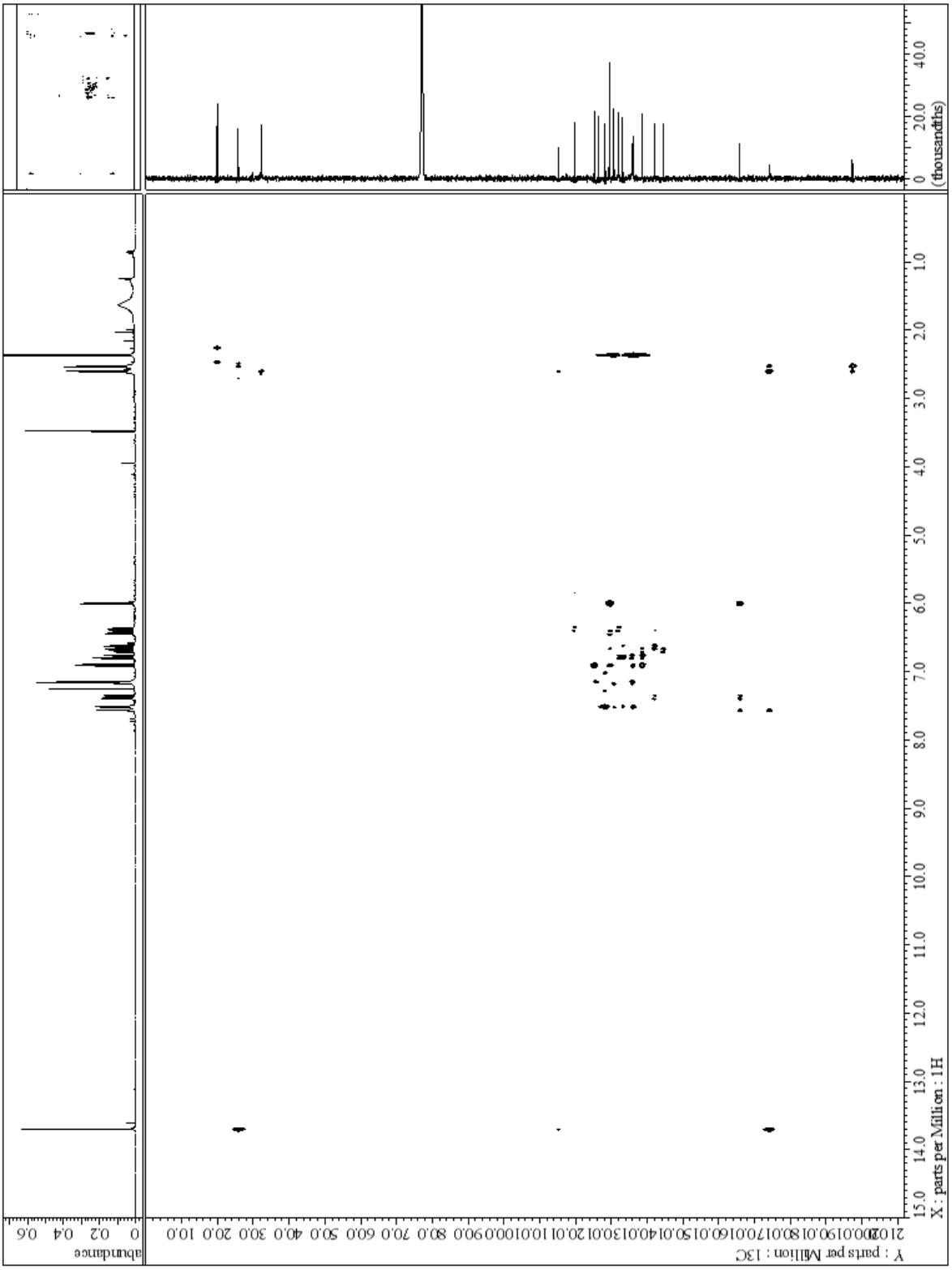


Figure S5. HMBC ($CDCl_3$) spectrum of **3**.

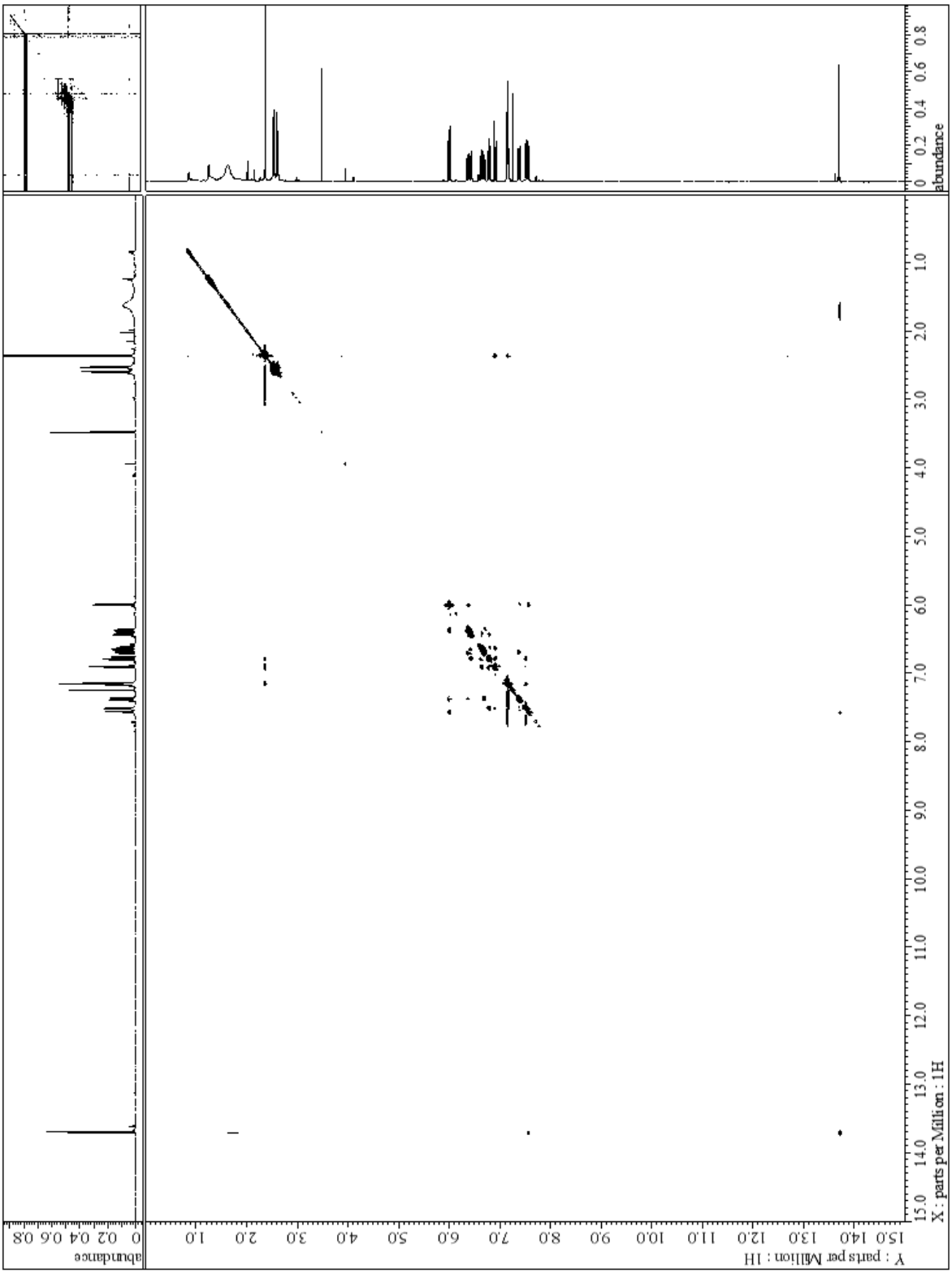


Figure S6. NOESY ($CDCl_3$) spectrum of **3**.

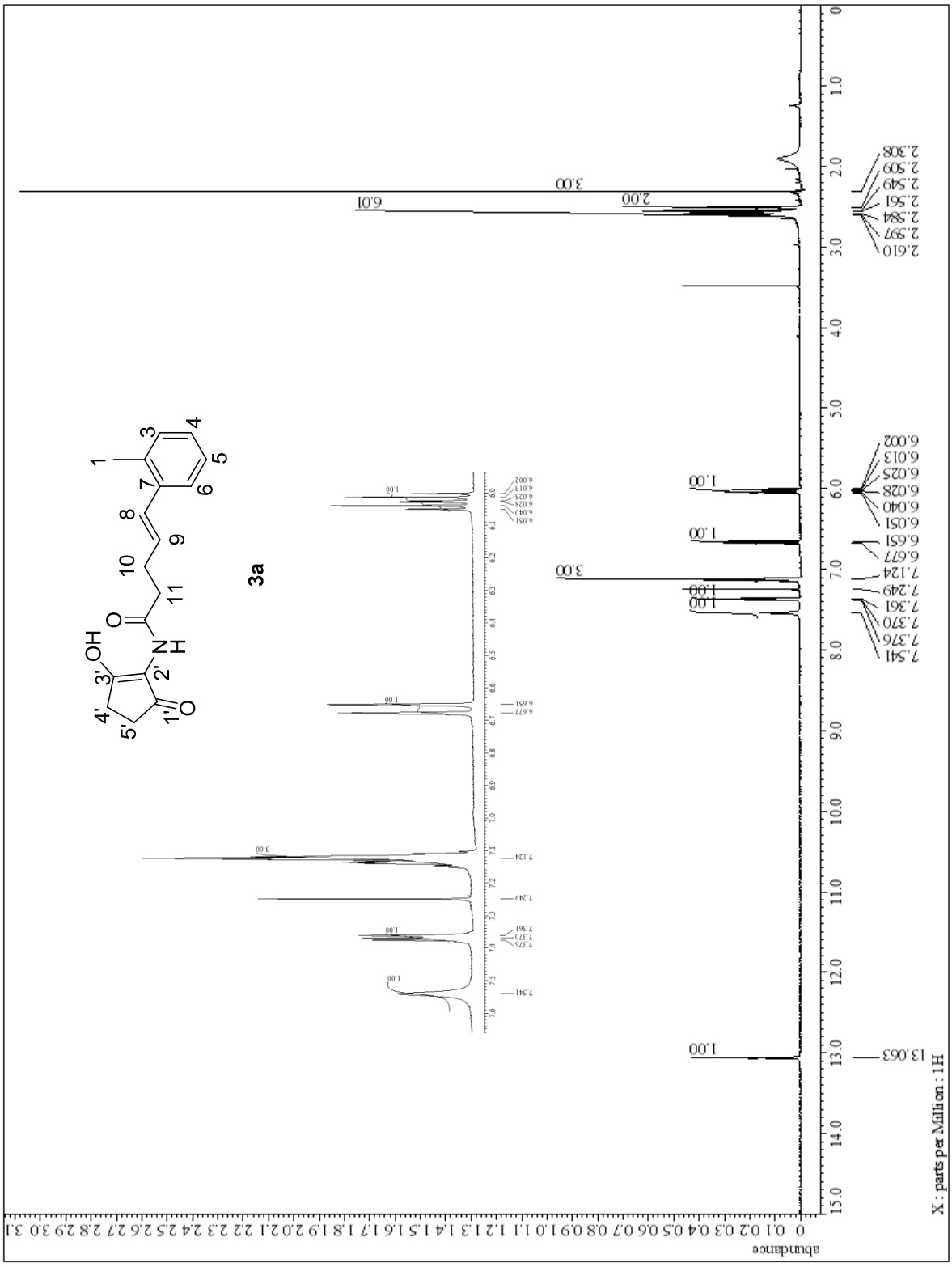


Figure S7. ^{1}H NMR (600 MHz, $CDCl_3$) spectrum of **3a**.

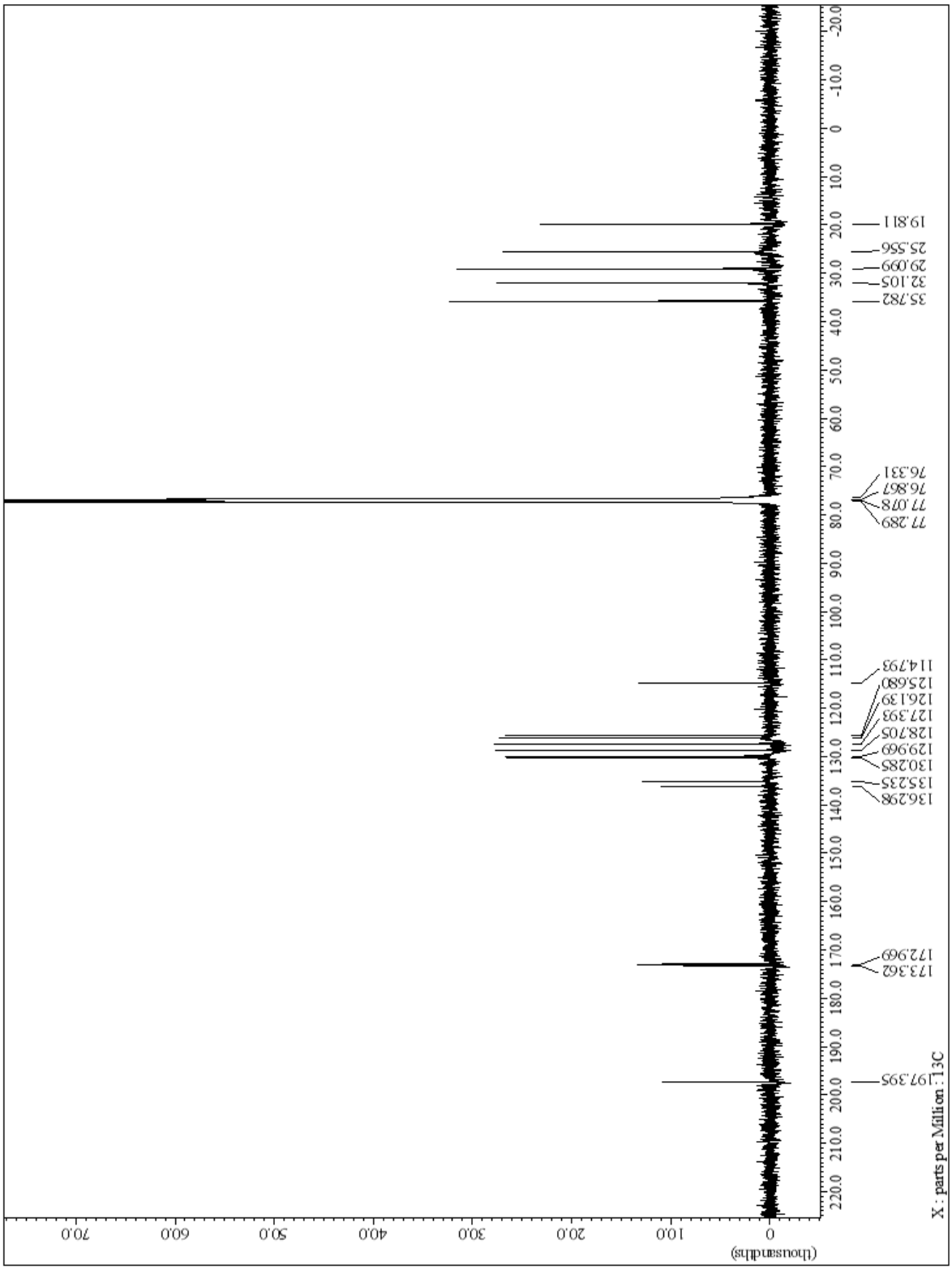


Figure S8. ^{13}C NMR (150 MHz, $CDCl_3$) spectrum of **3a**.

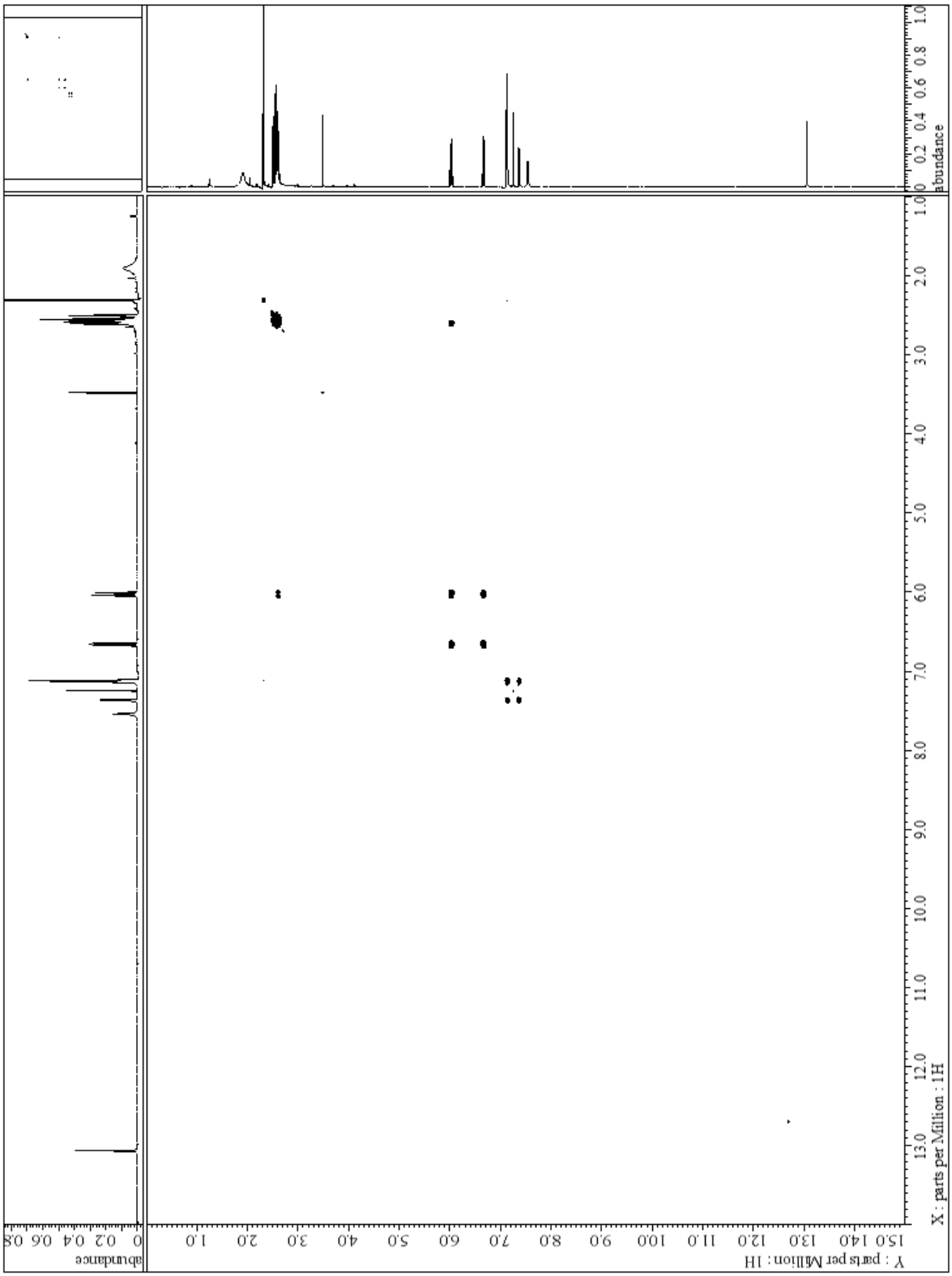


Figure S9. ^{1}H-^{1}H COSY ($CDCl_3$) spectrum of **3a**.

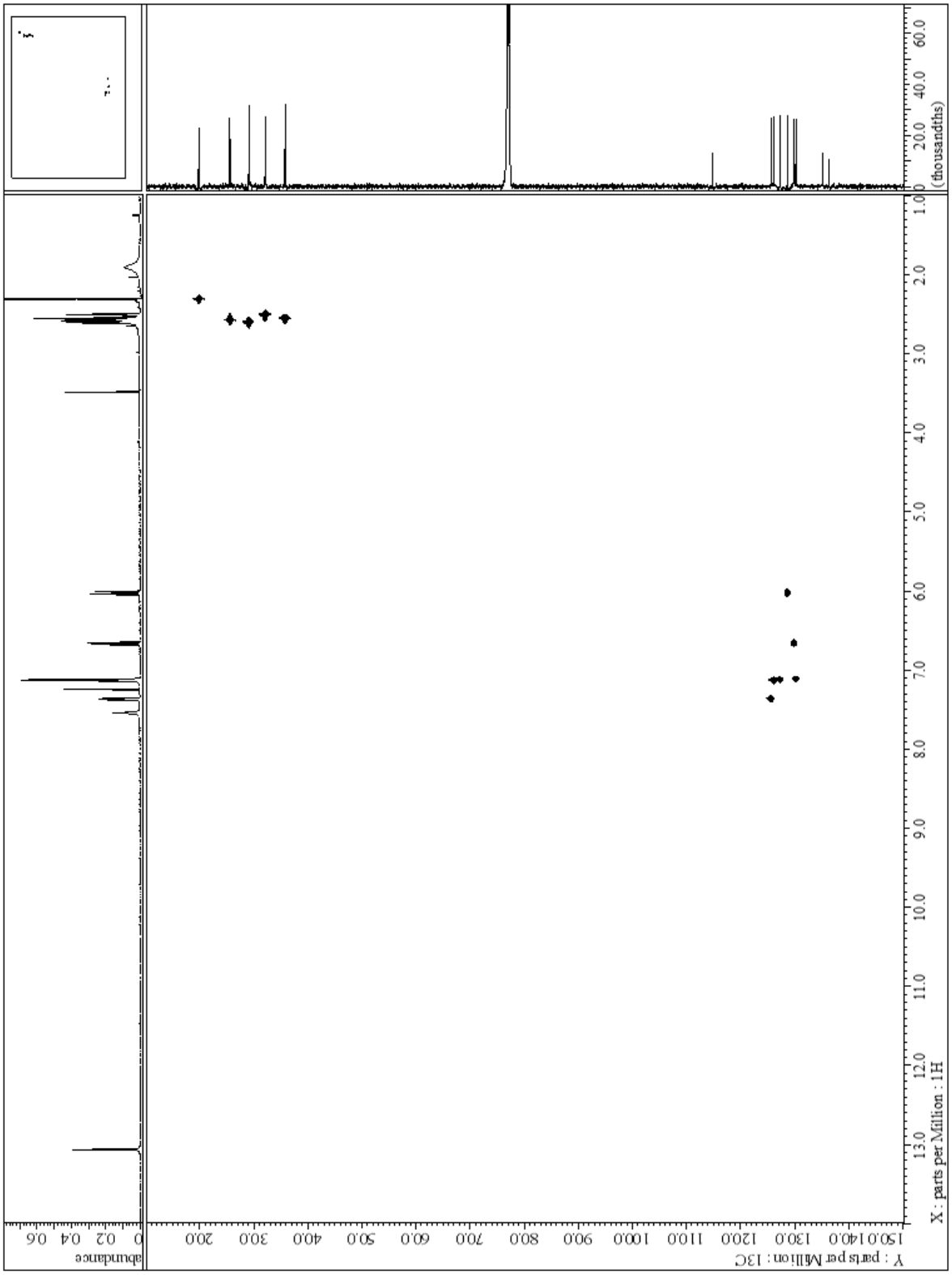


Figure S10. HSQC ($CDCl_3$) spectrum of **3a**.

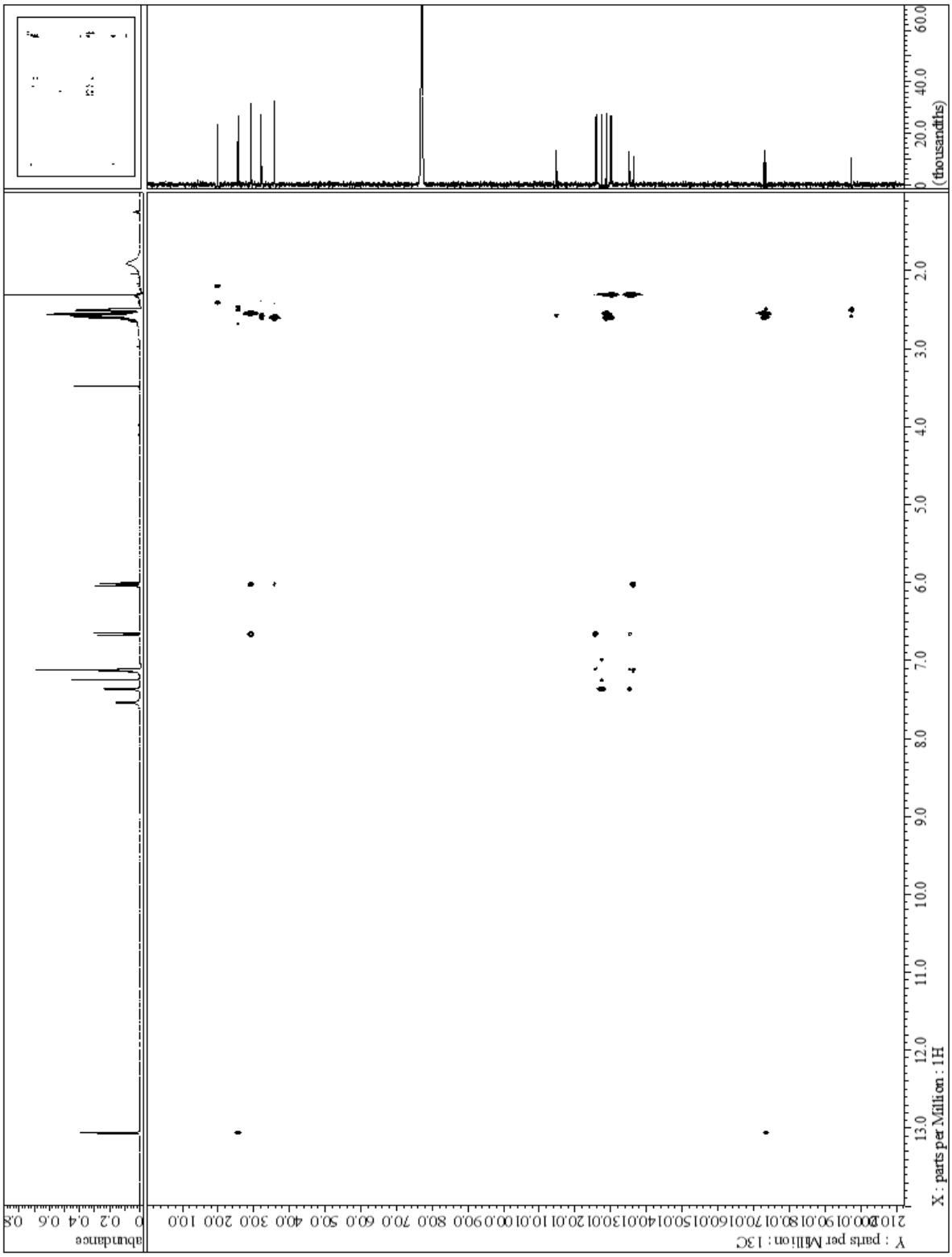


Figure S11. HMBC ($CDCl_3$) spectrum of **3a**.

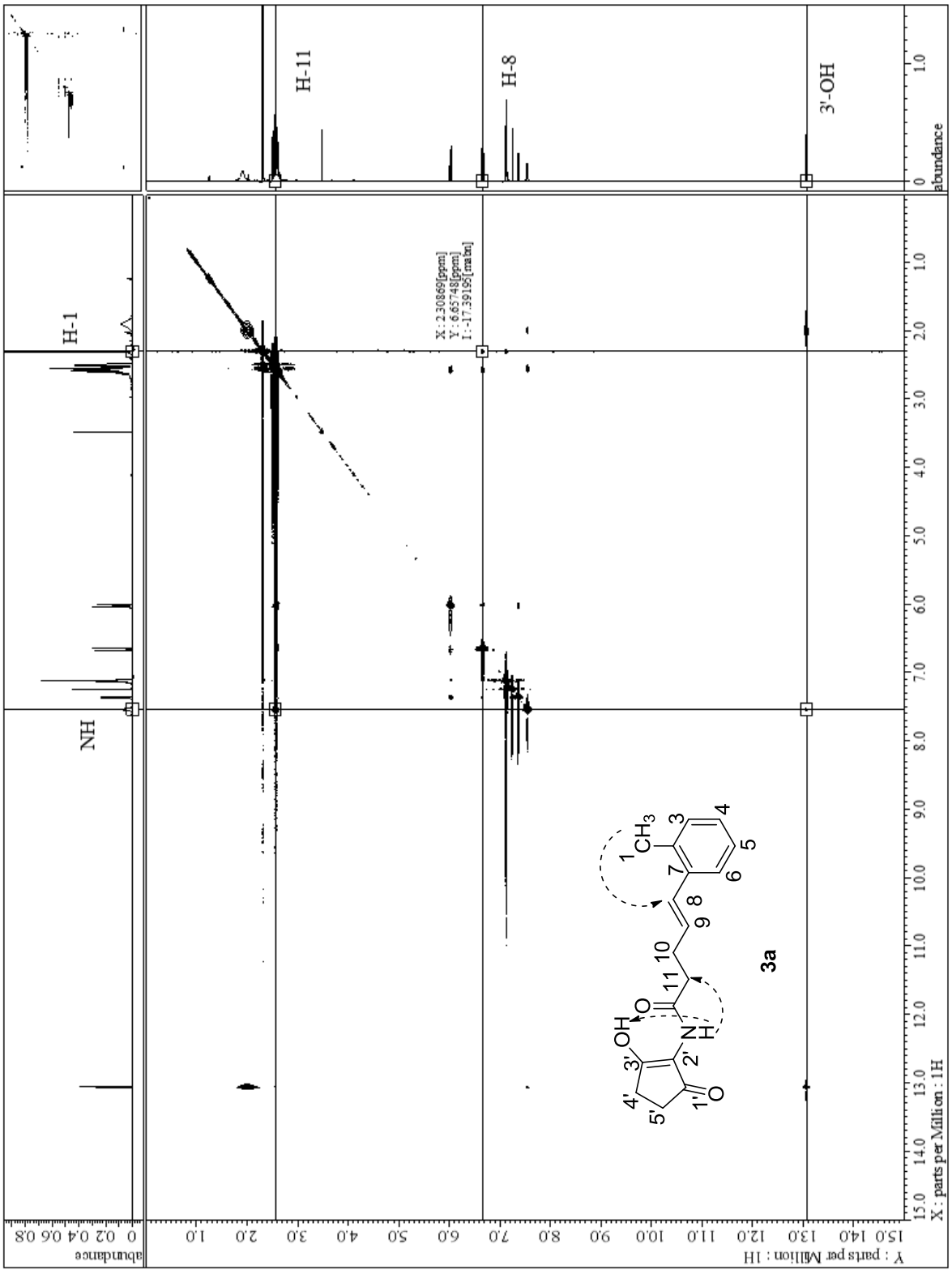


Figure S12. NOESY ($CDCl_3$) spectrum of **3a**.

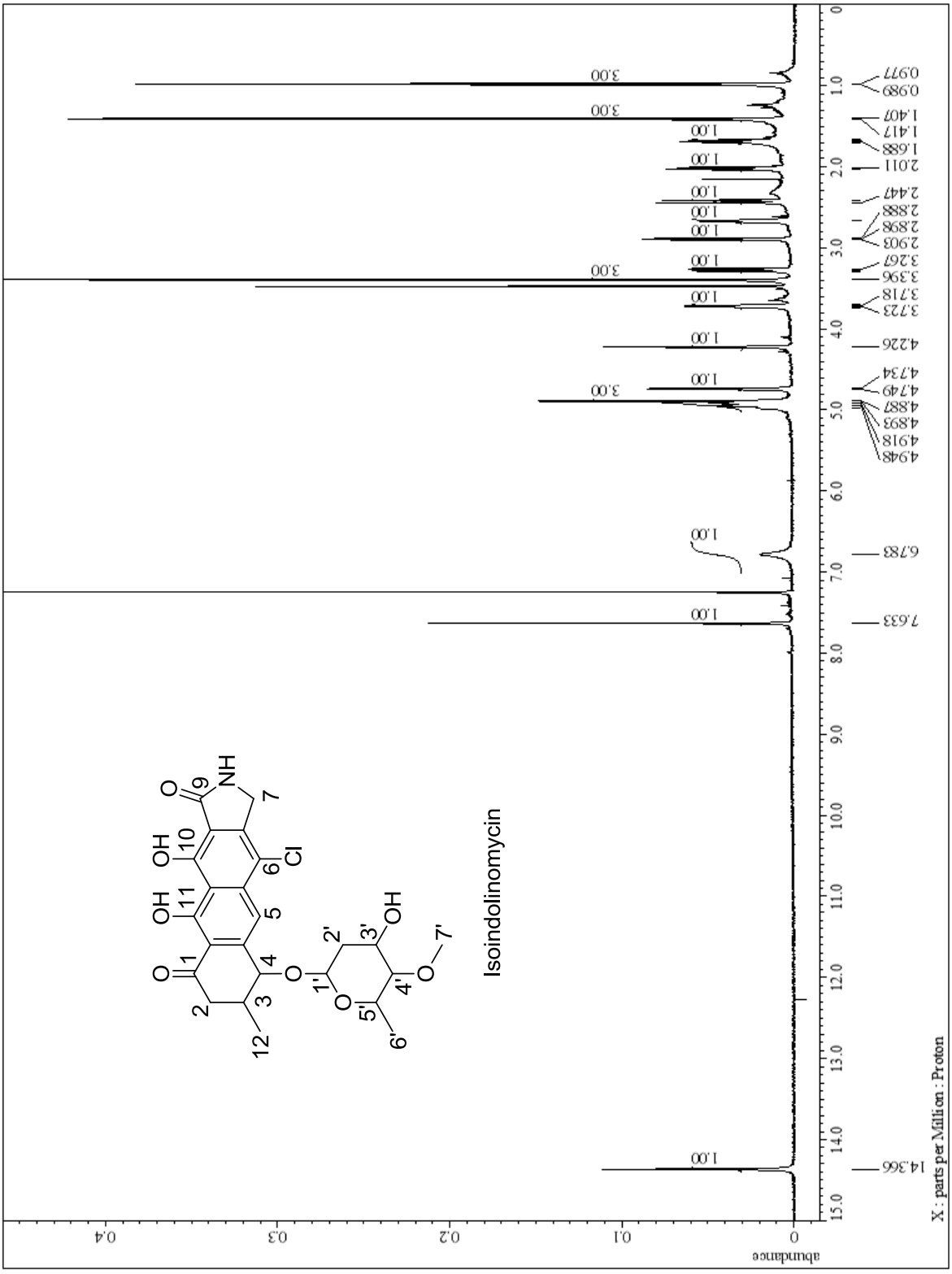


Figure S13. ^{1}H NMR (600 MHz, $CDCl_3$) spectrum of isoindolinomycin.

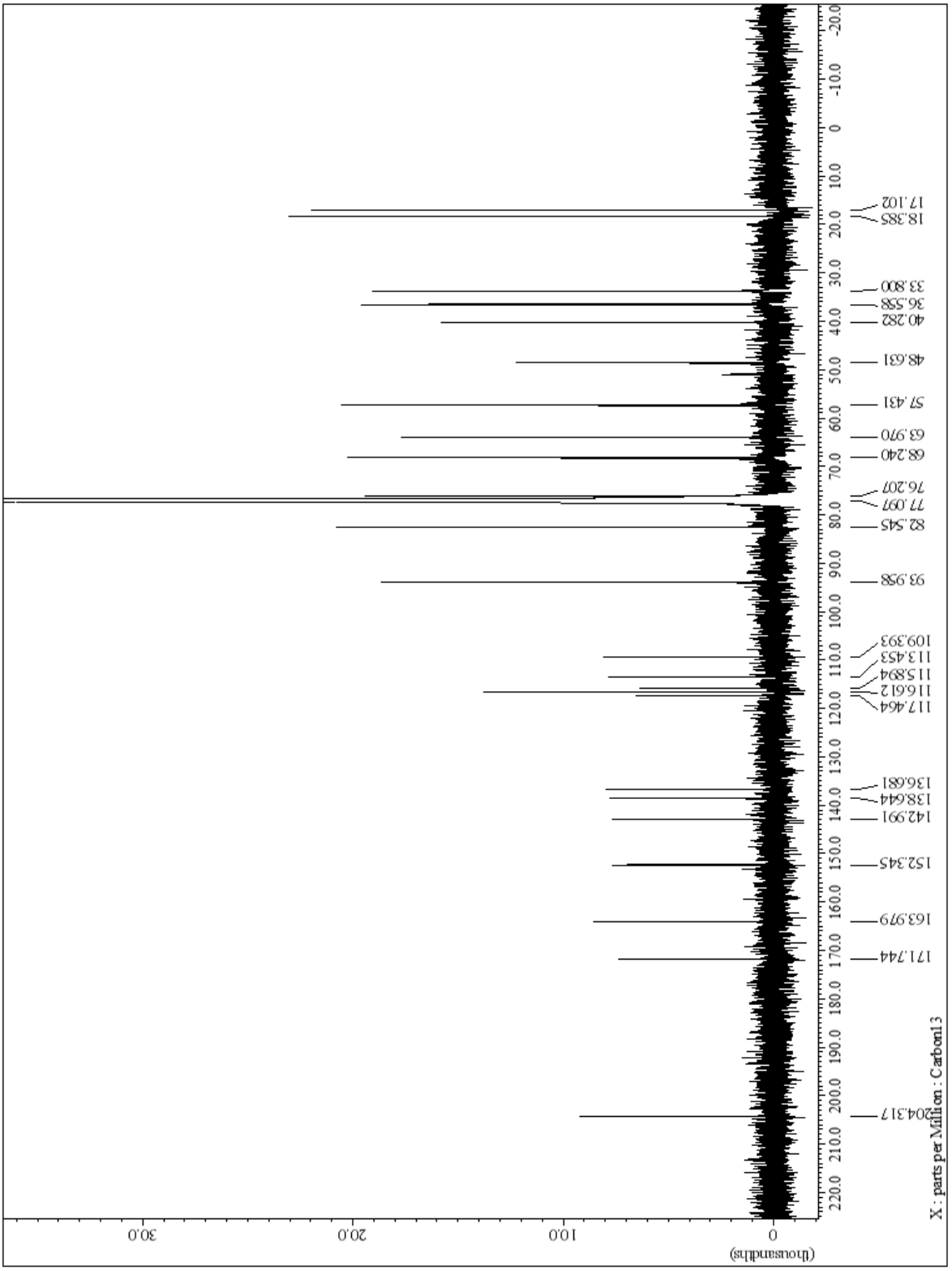


Figure S14. ^{13}C NMR (150 MHz, $CDCl_3$) spectrum of isoindolinomycin.

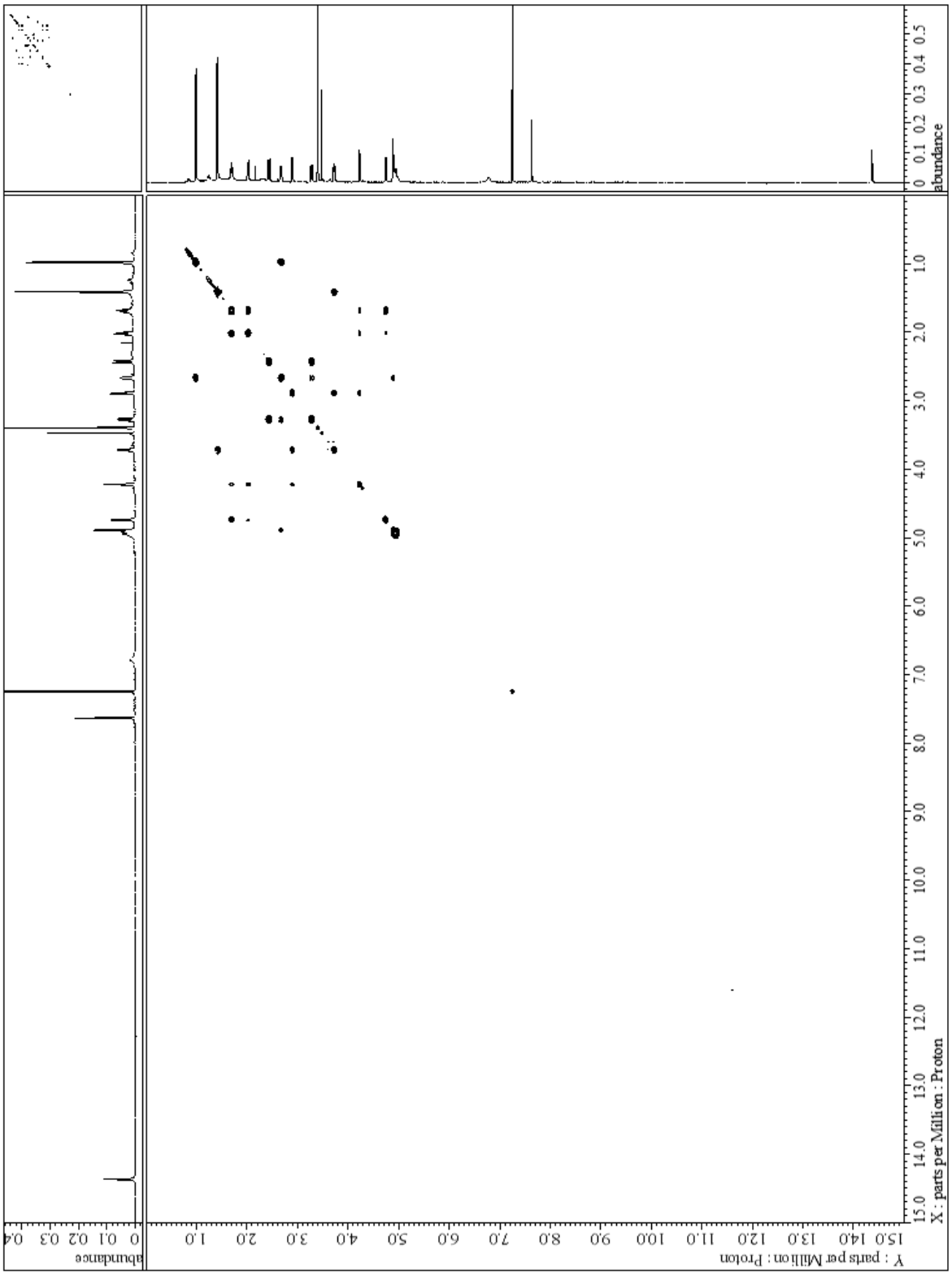


Figure S15. ^{1}H-^{1}H COSY ($CDCl_3$) spectrum of isoindolinomycin.

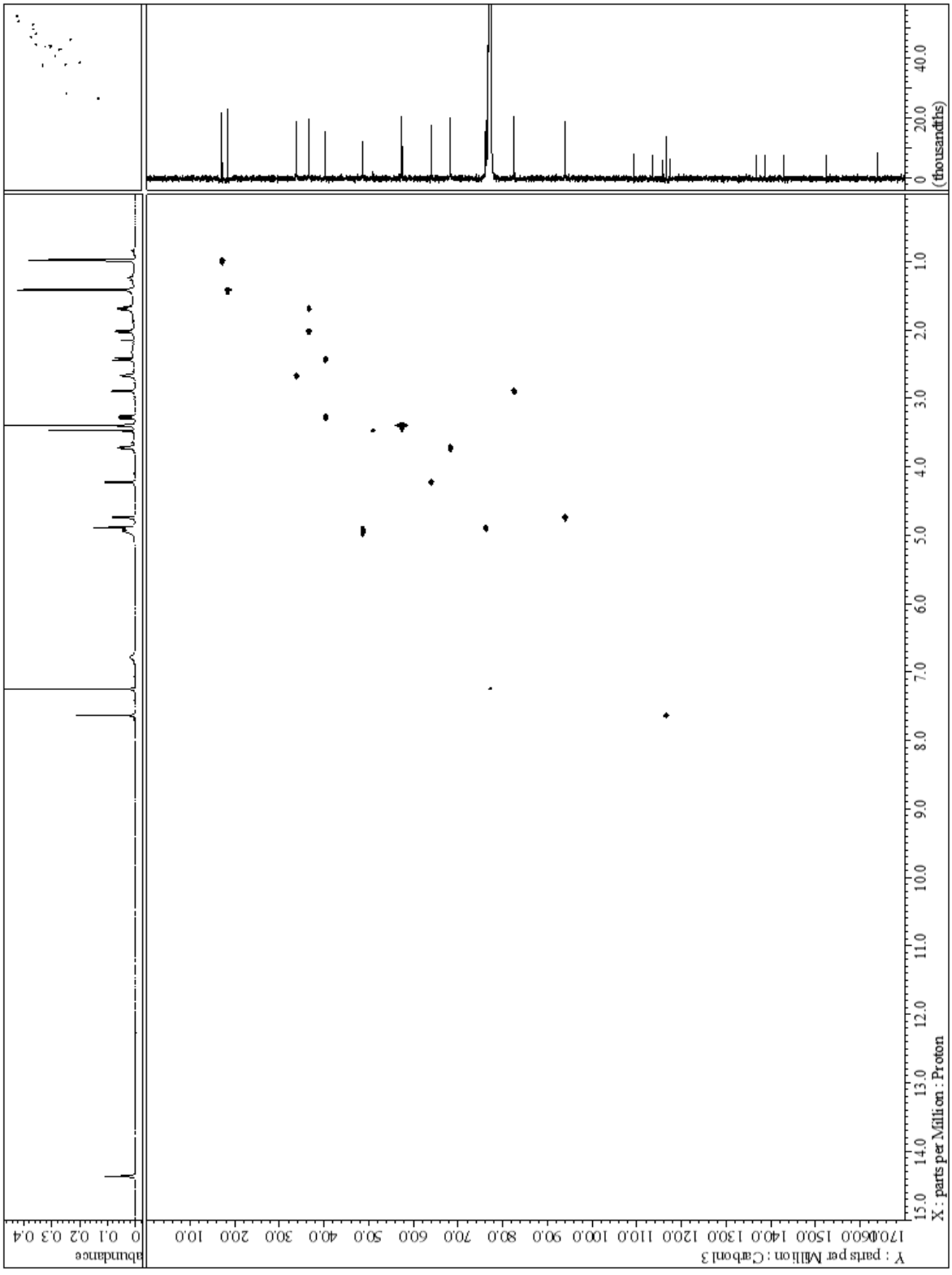


Figure S16. HSQC ($CDCl_3$) spectrum of isoindolinomycin.

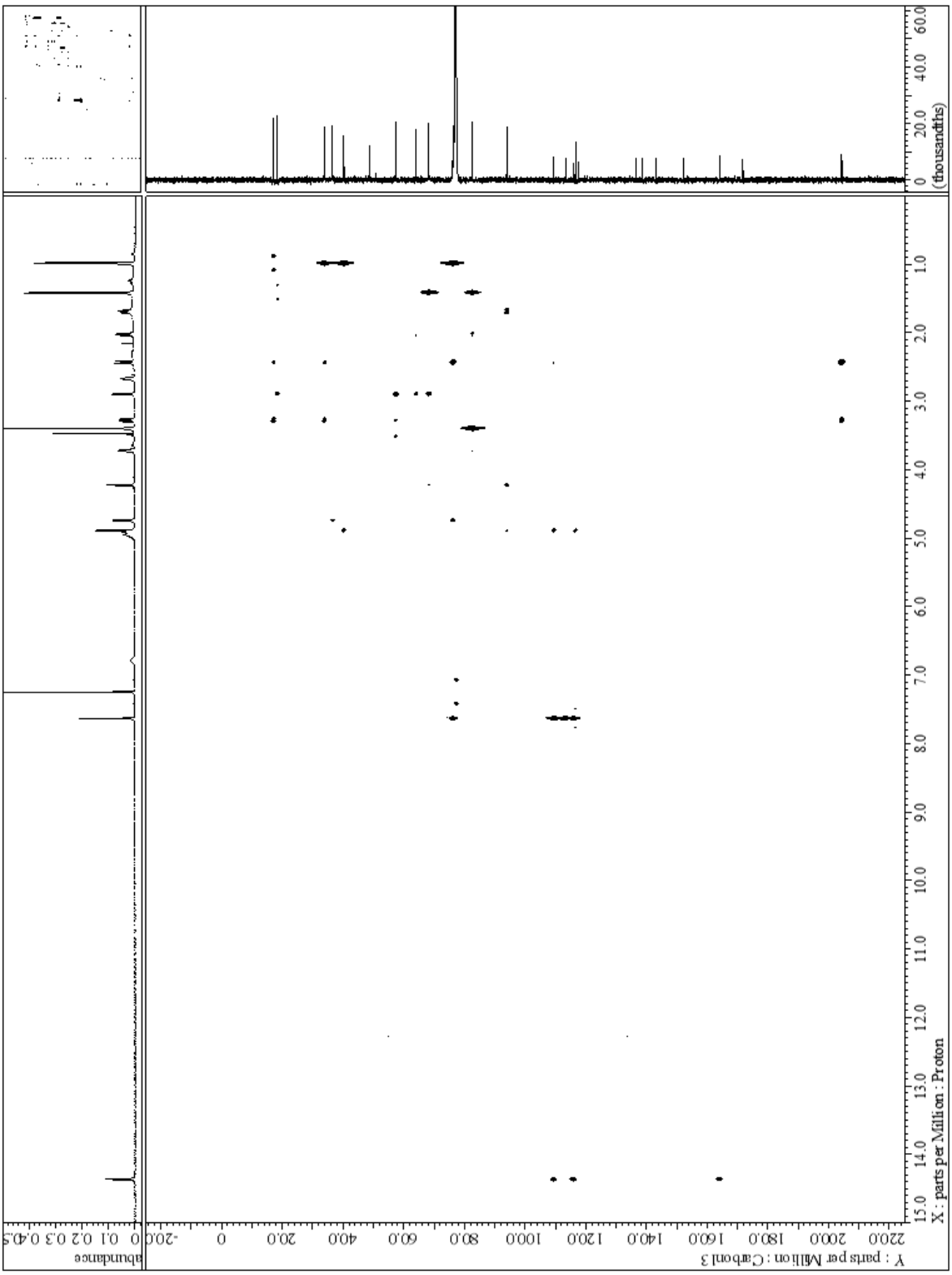


Figure S17. HMBC ($CDCl_3$) spectrum of isoindolinomycin.

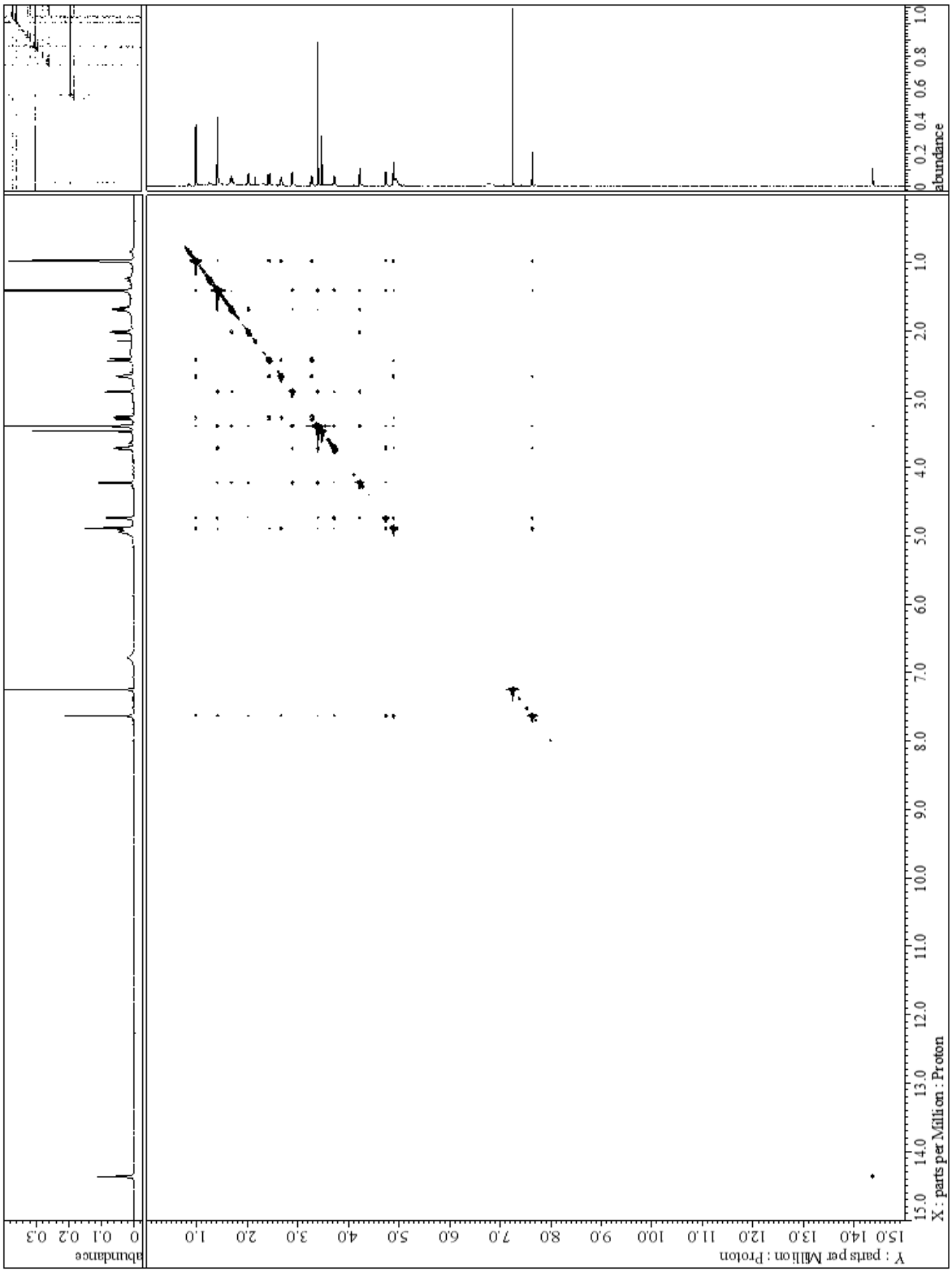


Figure S18. NOESY ($CDCl_3$) spectrum of isoindolinomycin.

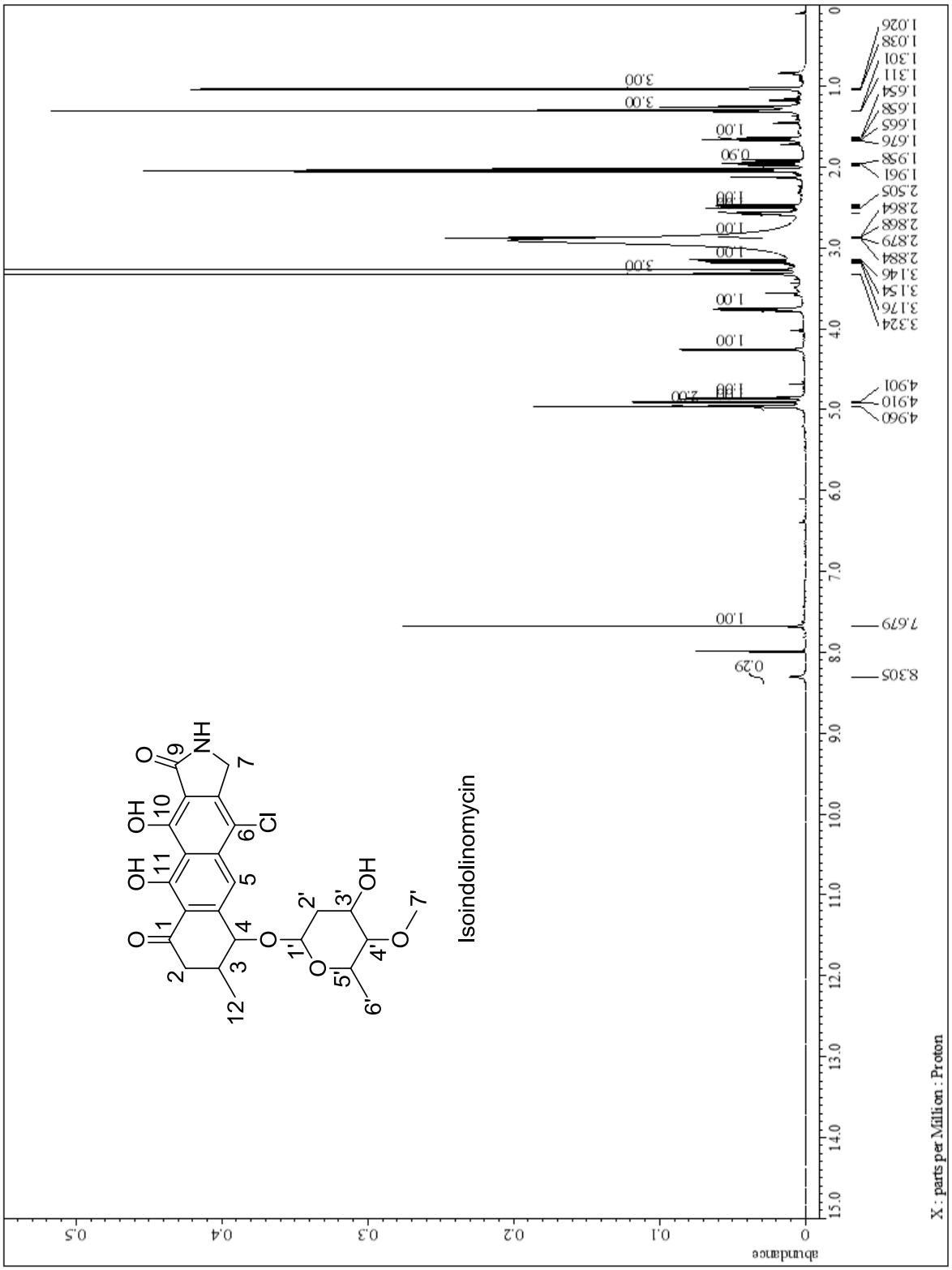


Figure S19. ^{1}H NMR (600 MHz, $(CD_3)_2CO$) spectrum of isoindolinomycin.

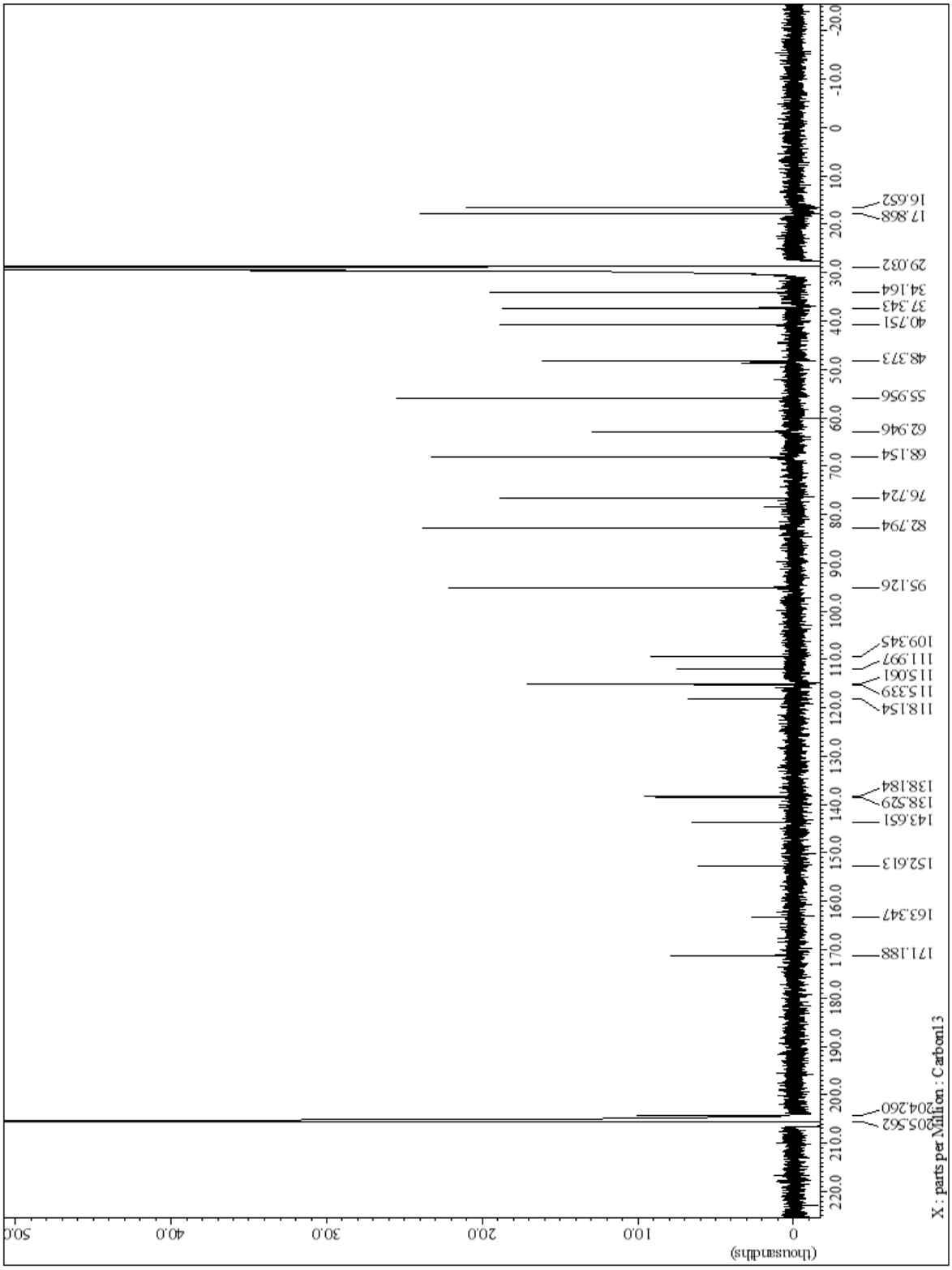


Figure S20. ^{13}C NMR (150 MHz, $(CD_3)_2CO$) spectrum of isoindolinomycin.

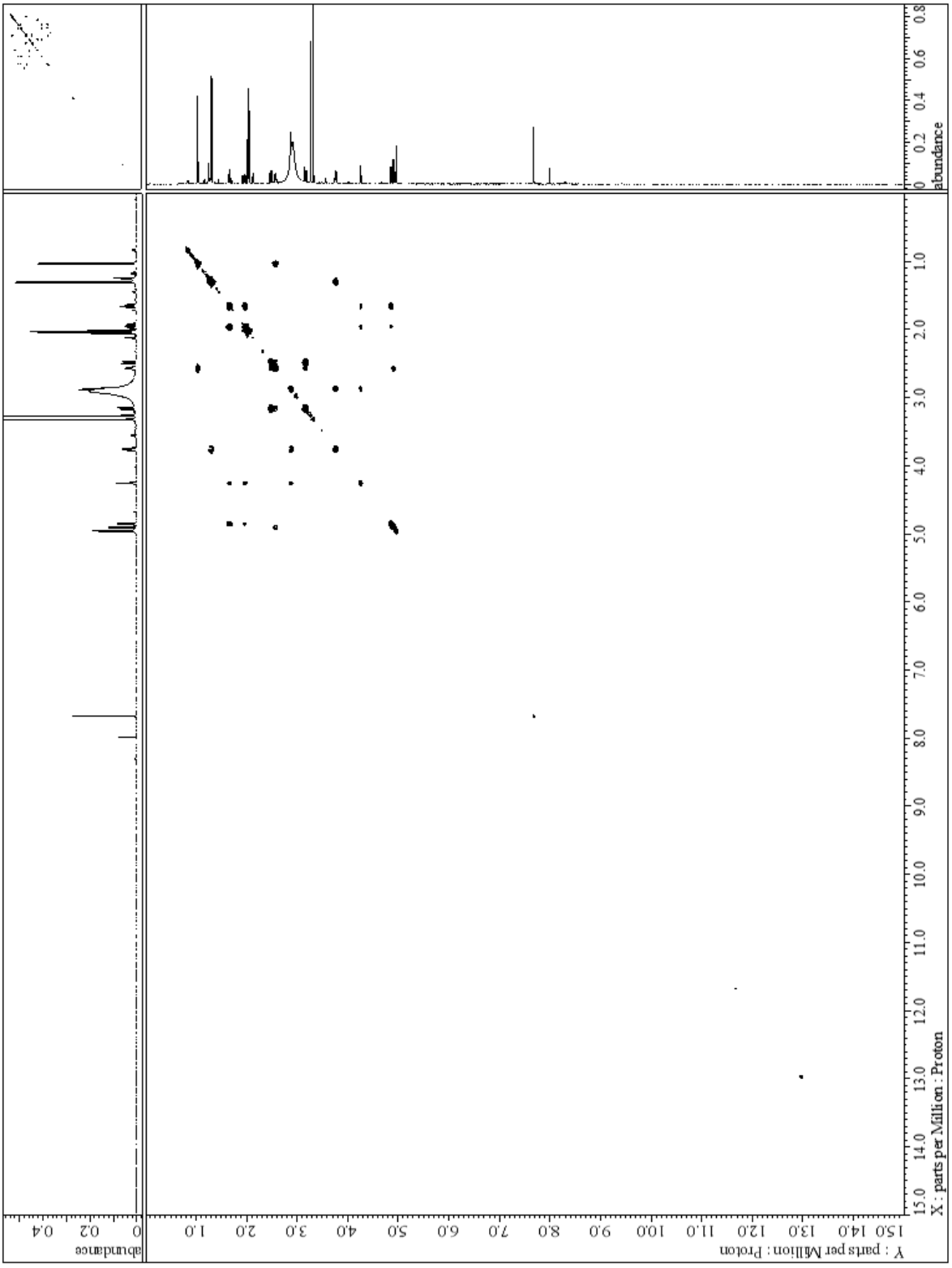


Figure S21. ^{1}H-^{1}H COSY ($(CD_3)_2CO$) spectrum of isoindolinomycin.

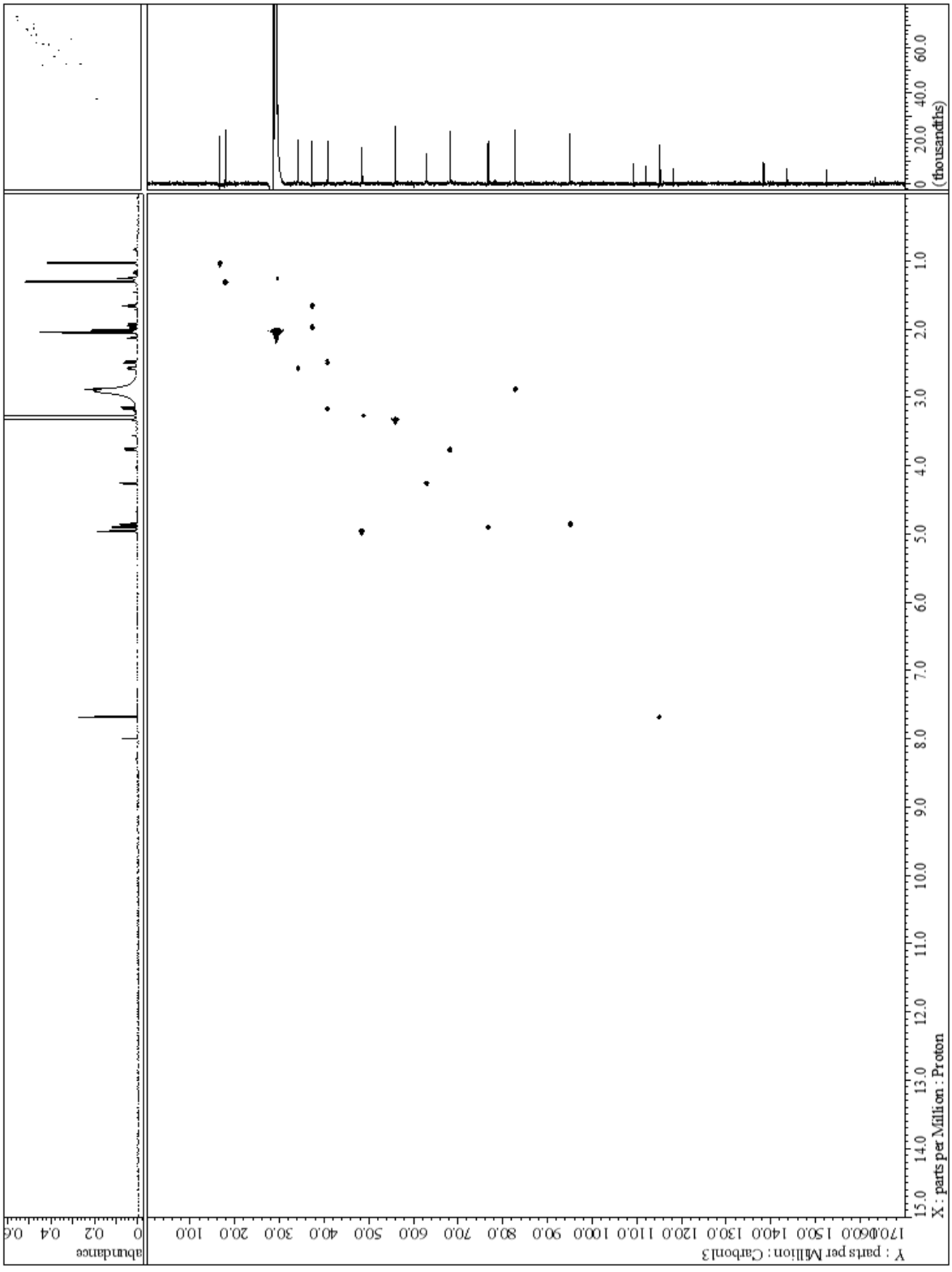


Figure S22. HSQC ($(CD_3)_2CO$) spectrum of isoindolinomycin.

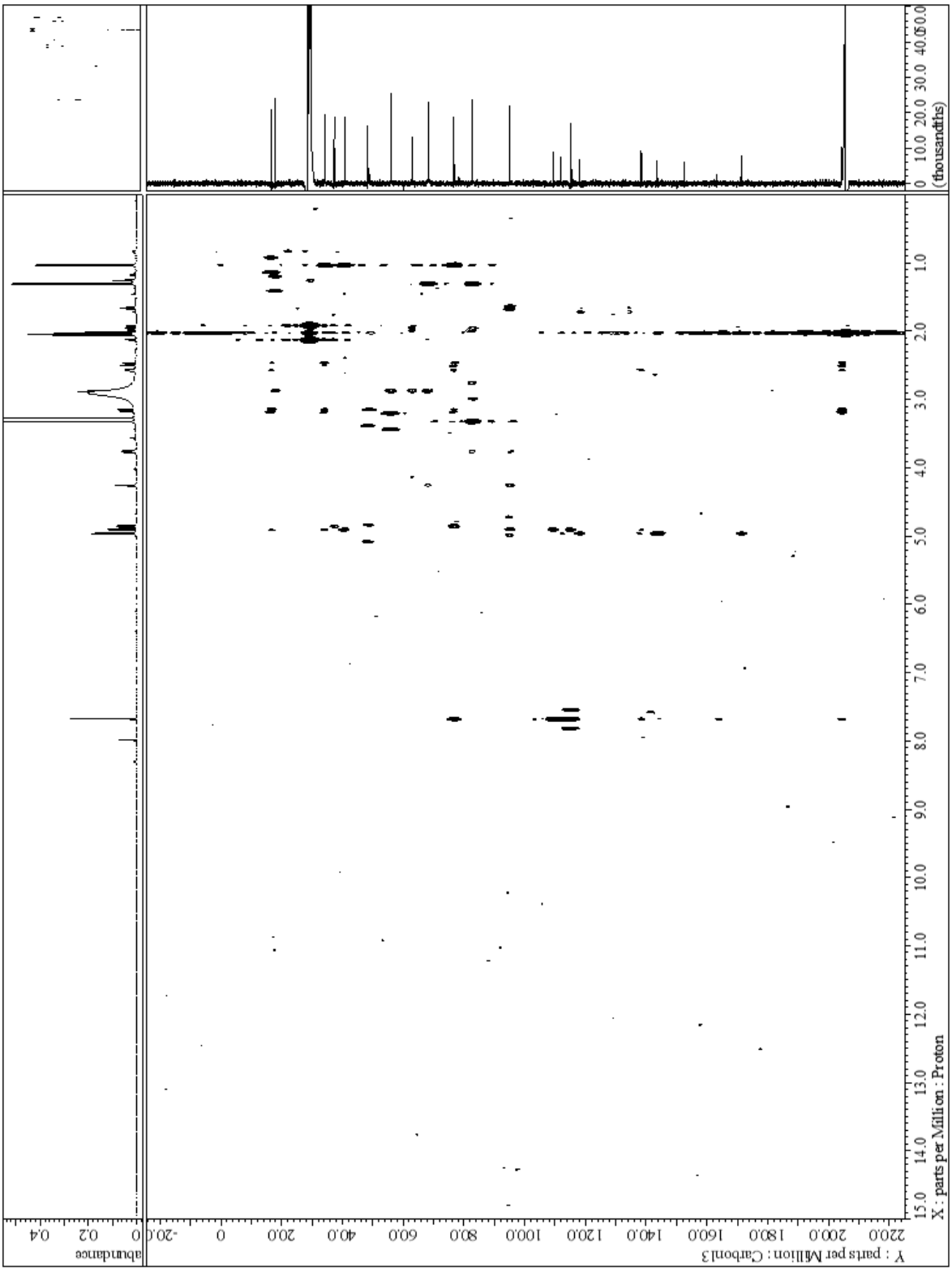


Figure S23. HMBC ($(CD_3)_2CO$) spectrum of isoindolinomycin.

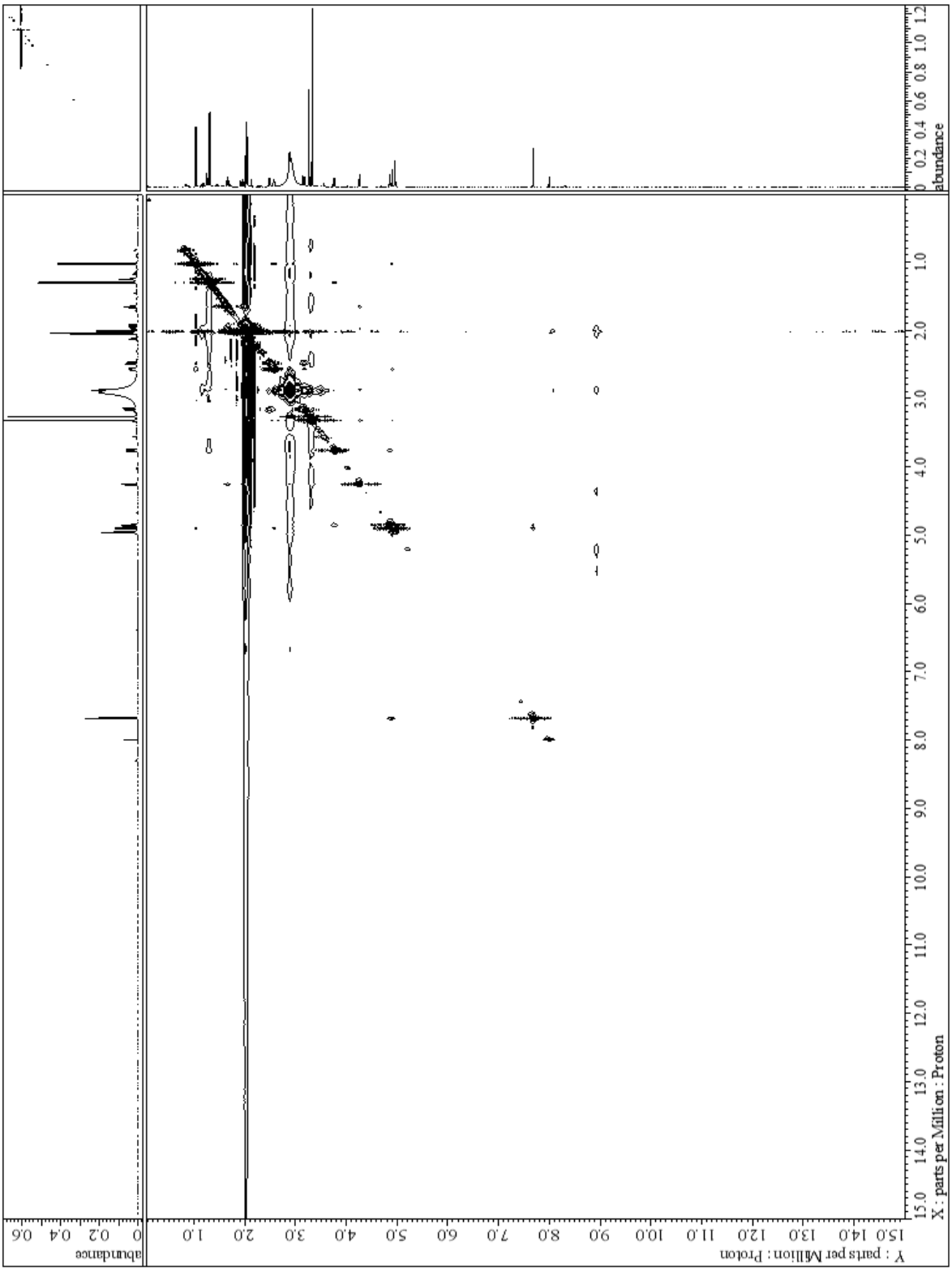


Figure S24. NOESY ($(CD_3)_2CO$) spectrum of isoindolinomycin.

REFERENCES

1. Berdy, J. (2005) Bioactive microbial metabolites. A personal view, *Journal of Antibiotics* **58**, 1-26.
2. Newman, D. J., and Cragg, G. M. (2012) Natural Products As Sources of New Drugs over the 30 Years from 1981 to 2010, *Journal of Natural Products* **75**, 311-335.
3. Berdy, J. (2012) Thoughts and facts about antibiotics: Where we are now and where we are heading, *Journal of Antibiotics* **65**, 385-395.
4. Ikeda, H., Ishikawa, J., Hanamoto, A., Shinose, M., Kikuchi, H., Shiba, T., Sakaki, Y., Hattori, M., and Omura, S. (2003) Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis*, *Nature Biotechnology* **21**, 526-531.
5. Ohnishi, Y., Ishikawa, J., Hara, H., Suzuki, H., Ikenoya, M., Ikeda, H., Yamashita, A., Hattori, M., and Horinouchi, S. (2008) Genome sequence of the streptomycin-producing microorganism *Streptomyces griseus* IFO 13350, *Journal of Bacteriology* **190**, 4050-4060.
6. Bentley, S. D., Chater, K. F., Cerdeno-Tarraga, A. M., Challis, G. L., Thomson, N. R., James, K. D., Harris, D. E., Quail, M. A., Kieser, H., Harper, D., Bateman, A., Brown, S., Chandra, G., Chen, C. W., Collins, M., Cronin, A., Fraser, A., Goble, A., Hidalgo, J., Hornsby, T., Howarth, S., Huang, C. H., Kieser, T., Larke, L., Murphy, L., Oliver, K., O'Neil, S., Rabbinowitsch, E., Rajandream, M. A., Rutherford, K., Rutter, S., Seeger, K., Saunders, D., Sharp, S., Squares, R., Squares, S., Taylor, K., Warren, T., Wietzorrek, A., Woodward, J., Barrell, B. G., Parkhill, J., and Hopwood, D. A. (2002) Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2), *Nature* **417**, 141-147.
7. Nett, M., Ikeda, H., and Moore, B. S. (2009) Genomic basis for natural product biosynthetic diversity in the actinomycetes, *Natural Product Reports* **26**, 1362-1384.
8. Bode, H. B., Bethe, B., Hofs, R., and Zeeck, A. (2002) Big effects from small changes: Possible ways to explore nature's chemical diversity, *Chembiochem* **3**, 619-627.
9. Takagi, H., Motohashi, K., Miyamoto, T., Shin-Ya, K., Furihata, K., and Seto, H. (2005) Studies on terpenoids produced by actinomycetes - Isolation and structural elucidation of antioxidative agents, naphterpins B and C, *Journal of Antibiotics* **58**, 275-278.

10. Imai, S., Furihata, K., Hayakawa, Y., Noguchi, T., and Seto, H. (1989) Lavanducyanin, a new antitumor substance produced by *Streptomyces* sp., *Journal of Antibiotics* **42**, 1196-1198.
11. Takatsuk.A, Arima, K., and Tamura, G. (1971) Tunicamycin, a new antibiotic. 1. Isolation and characterization of Tunicamycin, *Journal of Antibiotics* **24**, 215-223.
12. Murakami, R., Fujita, Y., Kizuka, M., Kagawa, T., Muramatsu, Y., Miyakoshi, S., Takatsu, T., and Inukai, M. (2008) A-94964, a novel inhibitor of bacterial translocase I, produced by *Streptomyces* sp SANK 60404, *Journal of Antibiotics* **61**, 537-544.
13. Omura, S., Eda, S., Funayama, S., Komiyama, K., Takahashi, Y., and Woodruff, H. B. (1989) Studies on a novel antitumor antibiotic, phenazinomycin - Taxonomy, fermentation, isolation and physicochemical and biological characteristics, *Journal of Antibiotics* **42**, 1037-1042
14. Ueda, J.-y., Chijiwa, S., Takagi, M., and Shin-ya, K. (2008) A Novel Versipelostatin Analogue, Versipelostatin F Isolated from *Streptomyces versipellis* 4083-SVS6, *Journal of Antibiotics* **61**, 752-755.
15. Takahashi, E., Kimura, T., Nakamura, K., Arahira, M., and Iida, M. (1995) Phosphonothrixin, a novel herbicidal antibiotic produced by *Saccharothrix* sp. ST-888. 1. Taxonomy, fermentation, isolation and biological properties, *Journal of Antibiotics* **48**, 1124-1129.
16. Hosoya, T., Hirokawa, T., Takagi, M., and Shin-ya, K. (2012) Trichostatin Analogues JBIR-109, JBIR-110, and JBIR-111 from the Marine Sponge-Derived *Streptomyces* sp RM72, *Journal of Natural Products* **75**, 285-289.
17. Shirai, M., Okuda, M., Motohashi, K., Imoto, M., Furihata, K., Matsuo, Y., Katsuta, A., Shizuri, Y., and Seto, H. (2010) Terpenoids produced by actinomycetes: isolation, structural elucidation and biosynthesis of new diterpenes, gifhornenolones A and B from *Verrucosispora gifhornensis* YM28-088, *Journal of Antibiotics* **63**, 245-250
18. Kurosawa, K., Takahashi, K., and Tsuda, E. (2001) SNF4435C and D, novel immunosuppressants produced by a strain of *Streptomyces spectabilis* - I. Taxonomy, fermentation, isolation and biological activities, *Journal of Antibiotics* **54**, 541-547.
19. Omura, S., Otoguro, K., Nishikiori, T., Oiwa, R., and Iwai, Y. (1981) Setamycin, a new antibiotic, *Journal of Antibiotics* **34**, 1253-1256.

20. Shinya, K., Shimizu, S., Kunigami, T., Furihata, K., and Seto, H. (1995) A new neuronal cell protecting substance, Lavanduquinocin, produced by *Streptomyces viridochromogenes*, *Journal of Antibiotics* **48**, 574-578.
21. Firn, R. D., and Jones, C. G. (2000) The evolution of secondary metabolism - a unifying model, *Molecular Microbiology* **37**, 989-994.
22. Milshteyn, A., Schneider, J. S., and Brady, S. F. (2014) Mining the Metabiome: Identifying Novel Natural Products from Microbial Communities, *Chemistry & Biology* **21**, 1211-1223.
23. Zarins-Tutt, J. S., Barberi, T. T., Gao, H., Mearns-Spragg, A., Zhang, L., Newman, D. J., and Goss, R. J. (2015) Prospecting for new bacterial metabolites: a glossary of approaches for inducing, activating and upregulating the biosynthesis of bacterial cryptic or silent natural products, *Natural Product Reports* **33**, 54-72.
24. Nguyen, D. D., Wu, C. H., Moree, W. J., Lamsa, A., Medema, M. H., Zhao, X., Gavilan, R. G., Aparicio, M., Atencio, L., Jackson, C., Ballesteros, J., Sanchez, J., Watrous, J. D., Phelan, V. V., van de Wiel, C., Kersten, R. D., Mehnaz, S., De Mot, R., Shank, E. A., Charusanti, P., Nagarajan, H., Duggan, B. M., Moore, B. S., Bandeira, N., Palsson, B., Pogliano, K., Gutiérrez, M., and Dorrestein, P. C. (2013) MS/MS networking guided analysis of molecule and gene cluster families, *Proceedings of the National Academy of Sciences of the United States of America* **110**, 2611-2620.
25. Liu, W. T., Lamsa, A., Wong, W. R., Boudreau, P. D., Kersten, R., Peng, Y., Moree, W. J., Duggan, B. M., Moore, B. S., Gerwick, W. H., Linington, R. G., Pogliano, K., and Dorrestein, P. C. (2014) MS/MS-based networking and peptidogenomics guided genome mining revealed the stenothricin gene cluster in *Streptomyces roseosporus*, *Journal of Antibiotics* **67**, 99-104.
26. Wang, L., Hu, Y., Zhang, Y., Wang, S., Cui, Z., Bao, Y., Jiang, W., and Hong, B. (2009) Role of sgcR3 in positive regulation of enediyne antibiotic C-1027 production of *Streptomyces globisporus* C-1027, *BMC Microbiology* **9**, 1-12.
27. Chen, Y., Yin, M., Horsman, G. P., Huang, S., and Shen, B. (2010) Manipulation of pathway regulation in *Streptomyces globisporus* for overproduction of the enediyne antitumor antibiotic C-1027, *Journal of Antibiotics* **63**, 482-485.

28. Komatsu, M., Uchiyama, T., Omura, S., Cane, D. E., and Ikeda, H. (2010) Genome-minimized *Streptomyces* host for the heterologous expression of secondary metabolism, *Proceedings of the National Academy of Sciences of the United States of America* **107**, 2646-2651.
29. Luo, Y., Huang, H., Liang, J., Wang, M., Lu, L., Shao, Z., Cobb, R. E., and Zhao, H. (2013) Activation and characterization of a cryptic polycyclic tetramate macrolactam biosynthetic gene cluster, *Nature Communication* **4**, 1-8.
30. Fdhila, F., Vázquez, V., Sánchez, J. L., and Riguera, R. (2003) dd-diketopiperazines: antibiotics active against *Vibrio anguillarum* isolated from marine bacteria associated with cultures of Pecten maximus, *Journal of Natural Products* **66**, 1299-1301.
31. Cueto, M., Jensen, P. R., Kauffman, C., Fenical, W., Lobkovsky, E., and Clardy, J. (2001) Pestalone, a new antibiotic produced by a marine fungus in response to bacterial challenge, *Journal of Natural Products* **64**, 1444-1446.
32. Onaka, H., Mori, Y., Igarashi, Y., and Furumai, T. (2011) Mycolic acid-containing bacteria induce natural-product biosynthesis in *Streptomyces* species, *Applied and Environmental Microbiology* **77**, 400-406.
33. Hu, H. F., Zhang, Q., and Ochi, K. (2002) Activation of antibiotic biosynthesis by specified mutations in the rpoB gene (encoding the RNA polymerase beta subunit) of *Streptomyces lividans*, *Journal of Bacteriology* **184**, 3984-3991.
34. Goldstein, B. P. (2014) Resistance to rifampicin: a review, *Journal of Antibiotics* **67**, 625-630.
35. Jin, D. J., and Gross, C. A. (1989) Characterization of the pleiotropic phenotypes of rifampin-resistant *rpoB* mutants of *Escherichia coli*, *Journal of Bacteriology* **171**, 5229-5231.
36. Sonenshein, A. L., Cami, B., Brevet, J., and Cote, R. (1974) Isolation and characterization of rifampin-resistant and streptolydigin-resistant mutants of *Bacillus subtilis* with altered sporulation properties, *Journal of Bacteriology* **120**, 253-265.
37. Perkins, A. E., and Nicholson, W. L. (2008) Uncovering new metabolic capabilities of *Bacillus subtilis* using phenotype profiling of rifampin-resistant *rpoB* mutants, *Journal of Bacteriology* **190**, 807-814.

38. Tanaka, Y., Kasahara, K., Hirose, Y., Murakami, K., Kugimiya, R., and Ochi, K. (2013) Activation and products of the cryptic secondary metabolite biosynthetic gene clusters by rifampin resistance (*rpoB*) mutations in actinomycetes, *Journal of Bacteriology* **195**, 2959-2970.
39. Ochi, K., Okamoto, S., Tozawa, Y., Inaoka, T., Hosaka, T., Xu, J., and Kurosawa, K. (2004) Ribosome engineering and secondary metabolite production, *Advances in Applied Microbiology* **56**, 155-184.
40. Chakraburtty, R., and Bibb, M. (1997) The ppGpp synthetase gene (*relA*) of *Streptomyces coelicolor* A3(2) play's a conditional role in antibiotic production and morphological differentiation, *Journal of Bacteriology* **179**, 5854-5861.
41. Hosaka, T., Ohnishi-Kameyama, M., Muramatsu, H., Murakami, K., Tsurumi, Y., Kodani, S., Yoshida, M., Fujie, A., and Ochi, K. (2009) Antibacterial discovery in actinomycetes strains with mutations in RNA polymerase or ribosomal protein S12, *Nature Biotechnology* **27**, 462-464.
42. Zhang, Y., Huang, H., Xu, S., Wang, B., Ju, J., Tan, H., and Li, W. (2015) Activation and enhancement of Fredericamycin A production in deepsea-derived *Streptomyces somaliensis* SCSIO ZH66 by using ribosome engineering and response surface methodology, *Microbial Cell Factories* **14**, 1-11.
43. Xu, J., Tozawa, Y., Lai, C., Hayashi, H., and Ochi, K. (2002) A rifampicin resistance mutation in the *rpoB* gene confers ppGpp-independent antibiotic production in *Streptomyces coelicolor* A3(2), *Molecular Genetics and Genomics* **268**, 179-189.
44. Shinya, K., Imai, S., Furihata, K., Hayakawa, Y., Kato, Y., Vanduyne, G. D., Clardy, J., and Seto, H. (1990) Isolation and structural elucidation of an antioxidative agent, naphterpin, *Journal of Antibiotics* **43**, 444-447.
45. Calvori, C., Frontali, L., Leoni, L., and Tecce, G. (1965) Effect of Rifamycin on protein synthesis, *Nature* **207**, 417-418.
46. Campbell, E. A., Korzheva, N., Mustaev, A., Murakami, K., Nair, S., Goldfarb, A., and Darst, S. A. (2001) Structural mechanism for rifampicin inhibition of bacterial RNA polymerase, *Cell* **104**, 901-912.
47. Thong, W. L. (2013) Studies on secondary metabolite production by actinomycetes with rifampicin-induced *rpoB* gene mutation, In *Department of Biotechnology*, The University of Tokyo, Tokyo, Japan.

48. Kinjo, J., Yokomizo, K., Awata, Y., Shibata, M., Nohara, T., Teramine, T., and Takahashi, K. (1987) Structures of phytotoxins, AV-toxins C, D and E, produced by zonate leaf spot fungus of mulberry, *Tetrahedron Letters* **28**, 3697-3698.

49. Gerber, N. N. (1967) Phenazines, phenoxazinones and dioxopiperazines from *Streptomyces thioluteus*, *Journal of Organic Chemistry* **32**, 4055-4056.

50. Liu, W. H., Li, M. G., Li, Y. Q., Zhao, J. Y., Ding, Z. G., Yang, P. W., Cui, X. L., and Wen, M. L. (2008) Cytotoxic metabolites of *Streptimonospora salina*, *Chemistry of Natural Compounds* **44**, 503-505.

51. Giurg, M., Piekielska, K., Gebala, M., Ditkowski, B., Wolanski, M., Peczynska-Czoch, W., and Mlochowski, J. (2007) Catalytic oxidative cyclocondensation of o-aminophenols to 2-amino-3H-phenoxazin-3-ones, *Synthetic Communications* **37**, 1779-1789.

52. Kakinuma, K., Ikekawa, N., Nakagawa, A., and Omura, S. (1979) The structure of asukamycin, a possible shunt metabolite from 3-dehydroquinic acid in the shikimate pathway, *Journal of the American Chemical Society* **101**, 3402-3404.

53. Zeeck, A., Schroder, K., Frobel, K., Grote, R., and Thiericke, R. (1987) The structure of manumycin I. Characterization, structure elucidation and biological-activity, *Journal of Antibiotics* **40**, 1530-1540.

54. V., M. V. J. R., P., M. R., Peter, M., Iris, G.-W., and Hartmut, L. (2002) 5-(2-Methylphenyl)-4-pentenoic Acid from a Terrestrial Streptomycete, *Z. Naturforschung* **57b**, 335-337.

55. Nakata, M., Saito, M., Inouye, Y., Nakamura, S., Hayakawa, Y., and Seto, H. (1992) A new anthracycline antibiotic, cinerubin R. Taxonomy, structural elucidation and biological activity, *Journal of Antibiotics* **45**, 1599-1608.

56. Johdo, O., Watanabe, Y., Ishikura, T., Yoshimoto, A., Naganawa, H., Sawa, T., and Takeuchi, T. (1991) Anthracycline metabolites from *Streptomyces violaceus* A262. II. New anthracycline epelmycins produced by a blocked mutant strain SU2-730, *Journal of Antibiotics* **44**, 1121-1129.

57. Menéndez, N., Nur-e-Alam, M., Braña, A. F., Rohr, J., Salas, J. A., and Méndez, C. (2004) Biosynthesis of the antitumor chromomycin A3 in Streptomyces griseus: analysis of the gene cluster and rational design of novel chromomycin analogs, *Chemistry & Biology* **11**, 21-32.

58. Matsumoto, N., Tsuchida, T., Nakamura, H., Sawa, R., Takahashi, Y., Naganawa, H., Iinuma, H., Sawa, T., Takeuchi, T., and Shiro, M. (1999) Lactonamycin, a new antimicrobial antibiotic produced by *Streptomyces rishiriensis* MJ773-88K4. II. Structure determination, *Journal of Antibiotics* **52**, 276-280.
59. Izumikawa, M., Kozone, I., Hashimoto, J., Kagaya, N., Takagi, M., Koiwai, H., Komatsu, M., Fujie, M., Satoh, N., Ikeda, H., and Shin-Ya, K. (2015) Novel thioviridamide derivative-JBIR-140: heterologous expression of the gene cluster for thioviridamide biosynthesis, *Journal of Antibiotics* **68**, 533-536.
60. Fuoco, D. (2012) Classification Framework and Chemical Biology of Tetracycline-Structure-Based Drugs, *Antibiotics* **1**, 1-13.
61. Suzuki, H., Furusho, Y., Higashi, T., Ohnishi, Y., and Horinouchi, S. (2006) A novel o-aminophenol oxidase responsible for formation of the phenoxazinone chromophore of grixazone, *Journal of Biological Chemistry* **281**, 824-833.
62. Zhang, W. J., Bolla, M. L., Kahne, D., and Walsh, C. T. (2010) A Three Enzyme Pathway for 2-Amino-3-hydroxycyclopent-2-enone Formation and Incorporation in Natural Product Biosynthesis, *Journal of the American Chemical Society* **132**, 6402-6411.
63. Fujii, I., and Ebizuka, Y. (1997) Anthracycline Biosynthesis in *Streptomyces galilaeus*, *Chemical Reviews* **97**, 2511-2524.
64. Shen, B. (2003) Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms, *Current Opinion in Chemical Biology* **7**, 285-295.
65. Meguro, A., Tomita, T., Nishiyama, M., and Kuzuyama, T. (2013) Identification and characterization of bacterial diterpene cyclases that synthesize the cembrane skeleton, *Chembiochem* **14**, 316-321.
66. Zerbino, D. R., and Birney, E. (2008) Velvet: algorithms for de novo short read assembly using de Bruijn graphs, *Genome Research* **18**, 821-829.
67. Ishikawa, J., and Hotta, K. (1999) FramePlot: a new implementation of the frame analysis for predicting protein-coding regions in bacterial DNA with a high G + C content, *FEMS Microbiology Letters* **174**, 251-253.
68. Hamano, Y., Maruyama, C., and Kimoto, H. (2006) Construction of a knockout mutant of the streptothricin-resistance gene in *Streptomyces albulus* by electroporation., *Actinomycetologica* **20**, 35-41.

69. Staunton, J., and Weissman, K. J. (2001) Polyketide biosynthesis: a millennium review, *Natural Product Reports* **18**, 380-416.
70. Moore, B. S., and Hertweck, C. (2002) Biosynthesis and attachment of novel bacterial polyketide synthase starter units, *Natural Product Reports* **19**, 70-99.
71. Kalan, L., Gessner, A., Thaker, M. N., Waglechner, N., Zhu, X., Szawiola, A., Bechthold, A., Wright, G. D., and Zechel, D. L. (2013) A cryptic polyene biosynthetic gene cluster in *Streptomyces calvus* is expressed upon complementation with a functional bldA gene, *Chemical & Biology* **20**, 1214-1224.
72. Thiericke, R., Zeeck, A., Nakagawa, A., Omura, S., Herrold, R. E., Wu, S. T. S., Beale, J. M., and Floss, H. G. (1990) Biosynthesis of the manumycin group antibiotics, *Journal of the American Chemical Society* **112**, 3979-3987.
73. Thiericke, R., Zeeck, A., Robinson, J. A., Beale, J. M., and Floss, H. G. (1989) The biosynthesis of manumycin; origin of the oxygen and nitrogen atoms, *Journal of the Chemical Society, Chemical Communications*, **7**, 402-403.
74. Hertweck, C. (2009) The biosynthetic logic of polyketide diversity, *Angewandte Chemie International Edition* **48**, 4688-4716.
75. Hernandez, D., François, P., Farinelli, L., Osterås, M., and Schrenzel, J. (2008) *De novo* bacterial genome sequencing: millions of very short reads assembled on a desktop computer, *Genome Research* **18**, 802-809.
76. Gago, G., Diacovich, L., Arabolaza, A., Tsai, S. C., and Gramajo, H. (2011) Fatty acid biosynthesis in actinomycetes, *FEMS Microbiology Reviews* **35**, 475-497.
77. Hertweck, C., Luzhetskyy, A., Rebets, Y., and Bechthold, A. (2007) Type II polyketide synthases: gaining a deeper insight into enzymatic teamwork, *Natural Product Reports* **24**, 162-190.
78. Pohle, S., Appelt, C., Roux, M., Fiedler, H. P., and Süssmuth, R. D. (2011) Biosynthetic gene cluster of the non-ribosomally synthesized cyclodepsipeptide skyllamycin: deciphering unprecedented ways of unusual hydroxylation reactions, *Journal of the American Chemical Society* **133**, 6194-6205.
79. Beaudry, C. M., Malerich, J. P., and Trauner, D. (2005) Biosynthetic and biomimetic electrocyclizations, *Chemical Reviews* **105**, 4757-4778.
80. Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0, *Molecular Biology and Evolution* **30**, 2725-2729.

81. Bilyk, O., Brötz, E., Tokovenko, B., Bechthold, A., Paululat, T., and Luzhetskyy, A. (2015) New Simocyclinones: Surprising Evolutionary and Biosynthetic Insights, *ACS Chemical Biology* **11**, 241-250.

82. Ames, B. D., Korman, T. P., Zhang, W., Smith, P., Vu, T., Tang, Y., and Tsai, S. C. (2008) Crystal structure and functional analysis of tetracenomycin ARO/CYC: implications for cyclization specificity of aromatic polyketides, *Proceedings of the National Academy of Sciences of the United States of America* **105**, 5349-5354.

83. Harris, C. M., Kannan, R., Kopecka, H., and Harris, T. M. (1985) The role of the chlorine substituents in the antibiotic vancomycin: preparation and characterization of mono- and didechlorovancomycin, *Journal of the American Chemical Society* **107**, 6652-6658.

84. Neumann, C. S., Fujimori, D. G., and Walsh, C. T. (2008) Halogenation strategies in natural product biosynthesis, *Chemical & Biology* **15**, 99-109.

85. Wagner, C., El Omari, M., and König, G. M. (2009) Biohalogenation: nature's way to synthesize halogenated metabolites, *Journal of Natural Products* **72**, 540-553.

86. Röttig, M., Medema, M. H., Blin, K., Weber, T., Rausch, C., and Kohlbacher, O. (2011) NRPSpredictor2-a web server for predicting NRPS adenylation domain specificity, *Nucleic Acids Research* **39**, 362-367.

87. Rafanan, E. R., Hutchinson, C. R., and Shen, B. (2000) Triple hydroxylation of tetracenomycin A2 to tetracenomycin C involving two molecules of O_2 and one molecule of H_2O, *Organic Letters* **2**, 3225-3227.

88. Maruyama, C., Toyoda, J., Kato, Y., Izumikawa, M., Takagi, M., Shin-ya, K., Katano, H., Utagawa, T., and Hamano, Y. (2012) A stand-alone adenylation domain forms amide bonds in streptothricin biosynthesis, *Nature Chemical Biology* **8**, 791-797.

89. Zhang, X., Alemany, L. B., Fiedler, H. P., Goodfellow, M., and Parry, R. J. (2008) Biosynthetic investigations of lactonamycin and lactonamycin z: cloning of the biosynthetic gene clusters and discovery of an unusual starter unit, *Antimicrobial Agents and Chemotherapy* **52**, 574-585.

90. Bao, W., Sheldon, P. J., Wendt-Pienkowski, E., and Hutchinson, C. R. (1999) The Streptomyces peucetius dpsC gene determines the choice of starter unit in biosynthesis of the daunorubicin polyketide, *Journal Bacteriology* **181**, 4690-4695.

91. Izumikawa, M., Cheng, Q., and Moore, B. S. (2006) Priming type II polyketide synthases via a type II nonribosomal peptide synthetase mechanism, *Journal of the American Chemical Society* **128**, 1428-1429.

92. Tang, Y., Koppisch, A. T., and Khosla, C. (2004) The acyltransferase homologue from the initiation module of the R1128 polyketide synthase is an acyl-ACP thioesterase that edits acetyl primer units, *Biochemistry* **43**, 9546-9555.

ACKNOWLEDGEMENT

This dissertation would not have been possible without the guidance and the help of many individuals who in one way or another contributed and extended their valuable assistance in the preparation and completion of this study.

First and foremost, my utmost gratitude to Professor Makoto Nishiyama of the Laboratory of Cell Biotechnology, Biotechnology Research Center, The University of Tokyo for allowing me to join this great laboratory and also his guidance, comments, and ideas.

Dr. Tomohisa Kuzuyama, Associate Professor of the Laboratory of Cell Biotechnology, Biotechnology Research Center, The University of Tokyo for his patience and dedication in guiding me throughout my study. Dr. Kuzuyama has been my inspiration as I hurdle all the obstacles in the completion of this research work.

Dr. Saori Kosono, Associate Professor of The University of Tokyo for her creative ideas and help throughout my study.

Dr. Takeo Tomita, Dr. Ayako Yoshida, and Mr. Taro Shiraishi, Assistant Professors of The University of Tokyo for all their supports and advice in my research.

Dr. Kazuo Furihata, Assistant Professor of The University of Tokyo for all his professional advice regarding NMR analysis.

Dr. Masahiro Tanaka of Daiichi-Sankyo Co. Ltd. and Dr. Kazuo Shin-ya of National Institute of Advance Industrial Science and Technology (AIST) for providing some of the actinomycetes strains used in this study. Dr. Miho Izumikawa also from AIST for kindly conducting the bioactivity assays

Dr. Susumu Okamoto of National Food Research Institute, Tsukuba for his professional advice regarding rifampicin-resistant mutant screening.

My colleagues in the Laboratory of Cell Biotechnology for their unfailing help, kindness, and laughter. A note of appreciation goes to Dr. Taro Ozaki, Dr. Jin-Soo Park, Takuya Hashimoto and Intan Timur Maisyarah for their suggestions and advice on countless matters in both my research and also personal life. It was a great experience being in this laboratory.

Last but not least, my heartiest appreciation to my mum Chin Sui Ching, my little bro Thong Yee Han and my fiancé Chu Ket Pin for their never-ending loves and supports throughout my study career. I love you guys! ☺